Table of Contents

SYLLABUS

Course Catalog

3 Credit hours (3 h lectures). Overview of wireless communications. Cellular systems: principles, trunking, grade of service and traffic capacity, power control, and handovers. Characterization of wireless channels: large scale and small scale propagation mechanisms, path loss, multipath and fading. Digital modulation techniques for wireless channels. Power efficiency, nonlinear amplifiers, diversity. Performance in multipath fading channels. Multiple access: fixed (FDMA, TDMA, CDMA) and random (ALOHA, CSMA) access methods.

Textbook

Theodore S. Rappaport (2002). Wireless Communications: Principles and Practice, 2nd ed. Printice Hall.

References

1. Gordon L. Stuber, *Principles of Mobile Communication,* Springer, 2017
2. Vijay Kumar Garg, *Wireless communications and networking,* Elsevier, 2007

Instructor

Instructor: Dr. Mohammad M. Banat

Email Address: banat@just.edu.jo

Prerequisites

Prerequisites by topic Digital Communications

Prerequisites by course EE 551

Topics Covered

Week	Topics	Chepters in Text
1	Introduction to Wireless communication systems	1
2-4	Cellular Systems Design Fundamentals	2
5-7	Mobile Radio Propagation, Fading and Multipath	3
8-9	Modulation techniques for mobile radio systems	4
10-11	Diversity techniques for mobile radio systems	5-9
12-13	Multiple access techniques for mobile systems	10
14-15	Mobile Systems and Standards	11
16	Next Generation Wireless Communications	Handouts

Objectives and Outcomes

Objectives	Outcomes
1. Ability to understand and apply cellular system components and design fundamentals. [1,2]	1.1. Understanding the cellular conept [1,2]. 1.2. Ability to apply understanding of cellular concept in solving traffic and coverage problems. [1,2]
2. Ability to understand models and characterizations of wireless channels. [1,2]	2.1. Understanding wireless channel characteristics. [1,2] 2.2. Ability to understand and apply mathematical models of wireless channels. [1,2] 2.3. Ability to use the gained knowledge in analyzing the performance of digital modulation techniques over wireless channels. [1,2]
3. Acquiring working knowledge of multiple access techniques for mobile systems. [1]	3.1. Understanding the principles and different types of multiple access. [1] 3.2. Understanding the use of multiple access techniques in wireless communications. [1]
4. Acquiring working knowledge of mobile systems and their standards. [1,3]	4.1. Understanding the role of standardization in wireless communications. [1] 4.2. Conducting a research on a topic related to mobile/wireless systems, write a technical report, and give an oral presentation on the results of the research through term projects. [3]

Evaluation

Assessment Tool	Expected Due Date	Weight
Mid-Term Exam	10 May 2022	20%
Project Report	19 May 2022	12%
Project Presentations	26 May 2021	8%
Class Work		10%
Final Exam		50%

Contribution of Course to Meeting the Professional Component

The course contributes to prepare engineers who are as up-to-date as possible on current and emerging wireless communication network technologies.

I. INTRODUCTION

The ability to communicate with people on the move has evolved remarkably since 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. 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 technologies which make portable radio equipment smaller, cheaper, and more reliable.

Commercial wireless systems and services have undergone rapid development and deployment since first generation cellular telephone systems were introduced in the early 1980s. These first generation (1G) cellular telephone systems were based on analog frequency modulation (FM) technology, and were designed to carry narrow-band circuit-switched voice services[1]. The first generation cellular service providers witnessed an exponential growth rate in their subscriptions, and by the late 1980s capacity limits were already reached in the largest markets with 1G cellular systems. In response to such heavy demand, second generation (2G) digital cellular systems were developed and introduced in the early 1990s. These 2G cellular systems were based on either time division multiple access (TDMA) or code division multiple access (CDMA) technologies, and were initially designed to carry circuit-switched voice and data. During the 1990s, these 2G systems were enhanced to provide packet-switched[2] data in addition to circuit-switched voice. These transitional 2G cellular systems with their enhanced data transmission capabilities later became known as 2.5G systems. Third generation (3G) cellular systems were introduced after the year 2000, and they allowed simultaneous use of speech and data services and still higher data rates. These higher data rate capabilities were supplemented by geolocation information, giving rise to location dependent services. Fourth generation (4G) cellular systems use voice over Internet Protocol (VoIP) and multimedia applications with broadband access. These 4G systems are based on multicarrier modulation/multiplexing techniques such as orthogonal frequency division multiple access (OFDMA), or advanced single-carrier modulation/multiplexing techniques such as single-carrier frequency division multiple access (SC-FDMA). Fifth generation (5G) wireless systems are currently in their initial deployment stages.

[1] Circuit switching is a type of network in which a physical path is dedicated to a single connection between two end-points in the network for the duration of the connection. Ordinary voice phone service is circuit-switched. The telephone company reserves a specific physical path to the number you are calling for the duration of your call. During that time, no one else can use the physical lines involved.

[2] Packet switching is a method of grouping data that is transmitted over a digital network into packets. Packets are made of a header and a payload. Data in the header is used by networking hardware to direct the packet to its destination, where the payload is extracted and used by application software. Packet switching is the primary basis for data communications in computer networks worldwide.

I.1. Evolution of Wireless Systems and Standards

I.1.A. First Generation (1G) Cellular Systems

The early 1970s saw the emergence of the radio technology that was needed for the deployment of mobile radio systems in the 800/900MHz band at a reasonable cost. In the early 1980s, many countries deployed incompatible first generation (1G) cellular systems based on frequency division multiple access (FDMA) and analog FM technology. 1G systems were designed to carry narrow-band circuit switched voice services. With FDMA there is a single traffic channel per radio frequency carrier. When a user accesses the network two carriers (channels) are actually assigned, one for the forward (base-to-mobile) link and one for the reverse (mobile-to-base) link. Separation of the forward and reverse carrier frequencies is necessary to allow implementation of a duplexer, an arrangement of filters that isolates the forward and reverse link channels, thus preventing a radio transceiver from jamming itself.

In 1979, the first analog cellular system, the Nippon Telephone and Telegraph (NTT) system, became operational. In 1981, Ericsson Radio Systems AB fielded the Nordic Mobile Telephone (NMT) 900 system, and in 1983 AT&T fielded the Advanced Mobile Phone Service (AMPS) as a trial in Chicago, IL. Many other first generation analog systems were also deployed in the early 1980s including TACS, ETACS, NMT 450, C-450, RTMS, and Radiocom 2000 in Europe, and JTACS/NTACS in Japan.

All 1G cellular systems are now extinct.

I.1.B. Second Generation (2G) Cellular Systems

European countries saw the deployment of incompatible 1G cellular systems that did not admit roaming throughout Europe. As a result, the Conference of European Postal and Telecommunications Administrations (CEPT) established Groupe Special Mobile (GSM) in 1982 with the mandate of defining standards for future Pan-European cellular radio systems.

Second generation (2G) digital cellular systems were developed and introduced in the early 1990s. These included the GSM/DCS1800/PCS1900 standard in Europe, the Personal Digital Cellular (PDC) standard in Japan, and the IS-54/136 and IS-95 standards in the USA.

2G cellular systems were based on either time division multiple access (TDMA) or code division multiple access (CDMA) technologies. They were initially designed to carry circuit-switched voice and data. During the 1990s, these 2G systems were enhanced to provide packet-switched data in addition to circuit-switched voice. These transitional 2G cellular systems with their enhanced data transmission capabilities later became known as 2.5G systems.

The GSM system (now "Global System for Mobile Communications") was developed to operate in a new frequency allocation, and made improved quality, Pan-European roaming, and the support of data services its primary objectives. GSM was deployed in late 1992 as the world's first digital cellular system. GSM can support data services at 2.4, 4.8, and 9.6 kb/s. GSM uses TDMA with 200 kHz carrier spacings.

Variants of GSM have also been developed to operate in the 900MHz and 1,800MHz bands in Europe, and the 850MHz and 1,900MHz bands in North America. GSM has been a phenomenal success. In late 1993, over a million subscribers were using GSM phone networks. By mid 2010, GSM had over 4.3 billion subscribers across more than 212 countries and territories.

Newer versions of the standard have been developed that are backward compatible with the original GSM system. GSM Release '97 added packet data capabilities. This enhancement provides data rates up to 140 kb/s and is called General Packet Radio Service (GPRS). GSM Release '99 introduced higher speed data transmission using a higher-level 8-PSK modulation format. This enhancement is called Enhanced Data Rates for GSM Evolution (EDGE). GPRS and EDGE are generally branded as 2.5G systems.

In North America, the primary driver for second generation systems was the capacity limit felt by some AMPS operators in the largest US markets by the late 1980s. One of the key objectives established by the Cellular Telephone Industry Association (CTIA) at that time was a tenfold increase in capacity over AMPS. Furthermore, since AMPS was already deployed extensively throughout North America, it was desirable that any second generation cellular system be reverse compatible with AMPS. This eventually lead to the development of dual-mode cellular standards in North America.

While Europe saw a convergence to the GSM standard based on TDMA technology, North America saw a divergence to two second generation digital cellular standards, IS-54/136 and IS-95, based on TDMA and CDMA technology, respectively. The IS-54 standard, adopted in 1990, was based on TDMA with 30 kHz carrier spacings (the same as AMPS) and $\pi/4$ phase-shifted quadrature differential phase shift keyed ($\pi/4$-DQPSK) modulation with a raw bit rate of 48.6 kb/s.

Just after the CTIA adopted IS-54 in 1990, another second generation digital cellular standard was proposed by Qualcomm based on CDMA technology. In March 1992, CDMA was adopted as the IS-95 standard. The introduction of IS-95 saw considerable debate and spirited exchanges over the relative capacity and merits of TDMA and CDMA cellular systems. Initial capacity claims for IS-95 were 40 times AMPS. However, commercial deployments eventually realized a capacity gain of six to ten times AMPS. The introduction of IS-95 CDMA cellular was of historical significance, because 3G cellular systems are based on CDMA technology.

With IS-95, the basic user data rate is 9.6 kb/s for Rate Set 1 (RS1) and 14.4 kb/s for Rate Set 2 (RS2), which is spread using pseudonoise (PN) sequence with a chip rate of 1.2288 Mchips/s. The forward channel supports coherent detection using an unmodulated pilot channel for channel estimation. Information on the forward link is encoded using a rate-1/2 convolutional code, interleaved, spread using one of 64 orthogonal Walsh codes, and transmitted in 20 ms bursts. Each MS in a cell is assigned a different Walsh code, thus providing complete orthogonality under ideal channel conditions. Final spreading with a PN code of length 2^{15}, having a phase offset that depends on each BS, is used to mitigate the multiple access interference to and from other cells.

The information on the IS-95 reverse link is encoded using a rate-1/3 convolutional code, interleaved, and mapped onto one of 64 Walsh codes. Unlike the IS-95 forward channel that uses Walsh codes for spreading, the reverse link uses Walsh codes for 64-ary orthogonal modulation. The BS receiver uses noncoherent detection, since no pilot signal is transmitted on the reverse link. Both the BSs and the MSs use RAKE receivers to provide multipath diversity. To ensure that the power control algorithm is stable, CDMA cellular systems must use soft handoff, where the MS maintains a radio link with multiple BSs when traversing between cells.

I.1.C. Third Generation (3G) Cellular Systems

In March 1992, the World Allocation Radio Conference (WARC) approved a worldwide spectral allocation in support of IMT-2000 (International Mobile Telephone by the Year 2000) in the 1,885–2,200MHz band. The IMT-2000 standard was developed by the International Telecommunications Union Radio Communications (ITU-R) and Telecommunications (ITU-T) sectors. Various standards bodies around the world have provided inputs to the IMT-2000 standard definition. IMT-2000 was envisioned as a ubiquitous wireless system that could support voice, multimedia, and high-speed data communication. The ITU provided no clear definition of the minimum or average rates to expect from 3G equipment or providers. However, it was generally expected that 3G networks would provide a minimum downlink peak data rate of 2 Mbit/s for stationary or walking users, and 384 kbit/s for users in moving vehicles.

IMT-2000 is actually a family of standards. Two of the standards, namely EDGE and Digital Enhanced Cordless Telephone (DECT), are based on TDMA. While the EDGE standard fulfils the requirements for IMT-2000, EDGE networks are typically branded as 2.5G networks rather than 3G networks. The most predominant forms of IMT-2000 are cdma2000 developed by 3GPP2 and the Universal Mobile Telecommunications System (UMTS) family of standards, which includes Wideband Code Division Multiple Access (WCDMA), developed by 3GPP. Sometimes WCDMA is used synonymously with UMTS. Mobile WiMAX (Worldwide Interoperability for Microwave Access), developed by the IEEE802.16 working group, is also included under the IMT-2000 umbrella as a 3.5G standard. WiMAX is a multicarrier scheme based on OFDMA.

Third generation (3G) cellular systems were introduced after the year 2000. They allowed simultaneous use of speech and data services and still higher data rates. Most wireless providers deployed 3G networks that operate using one of two different standards:

- One is GSM/GPRS/EDGE/WCDMA/HSPA which has been developed by the Third Generation Partnership Project (3GPP) and accounts for roughly 80% of the global market.
- The other is IS-95A/B/cdma20001x/cdma2000EV-DO, which has been developed by the Third Generation Partnership Project 2 (3GPP2) and accounts for the remaining 20% of the global market.

The 3G standard HSPA+ has a peak downlink speed of 21 Mbps in 5MHz bandwidth based on single-carrier TDM/CDMA technology.

I.1.D. Fourth Generation (4G) Cellular Systems

4G cellular systems use voice over Internet Protocol (VoIP) and multimedia applications with ultra-broadband (gigabit peak speed) access. Most 4G systems are based on multicarrier modulation/multiplexing techniques such as orthogonal frequency division multiple access (OFDMA), or advanced single-carrier modulation/multiplexing techniques such as single-carrier frequency division multiple access (SC-FDMA).

Currently, fourth generation (4G) cellular systems are commercially available, known as Long-Term Evolution (LTE) and Long-Term Evolution – Advanced (LTE-A). Unlike the 3G cellular systems that are based on CDMA technology, the 4G cellular system proposals are based on OFDMA and SC-FDMA technology.

I.2. Overview of Mobile Systems

Wireless data network users are not confined to the locations of "wired" data jacks, and enjoy connectivity that is less restrictive and therefore well suited to meet the needs of today's mobile users.

The cellular system employs a different design approach to what is used in most commercial radio and television systems. Radio and television systems typically operate at maximum power and with the tallest antennas allowed by the regulatory agency of the country. In the cellular system, the service area is divided into cells. A transmitter is designed to serve an individual cell. The system seeks to make efficient use of available channels by using low-power transmitters to allow frequency reuse at much smaller distances. Maximizing the number of times each channel can be reused in a given geographic area is the key to an efficient cellular system design.

Examples of wireless systems include wireless wide-area networks (WWAN) [i.e., cellular systems], wireless local area networks (WLAN), and wireless personal area networks (WPAN). The handsets used in all of these systems possess complex functionality, yet they have become small, low-power consuming devices that are mass produced at low cost, which has in turn accelerated their widespread use. The recent advancements in Internet technology have increased network traffic considerably, resulting in a rapid growth of data rates. This phenomenon has also had an impact on mobile systems, resulting in the extraordinary growth of the mobile Internet.

As mobile networks evolve to offer both circuit and packet-switched services, users will be connected permanently (always on) via their personal terminal of choice to the network. Transmission rates will be increased and there will be far more efficient use of network bandwidth and resources. As the number of IP-based mobile applications grows, future systems will offer flexible access technology because it allows for mobile, office, and residential use in a wide range of public and nonpublic networks.

I.2.A. FDD AND TDD

Frequency division duplexing (FDD) provides simultaneous radio transmission channels for the subscriber and the base station, so that they both may constantly transmit while simultaneously receiving signals from one another. At the base station, separate transmit and receive antennas are used to accommodate the two separate channels. 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 inside the subscriber unit to enable the same antenna to be used for simultaneous transmission and reception. To facilitate FDD, it is necessary to separate the transmit and receive frequencies by about 5% of the nominal RF frequency, so that the duplexer can provide sufficient isolation while being inexpensively manufactured.

In FDD, a pair of simplex channels with a fixed and known frequency separation is used to define a specific radio channel in the system. The channel used to convey traffic to the mobile user from a base station is called the forward channel (or downlink), while the channel used to carry traffic from the mobile user to a base station is called the reverse channel (or uplink). Full duplex mobile radio systems provide many of the capabilities of the standard telephone, with the added convenience of mobility. FDD is used exclusively in analog mobile radio systems.

Time division duplexing (TDD) uses the fact 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, and the remaining time is used to transmit from the mobile to the base station. TDD is only possible with digital transmission formats and digital modulation, and is very sensitive to timing. It is for this reason that TDD has only recently been used, and only for indoor or small area wireless applications where the physical coverage distances (and thus the radio propagation time delays) are much smaller than the many kilometers used in conventional cellular telephone systems.

- In frequency division duplex (FDD) mode the bandwidth is split into two separate parts: one for the uplink and one for the downlink.
- In time-division duplex (TDD) mode the whole bandwidth is used for both downlink and uplink transmission, but separated in time.

TDD is the preferable mode in next generation wireless systems.

I.2.B. CELLULAR TELEPHONE SYSTEMS

A cellular telephone system provides a wireless connection to the public switched telephone network (PSTN) for any user location within the radio range of the system. Cellular systems accommodate a large number of users over a large geographic area, within a limited frequency spectrum. Cellular radio systems provide high quality service that is often comparable to that of landline telephone systems. High capacity is achieved 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. A sophisticated switching technique called a handoff enables a call to proceed uninterrupted when the user moves from one cell to another.

A basic cellular system consists of mobile stations (MS's), base stations (BS's) and a mobile switching center (MSC). The Mobile Switching Center is sometimes called a mobile telephone switching office (MTSO), since it is responsible for connecting all mobiles to the PSTN in a cellular system. Each mobile communicates via radio with one of the base stations and may be handed off to any number of base stations throughout the duration of a call. The MS 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 simultaneously 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 simultaneous mobile calls via telephone lines or microwave links to the MSC. The MSC coordinates the activities of all of the base stations and connects the entire cellular system to the PSTN. A typical, MSC handles 100,000 cellular subscribers and 5,000 simultaneous conversations at a time, and accommodates all billing and system maintenance functions, as well. In large cities, several MSCs are used by a single carrier.

Communication between the base station and the mobiles is defined by a standard common air interface (CAl) that specifies four different channels. The channels used for voice transmission from the base station to mobiles are called forward voice channels (FVC's) and the channels used for voice transmission from mobiles to the base station are called reverse voice channels (RVC's). The two channels responsible for initiating mobile calls are the forward control channels (FCC's) and reverse control channels (RCC's). Control channels are often called setup channels because they are only involved in setting up a call and moving it to an unused voice channel. Control channels transmit and receive data messages that carry call initiation and service requests, and are monitored by mobiles when they do not have a call in progress.

How a Cellular Telephone Call is Made

When a cellular phone is turned on, but is not yet engaged in a call, it first scans the group of forward control channels to determine the one with the strongest signal, and then monitors that control channel until the signal drops below a usable level. At this point it again scans the control channels in search of the strongest base station signal. The control channels are defined and standardized over the entire geographic area covered and typically make up about 5% of the total number of channels available in the system (the other 95% are dedicated to voice and data traffic for the end-users).

Since the control channels are standardized and are identical throughout different markets within the country or continent, every phone scans the same channels while idle. When a telephone call is placed to a mobile user, the MSC dispatches the request to all base stations in the cellular system. The mobile identification number (MIN), which is the subscriber's telephone number, is then broadcast as a paging message over all of the forward control channels throughout the cellular system. The mobile receives the paging message sent by the base station which it monitors, and responds by identifying itself over the reverse control channel. The base station relays the acknowledgment sent by the mobile and informs the MSC of the handshake. Then, the MSC instructs the base station to move the call to an unused voice channel within the cell (typically between ten to sixty voice channels and just one control channel are used in each base station). At this point the base station signals the mobile to change frequencies to an unused forward and reverse voice channel pair, at which point another data message (called an alert) is transmitted over the forward voice channel to instruct the mobile telephone to ring, thereby instructing the mobile user to answer the phone. All of these events occur within a few seconds and are not noticeable by the user.

Once a call is in progress, the MSC adjusts the transmitted power of the mobile and changes the channel of the mobile unit and base stations in order to maintain call quality as the subscriber moves in and out of range of each base station. This is called a handoff. Special control signaling is applied to the voice channels so that the mobile unit may be controlled by the base station and the MSC while a call is in progress.

When a mobile originates a call, a call initiation request is sent on the reverse control channel. With this request the mobile unit transmits its telephone number (MIN), electronic serial number (ESN), and the telephone number of the called party. The mobile also transmits a station class mark (SCM) which indicates what the maximum transmitter power level is for the particular user. The cell base station receives this data and sends it to the MSC. The MSC validates the request, makes connection to the called party through the PSTN, and instructs the base station and mobile user to move to an unused forward and reverse voice channel pair to allow the conversation to begin.

All cellular systems provide a service called roaming. This allows subscribers to operate in service areas other than the one from which service is subscribed. When a mobile enters a city or geographic area that is different from its home service area, it is registered as a roamer in the new service area. This is accomplished over the FCC, since each roamer is camped on to a FCC at all times. Every several minutes, the MSC issues a global command over each FCC in the system, asking for all mobiles which are previously unregistered to report their MIN and ESN over the RCC. New unregistered mobiles in the system periodically report back their subscriber information upon receiving the registration request, and the MSC then uses the MIN/ESN data to request billing status from the home location register (HLR) for each roaming mobile. If a particular roamer has roaming authorization for billing purposes, the MSC registers the subscriber as a valid roamer. Once registered, roaming mobiles are allowed to receive and place calls from that area, and billing is routed automatically to the subscriber's home service provider.

***

II. THE CELLULAR CONCEPT

II.1. Introduction

The design objective of early mobile radio systems was to achieve a large coverage area by using a single, high powered transmitter with an antenna mounted on a tall tower. While this approach achieved very good coverage, it also meant that it was impossible to reuse those same frequencies throughout the system, since any attempts to achieve frequency reuse would result in interference. Faced with the fact that government regulatory agencies could not make spectrum allocations in proportion to the increasing demand for mobile services, it became imperative to restructure the radio telephone system to achieve high capacity with limited radio spectrum, while at the same time covering very large areas.

The cellular concept was a major breakthrough in solving the problem of spectral congestion and user capacity. It offered very high capacity in a limited spectrum allocation without any major 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 transmitters (small cells) each providing 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 stations are assigned different groups of channels so that all the available channels are assigned to a relatively small number of neighboring base stations. Neighboring base stations are assigned different groups of channels so that the interference between base stations (and the mobile users under their control) is minimized. By systematically spacing base stations and their channel groups, the available channels are distributed throughout the geographic region and may be reused as many times as necessary, so long as the interference between co-channel stations is kept below acceptable levels.

As the demand for service increases (i.e., as more channels are needed within a particular market), the number of base stations may be increased (along with a corresponding decrease in transmitter power to avoid added interference), thereby providing additional radio capacity with no additional increase in radio spectrum. This fundamental principle is the foundation for all modern wireless communication systems, since it enables a fixed number of channels to serve an arbitrarily large number of subscribers by reusing the channels throughout the coverage region. Furthermore, the cellular concept allows every piece of subscriber equipment within a country or continent to be manufactured with the same set of channels, so that any mobile may be used anywhere within the region.

II.2. Frequency Reuse

Cellular radio systems rely on an intelligent allocation and reuse of channels throughout a coverage region. Each cellular base station is allocated a group of radio channels to be used within a small geographic area called a cell. Base stations in adjacent cells are assigned channel groups which contain completely different channels than neighboring cells. The base station antennas are designed to achieve the desired coverage within the particular cell by limiting the coverage area to within the boundaries of a cell.

Frequency reuse in cellular systems creates co-channel interference (CCI), which limits the cellular spectral efficiency. When CCI is generated, the power spectral densities of desired and interfering signals overlap. The group of channels used in a cell may be used in other cells that are separated from one another by distances large enough to keep CCI levels within tolerable limits. The process of selecting and allocating channel groups for all of the cellular base stations within a system is called frequency reuse or frequency planning.

Frequency reuse also introduces adjacent channel interference (ACI). This type of interference arises when neighboring cells use carrier frequencies that are spectrally adjacent to each other. In this case, the power spectral densities of the desired and interfering signals overlap partially. Since adjacent frequencies are used at close distances, the interference can be significant. As a result, the transmit power is regulated to fit within certain upper limits.

Figure II.1 illustrates cellular frequency reuse, where cells labeled with the same letter use the same group of channels. The hexagonal cell shape is conceptual, and is a simplified model of the radio coverage of each base station. However, it has been universally adopted since the hexagon permits easy and manageable analysis of cellular systems. The actual radio coverage of a cell is called the footprint and is determined from field measurements or propagation prediction models.

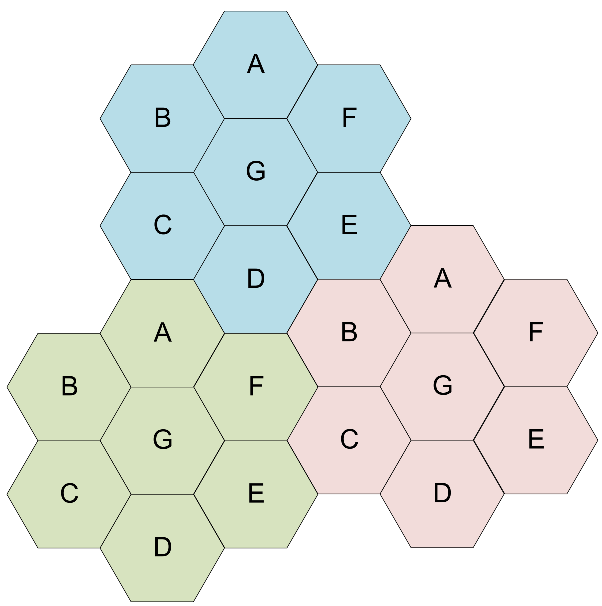

Figure II.1: Cellular frequency reuse

While it might appear to be natural to choose a circle to represent the coverage area of a base station, adjacent circles can not be overlaid upon a map without leaving gaps or creating overlapping regions. Thus, when considering geometric shapes which cover an entire region without overlap and with equal area, there are three sensible choices: a square; an equilateral triangle; and a hexagon. A cell must be designed to serve the weakest mobiles within the footprint, and these are typically located at the edge of the cell. For a given distance between the center of a polygon and its farthest perimeter points, the hexagon has the largest area of the three. Thus, by using the hexagon geometry, the fewest number of cells can cover a geographic region, and the hexagon closely approximates a circular radiation pattern which would occur for an omni-directional base station antenna and free space propagation. Of course, the actual cellular footprint is determined by the contour in which a given transmitter serves the mobiles successfully.

When using hexagons to model coverage areas, base station transmitters are depicted as either being in the center of the cell (center-excited cells) or on three of the six cell vertices (edge-excited cells). Normally, omni-directional antennas are used in center-excited cells and sectored directional antennas are used in corner-excited cells. Practical considerations usually do not allow base stations to be placed exactly as they appear in the hexagonal layout. Most system designs permit a base station to be positioned up to one-fourth the cell radius away from the ideal location.

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 divided 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

$$S = kN \qquad \text{(II.1)}$$

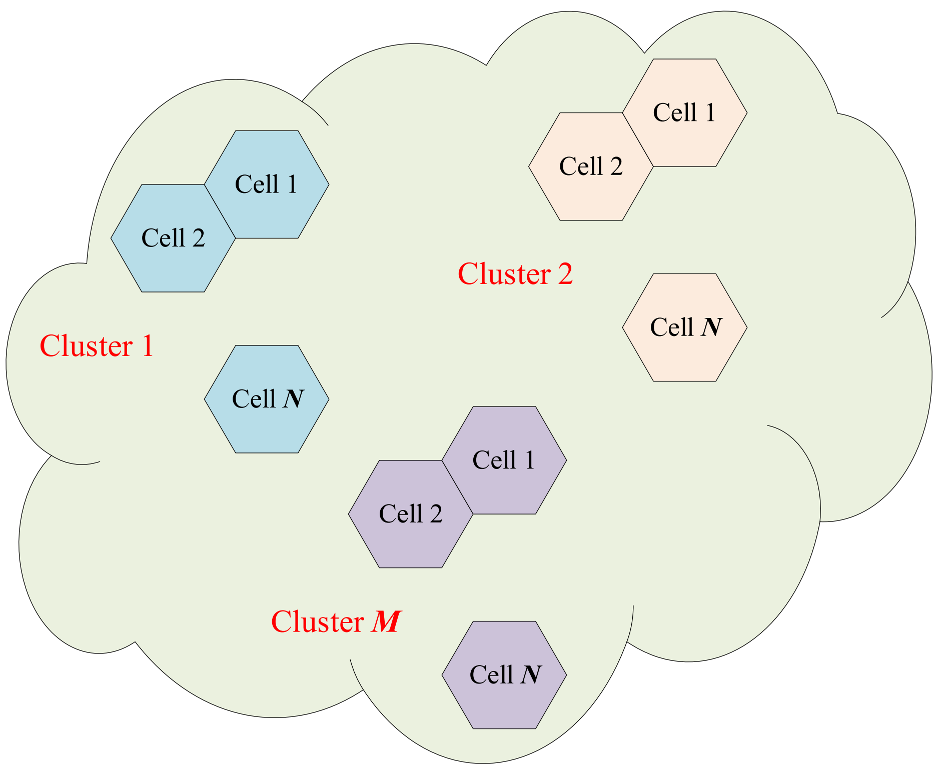

Figure II.2: Coverage area, clusters and cells

II.2-Frequency Reuse

The N cells which collectively use the complete set of available frequencies form what is called a cluster. If a cluster is replicated M times within the system, the total number of cells in the coverage area is equal to

$$N_c = MN \tag{II.2}$$

The total number of channels C can be used as a measure of capacity and is given by

$$\begin{aligned} C &= MS \\ &= kN_c \\ &= kMN \end{aligned} \tag{II.3}$$

As seen from (II.3), the capacity of a cellular system is directly proportional to the number of times a cluster is replicated in a fixed service area. The factor N is called the cluster size and is typically equal to 4, 7, or 12. If the cluster size N is reduced while the cell area is kept constant, more clusters are required to cover a given area and hence more capacity (a larger value of C) is achieved.

A large cluster size indicates that the ratio of the distance between co-channel cells to the cell radius is large. Conversely, a small cluster size indicates that co-channel cells are located much closer together. The value of N is a function of how much interference a mobile or base station can tolerate while maintaining a sufficient quality of communications. From a design viewpoint, the smallest possible value of N is desirable in order to maximize capacity over a given coverage area (i.e., to maximize C in (II.3)).

To understand how C and N are related, assume the overall coverage area is fixed, and is equal to A. Then, using M clusters, means that each cluster had an area of

$$A_{\text{cluster}} = \frac{A}{M} \tag{II.4}$$

If each cluster consists of N cells, then the cell area is equal to

$$A_{\text{cell}} = \frac{A_{\text{cluster}}}{N} \tag{II.5}$$

Equation (II.5) can be rewritten as

$$A_{\text{cluster}} = NA_{\text{cell}} \tag{II.6}$$

Let's assume the cell area is fixed, then decreasing the number of cells per cluster implies decreasing the cluster area. From (II.4), this mean the total number of clusters M goes up. From the first line of (II.3), the capacity C also goes up. In summary, with a fixed cell area, C is inversely proportional to N, provided that the coverage area and the number of channels in the system are fixed.

The frequency reuse factor of a cellular system is given by $1/N$, since each cell within a cluster is only assigned $1/N$ of the total available channels in the system.

Due to the fact that the hexagonal geometry of Figure II.1 has exactly six equidistant neighbors and that the lines joining the centers of any cell and each of its neighbors are separated by multiples of 60 degrees, there are only certain cluster sizes and cell layouts which are possible. To connect without gaps between adjacent cells, the geometry of hexagons is such that the number of cells per cluster N can only have values which satisfy

$$N = i^2 + ij + j^2 \quad \text{(II.7)}$$

where i and j are non-negative integers ($i = 0$ or $j = 0$ is allowed).

To find the nearest co-channel neighbors of a particular cell, one must do the following:

(1) move i cells along any chain of hexagons and then
(2) turn 60 degrees counter-clockwise and move j cells.

This is illustrated in Figure 2.2 for $i = 3$ and $j = 2$ ($N = 19$).

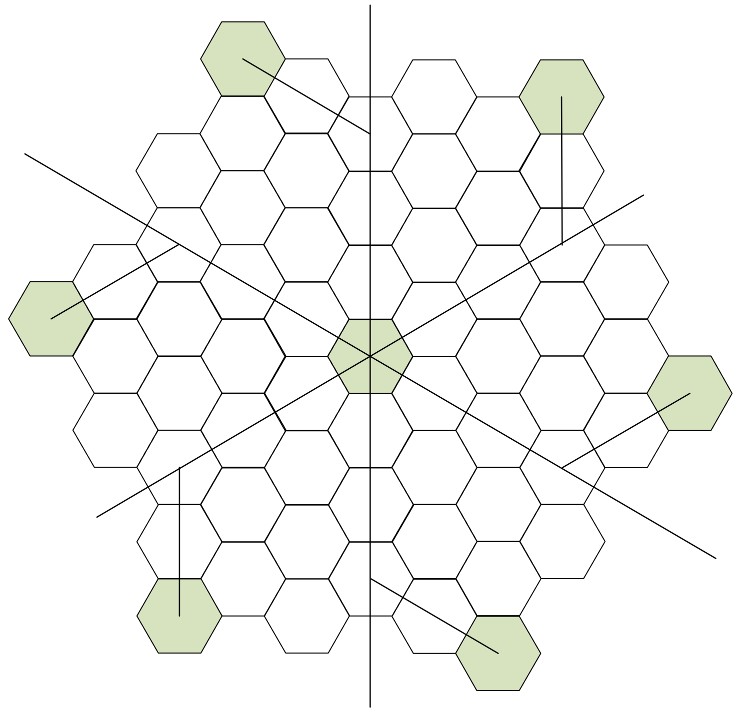

Figure II.3: Method of locating co-channel cells in a cellular system

Example II-1

If a total of 33 MHz is allocated to an FDD cellular 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) 4-ce1l reuse
(b) 7-ce1l reuse
(c) 12-ce1l reuse
(d) 19-cell reuse

If 1 MHz of the allocated spectrum is dedicated to control channels, determine an equitable distribution of control channels and voice channels in each cell for each of the three systems.

Solution

Channel bandwidth = 2×25 = 50 kHz.

Number of available channels = $\frac{33,000}{50}$ = 660.

Number of available channels per cell

(a) $\frac{660}{4} = 165$

(b) $\frac{660}{7} \approx 95$

(c) $\frac{660}{12} = 55$

A 1 MHz spectrum for control channels implies that there are 1000/50 = 20 control channels out of the 660 channels available. To evenly distribute the control and voice channels, simply allocate the same number of channels in each cell wherever possible. Here, the 660 channels must be evenly distributed to each cell within the cluster. In practice, only the 640 voice channels would be allocated, since the control channels are allocated separately as 1 per cell.

(d) For $N=4$, we can have 5 control channels and 160 voice channels per cell. In practice, however, each cell only needs a single control channel (the control channels have a greater reuse distance than the voice channels). Thus, one control channel and 160 voice channels would be assigned to each cell.

(e) For $N=7$, 4 cells with 3 control channels and 92 voice channels, 2 cells with 3 control channels and 90 voice channels, and 1 cell with 2 control channels and 92 voice channels could be allocated. In practice, however, each cell would have one control channel, four cells would have 91 voice channels, and three cells would have 92 voice channels.

(f) For $N=12$, we can have 8 cells with 2 control channels and 53 voice channels, and 4 cells with 1 control channel and 54 voice channels each. In an actual system, each cell would have 1 control channel, 8 cells would have 53 voice channels, and 4 cells would have 54 voice channels.

(g) To be done by students.

For efficient utilization of the radio spectrum, a frequency reuse scheme that is consistent with the objectives of increasing capacity and minimizing interference is required. A variety of channel assignment strategies have been developed to achieve these objectives. Channel assignment strategies can be classified as either fixed or dynamic. 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.

In a fixed channel assignment strategy, each cell is allocated a predetermined set of voice channels. Any call attempt within the cell can only be served by the unused channels in that particular cell. If all the channels in that cell are occupied, the call is blocked and the subscriber does not receive service. Several variations of the fixed assignment strategy exist. In one approach, called the borrowing strategy, a cell is allowed to borrow channels from a neighboring cell if all of its own channels are already occupied. The mobile switching center (MSC) supervises such borrowing procedures and ensures that the borrowing of a channel does not disrupt or interfere with any of the calls in progress in the donor cell.

In a dynamic channel assignment strategy, voice channels are not allocated to different cells permanently. Instead, each time a call request is made, the serving base station requests a channel from the MSC. The switch then allocates a channel to the requested cell following an algorithm that takes into account the likelihood of future blocking within the cell, the frequency of use of the candidate channel, the reuse distance of the channel, and other cost functions.

Accordingly, the MSC only allocates a given frequency if that frequency is not presently in use in the cell or any other cell which falls within the minimum restricted distance of frequency reuse to avoid co-channel interference. Dynamic channel assignment reduces the likelihood of blocking, which increases the trunking capacity of the system. Since all the available channels are accessible to all of the cells. 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 increases the storage and computational load on the system but provides the advantage of increased channel utilization and decreased probability of a blocked call.

II.3. Handoff Strategies

II.3.A. Handoff Concept

When a mobile moves into a different cell while a conversation is in progress, the MSC automatically transfers the call to a new channel belonging to the new base station. This handoff operation not only involves identifying a new base station, but also requires that the voice and control signals be allocated to channels associated with the new base station.

Processing handoffs is an important task in any cellular radio system. Many handoff strategies prioritize handoff requests over call initiation requests when allocating unused channels in a cell site. Handoffs must be performed successfully and as infrequently as possible, and be imperceptible to the users. In order to meet these requirements, system designers must specify an optimum signal level at which to initiate a handoff.

Once a particular signal level $P_{r,\text{usable}}$ is specified as the minimum usable signal for acceptable voice quality at the base station receiver (normally taken as between -90 dBm and -100 dBm), a slightly stronger signal level $P_{r,\text{handoff}}$ is used as a threshold at which a handoff is made. Let's define the margin

$$\Delta = P_{r,\text{handoff}} - P_{r,\text{usable}} \tag{II.8}$$

If Δ is too large, unnecessary handoffs which burden the MSC may occur, and if Δ is too small, there may be insufficient time to complete a handoff before a call is lost due to weak signal conditions. Therefore, Δ is chosen carefully to meet these conflicting requirements. Figure II.4 illustrates a handoff situation. Upper part of Figure II.4 demonstrates the case where a handoff is not made and the signal drops below the minimum acceptable level to keep the channel active. This dropped call event can happen when there is an excessive delay by the MSC in assigning a hand off, or when the threshold Δ is set too small for the handoff time in the system. Excessive delays may occur during high traffic conditions due to computational loading at the MSC or due to the fact that no channels are available on any of the nearby base stations (thus forcing the MSC to wait until a channel in a nearby cell becomes free).

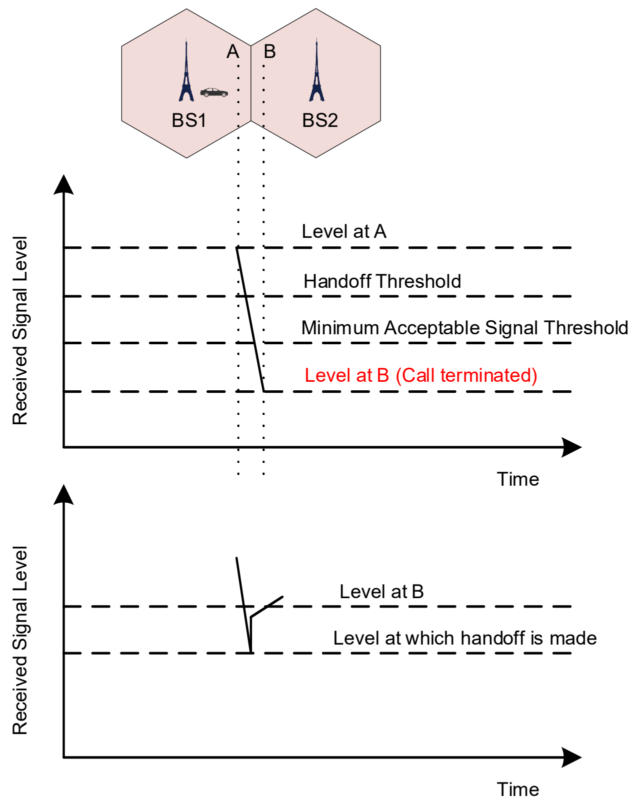

Figure II.4: Proper and improper handoff situations

In deciding when to handoff, it is important to ensure that the drop in the measured signal level is not due to momentary fading and that the mobile is actually moving away from the serving base station. In order to ensure this, the base station monitors the signal level for a certain period of time before a handoff is initiated. This running average measurement of signal strength should be optimized so that unnecessary handoffs are avoided, while ensuring that necessary handoffs are completed before a call is terminated due to poor signal level. The length of time needed to decide if a handoff is necessary depends on the speed at which the vehicle is moving. If the slope of the short-term average received signal level in a given time interval is steep, the handoff should be made quickly. Information about the vehicle speed, which can be useful in handoff decisions, can also be computed from the statistics of the received short-term fading signal at the base station.

The time over which a call may be maintained within a cell, without handoff, is called the dwell time. The dwell time of a particular user is governed by a number of factors, which include propagation, interference, distance between the subscriber and the base station, and other time varying effects. Even when a mobile user is stationary, ambient motion in the vicinity of the base station and the mobile can produce fading, thus even a stationary subscriber may have a random and finite dwell time.

Statistics of dwell time vary greatly, depending on the speed of the user and the type of radio coverage. For example, in cells which provide coverage for vehicular highway users, most users tend to have a relatively constant speed and travel along fixed and well-defined paths with good radio coverage. In such instances, the dwell time for an arbitrary user is a random variable with a distribution that is highly concentrated about the mean dwell time. On the other hand, for users in dense, cluttered microcell environments, there is typically a large variation of dwell time about the mean, and the dwell times are typically shorter than the cell geometry would otherwise suggest. It is apparent that the statistics of dwell time are important in the practical design of hand off algorithms

In first generation analog cellular systems, received signal strength information (RSSI) measurements are made by the base stations and supervised by the MSC. Each base station constantly monitors the signal strengths of all of its reverse voice channels to determine the relative location of each mobile user with respect to the base station tower. In addition to measuring the RSSI of calls in progress within the cell, a spare receiver in each base station, called the locator receiver, is used to determine signal strengths of mobile users which are in neighboring cells. The locator receiver is controlled by the MSC and is used to monitor the signal strength of users in neighboring cells which appear to be in need of handoff and reports all RSSI values to the MSC. Based on the locator receiver signal strength information from each base station, the MSC decides if a handoff is necessary or not.

In second generation systems that use digital TDMA technology, handoff decisions are mobile assisted. In mobile assisted handoff (MAHO), every mobile station measures the received power from surrounding base stations and continually reports the results of these measurements to the serving base station. A handoff is initiated when the power received from the base station of a neighboring cell begins to exceed the power received from the current base station by a certain level or for a certain period of time. The MAHO method enables the call to be handed over between base stations at a much faster rate than in first generation analog systems since the handoff measurements are made by each mobile, and the MSC no longer constantly monitors signal strengths. MAHO is particularly suited for microcellular environments where handoffs are more frequent.

During the course of a call, if a mobile moves from one cellular system to a different cellular system controlled by a different MSC, an intersystem handoff becomes necessary. An MSC engages in an intersystem handoff when a mobile signal becomes weak in a given cell and the MSC cannot find another cell within its system to which it can transfer the call in progress. There are many issues that must be addressed when implementing an intersystem handoff. For instance, a local call may become a long-distance call as the mobile moves out of its home system and becomes a roamer in a neighboring system. Also, compatibility between the two MSCs must be determined before implementing an intersystem handoff.

Different systems have different policies and methods for managing handoff requests. Some systems handle handoff requests in the same way they handle originating calls. In such systems, the probability that a handoff request will not be served by a new base station is equal to the blocking probability of incoming calls. However, from the user's point of view, having a call abruptly terminated while in the middle of a conversation is more annoying than being blocked occasionally on a new call attempt. To improve the quality of service as perceived by the users, various methods have been devised to prioritize handoff requests over call initiation requests when allocating voice channels.

II.3.B. Prioritizing Handoffs

One method for giving priority to handoffs is called the 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 handed off into the cell. This method has the disadvantage of reducing 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 demand-based allocation, are used.

Queuing of handoff requests is another method to decrease the probability of forced termination of a call due to lack of available channels. There is a tradeoff between the decrease in probability of forced termination and total carried traffic. Queuing of handoffs is possible due to the fact that there is a finite time interval between the time the received signal level drops below the handoff threshold and the time the call is terminated due to insufficient signal level. The delay time and size of the queue is determined from the traffic pattern of the particular service area. It should be noted that queuing does not guarantee a zero probability of forced termination, since large delays will cause the received signal level to drop below the minimum required level to maintain communication and hence lead to forced termination.

In practical cellular systems, several problems arise when attempting to design for a wide range of mobile velocities. High speed vehicles pass through the coverage region of a cell within a matter of seconds, whereas pedestrian users may never need a handoff during a call. Particularly with the addition of microcells to provide capacity, the MSC can quickly become burdened if high speed users are constantly being passed between very small cells. Several schemes have been devised to handle the simultaneous traffic of high speed and low speed users while minimizing the handoff intervention from the MSC.

Another practical handoff problem in microcell systems is known as cell dragging. Cell dragging results from pedestrian users that provide a very strong signal to the base station. Such a situation occurs in an urban environment when there is a line-of-sight (LOS) radio path between the subscriber and the base station. As the user travels away from the base station at a very slow speed, the average signal strength does not decay rapidly. Even when the user has traveled well beyond the designed range of the cell, the received signal at the base station may be above the handoff threshold, thus a handoff may not be made. This creates a potential interference and traffic management problem, since the user has meanwhile traveled deep within a neighboring cell. To solve the cell dragging problem, handoff thresholds and radio coverage parameters must be adjusted carefully.

In first generation analog cellular systems, the typical time to make a handoff, once the signal level is deemed to be below the handoff threshold, is about 10 seconds. This requires that the value for Δ be on the order of 6 dB to 12 dB. In digital cellular systems such as GSM, the mobile assists with the handoff procedure by determining the best handoff candidates, and the handoff, once the decision is made, typically requires only 1 or 2 seconds. Consequently, Δ is usually between 0 dB and 6 dB in 2G systems. The faster handoff process supports a much greater range of options for handling high speed and low speed users and provides the MSC with substantial time to “rescue” a call that is in need of handoff.

Another feature of newer cellular systems is the ability to make handoff decisions based on a wide range of metrics other than signal strength. The co-channel and adjacent channel interference levels may be measured at the base station or the mobile, and this information may be used with conventional signal strength data to provide a multi-dimensional algorithm for determining when a handoff is needed.

The IS-95 code division multiple access (CDMA) spread spectrum cellular system provides a unique handoff capability that cannot be provided with other wireless systems. Unlike channelized wireless systems that assign different radio channels during a handoff (called a hard handoff), spread spectrum mobiles share the same channel in every cell. Thus, the term handoff does not mean a physical change in the assigned channel, but rather that a different base station handles the radio communication task. By simultaneously evaluating the received signals from a single subscriber at several neighboring base stations, the MSC may actually decide which version of the user’s signal is best at any moment in time. This technique exploits macroscopic space diversity provided by the different physical locations of the base stations and allows the MSC to make a “soft” decision as to which version of the user’s signal to pass along to the PSTN at any instance. The ability to select between the instantaneous received signals from a variety of base stations is called soft handoff.

II.4. Interference and System Capacity

Interference is the major limiting factor in the performance of cellular radio systems. Sources of interference include another mobile in the same cell, a call in progress in a neighboring cell, other base stations operating in the same frequency band, or any non-cellular system which inadvertently leaks energy into the cellular frequency band. Interference on voice channels causes cross talk, where the subscriber hears interference in the background due to an undesired transmission. On control channels, interference leads to missed and blocked calls due to errors in the digital signaling. Interference is more severe in urban areas, due to the large number of base stations and mobiles.

Interference has been recognized as a major bottleneck in increasing capacity and is often responsible for dropped calls. The two major types of system-generated cellular interference are co-channel interference and adjacent channel interference. Even though interfering signals are often generated within the cellular system, they are difficult to control in practice (due to random propagation effects).

II.4.A. 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 called co-channel cells, and the interference between signals from these cells is called co-channel interference. Unlike thermal noise which can be overcome by increasing the signal-to-noise ratio (SNR), co-channel interference cannot be combated by simply increasing the carrier power of a transmitter. This is because an increase in carrier transmit power increases the interference to neighboring co-channel cells. To reduce co-channel interference, co-channel cells must be physically separated by a minimum distance to provide sufficient isolation due to propagation.

When the area of each cell is approximately the same, and the base stations 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 nearest co-channel cells D. By increasing the ratio D/R, the spatial separation between co-channel cells relative to the coverage distance of a cell is increased. Thus, interference is reduced from improved isolation of RF energy from the co-channel cell. The parameter Q, called the co-channel reuse ratio, is related to the cluster size. For a hexagonal geometry

$$\begin{aligned} Q &= \frac{D}{R} \\ &= \sqrt{3N} \end{aligned} \tag{II.9}$$

A smaller value of Q provides larger capacity since the cluster size N is smaller, whereas a larger value of Q improves the transmission quality, due to a smaller level of co-channel interference. A trade-off must be made between these two objectives in actual cellular design.

Let i_0 be the number of co-channel interfering cells. Then, the signal-to-interference ratio (S/I or SIR) for a mobile receiver which monitors a forward channel can be expressed as

$$\text{SIR} = \frac{S}{\sum_{i=1}^{i_0} I_i} \tag{II.10}$$

where S is the desired signal power from the desired base station, and I_i is the interference power caused by the ith interfering co-channel cell base station. If the signal levels of co-channel cells are known, then the SIR for the forward link can be found using equation (II.10).

Propagation measurements in a mobile radio channel show that the average received signal strength at any point decays as a power law of the distance of separation between a transmitter and receiver. The average received power P_r at a distance d from the transmitting antenna is approximated by

$$\begin{aligned} P_r &= P_0\left(\frac{d}{d_0}\right)^{-n} \\ &= P_0 X \end{aligned} \tag{II.11}$$

or

$$\begin{aligned} P_{r,\text{dBm}} &= P_{0,\text{dBm}} + X_{dB} \\ &= P_{0,\text{dBm}} - 10n\log_{10}\left(\frac{d}{d_0}\right) \end{aligned} \tag{II.12}$$

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 sent by the serving base station, and where the interference is due to 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}. The path loss exponent typically ranges between 2 and 4 in urban cellular systems.

When the transmit power of each base station is equal and the path loss exponent is the same throughout the coverage area, S/I for a mobile can be approximated as

$$\frac{S}{I} = \frac{R^{-n}}{\sum_{i=1}^{i_0} D_i^{-n}} \tag{II.13}$$

Considering only the first layer of interfering cells, if all the interfering base stations are equidistant from the desired base station and if this distance is equal to the distance D between cell centers, then (II.13) simplifies to

$$\begin{aligned}\frac{S}{I} &= \frac{(D/R)^n}{i_0} \\ &= \frac{\left(\sqrt{3N}\right)^n}{i_0}\end{aligned} \tag{II.14}$$

Equation (II.14) relates S/I to the cluster size N, which in tum determines the overall capacity of the system from (II.3). For example, we can assume that the six closest cells are close enough to create significant interference and that they are all approximately equal distance from the desired base station.

Example II-2

For the U.S. AMPS cellular system which uses FM and 30 kHz channels, sufficient voice quality is provided when S/I is greater than or equal to 18 dB. Using (II.14) with a path loss exponent of $n = 4$,

$$\frac{\left(\sqrt{3N}\right)^4}{6} \geq 10^{1.8} \Rightarrow$$
$$N \geq 6.49$$

Thus a minimum cluster size of 7 is required to meet an S/I requirement of 18 dB.

It should be noted that (II.14) is based on the hexagonal cell geometry where all the interfering cells are equidistant from the base station receiver, and hence provides an optimistic result in many cases. For some frequency reuse plans the closest interfering cells vary widely in their distances from the desired cell.

From Figure II.5, it can be seen for a 7-cell cluster, when the mobile unit is at the cell boundary, the mobile is a distance $D - R$ from the two nearest co-channel interfering cells and approximately $D + R/2$, D, $D - R/2$, and $D + R$ from the other interfering cells in the first tier. Using (II.13) and assuming $n = 4$, the signal-to-interference ratio for the worst case can be closely approximated as

$$\frac{S}{I} = \frac{R^{-4}}{2(D-R)^{-4} + 2D^{-4} + 2(D+R)^{-4}} \tag{II.15}$$

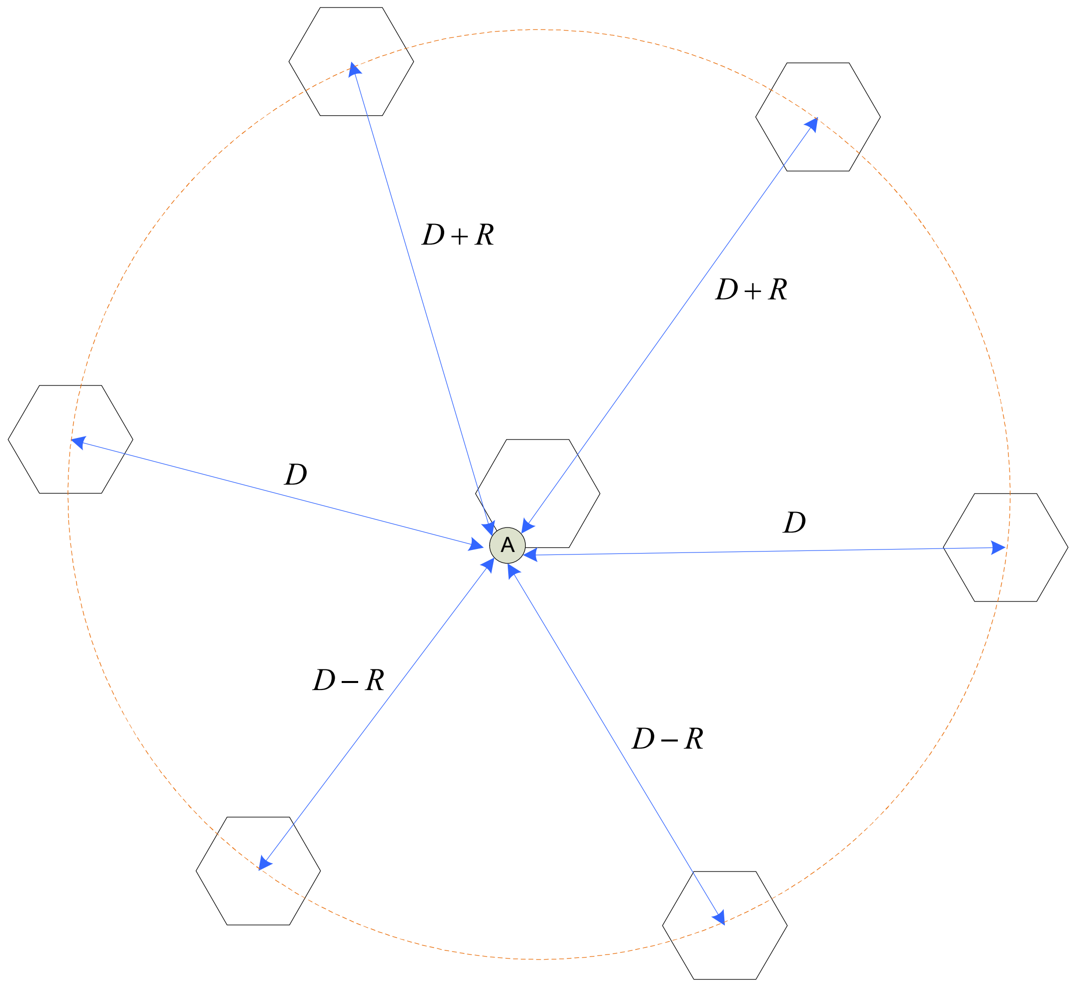

Figure II.5: First tier of co-channels for *N*=7

Equation (II.15) can be rewritten in terms of the co-channel reuse ratio Q as

$$\frac{S}{I}=\frac{R^{-4}/R^{-4}}{2(D-R)^{-4}/R^{-4}+2D^{-4}/R^{-4}+2(D+R)^{-4}/R^{-4}} \tag{II.16}$$

$$\frac{S}{I}=\frac{1}{2(Q-1)^{-4}+2Q^{-4}+2(Q+1)^{-4}} \tag{II.17}$$

$$\frac{S}{I}=\frac{1}{2\left(\sqrt{3N}-1\right)^{-4}+2\left(\sqrt{3N}\right)^{-4}+2\left(\sqrt{3N}+1\right)^{-4}} \tag{II.18}$$

For $N = 7$, the co-channel reuse ratio Q is $\sqrt{21} \approx 4.6$, and the worst case S/I is approximated as 49.56 (17 dB) using (II.17), whereas using (II.13) yields 17.8 dB. Hence for a 7-cell cluster, the S/I ratio is slightly less than 18 dB for the worst case. To design the cellular system for proper performance in the worst case, it would be necessary to increase N to the next largest size, which from (II.7) is found to be 12 (corresponding to $i = j = 2$). This obviously entails a significant decrease in capacity, since 12-cell reuse offers a spectrum utilization of $1/12$ within each cell, whereas 7-cell reuse offers a spectrum utilization of $1/7$. In practice, a capacity reduction of $7/12$ would not be tolerable to accommodate for the worst case situation which rarely occurs. From the above discussion it is clear that co-channel interference determines link performance, which in turn dictates the frequency reuse plan and the overall capacity of cellular systems.

Example II-3

If a signal to interference ratio 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 6 co-channels 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 7-cell reuse pattern. The co-channel reuse ratio is $Q = 4.583$. Using (II.14), the signal-to-noise interference ratio is given by $S/I = (4.583)^4/6 = 75.3 = 18.66 \text{ dB}$. Since this is greater than the minimum required S/I, $N = 7$ can be used.

b. $n = 3$: First, let us consider a 7-cell reuse pattern. The co-channel reuse ratio is $Q = 4.583$. Using (II.14), the signal-to-noise interference ratio is given by $S/I = (4.583)^3/6 = 16.04 = 12.05 \text{ dB}$. Since this is less than the minimum required S/I, a larger N should be used. The next possible value of N is 12. Therefore, $Q = \sqrt{3 \times 12} = 6$. Using (II.14), $S/I = (6)^3/6 = 36 = 15.56 \text{ dB}$. Since this is greater than the minimum required S/I, $N = 12$ can be used.

II.5. Trunking and Grade of Service

Cellular radio systems rely on trunking to accommodate a large number of users in a limited radio spectrum. The concept of trunking allows a large number of users to share the relatively small number of channels in a cell by providing access to each user, on demand, from a pool of available channels. In a trunked radio system, each user is allocated a channel on a per call basis, and upon termination of the call, the previously occupied channel is immediately returned to the pool of available channels.

Trunking exploits the statistical behavior of users so that a fixed number of channels or circuits may accommodate a large, random user community. There is a trade-off between the number of available telephone circuits and the likelihood of a particular user finding that no circuits are available during the peak calling time. In a trunked mobile radio system, when a particular user requests service and all of the radio channels are already in use, the user is blocked, or denied access to the system. In some systems, a queue may be used to hold the requesting users until a channel becomes available.

To design trunked radio systems that can handle a specific capacity at a specific "grade of service", it is essential to understand trunking theory and queuing theory. The fundamentals of trunking theory were developed by Erlang, a Danish mathematician in the late 19th century. Today, the measure of traffic intensity bears his name. One Erlang represents the amount of traffic intensity carried by a channel that is completely occupied (i.e., 1 call-hour per hour or 1 call-minute per minute). For example, a radio channel that is occupied for thirty minutes during an hour carries 0.5 Erlangs of traffic.

The grade of service (GOS) is a measure of the ability of a user to access a trunked system during the busiest hour. The grade of service is a benchmark used to define the desired performance of a particular trunked system by specifying a desired likelihood of a user obtaining channel access given a specific number of channels available in the system. It is the wireless designer's job to estimate the maximum required capacity and to allocate the proper number of channels in order to meet the GOS. GOS is typically given as the likelihood that a call is blocked, or the likelihood of a call experiencing a delay greater than a certain queuing time.

Request Rate: The average number of call requests per unit time. Denoted by λ seconds^{-1}.

Setup Time: The time required to allocate a trunked radio channel to a requesting user.

Blocked Call: Call which cannot be completed at time of request, due to congestion. Also referred to as a lost call.

Holding Time: Average duration of a typical call. Denoted by H (in seconds).

Traffic Intensity: Measure of channel time utilization, which is the average channel occupancy measured in Erlangs. This is a dimensionless quantity and may be used to measure the time utilization of single or multiple channels. Denoted by A. The traffic intensity offered by each user is equal to the call request rate multiplied by the holding time. That is, each user generates a traffic intensity of

$$A_u = \lambda H \qquad \text{(II.19)}$$

For a system containing U users and an unspecified number of channels, the total offered traffic intensity is given as

$$A = UA_u \tag{II.20}$$

If the system has C channels, and the traffic is equally distributed among the channels, then the traffic intensity per channel is given as

$$\begin{aligned} A_c &= \frac{A}{C} \\ &= \frac{UA_u}{C} \end{aligned} \tag{II.21}$$

Note that the offered traffic is not necessarily the traffic which is carried by the trunked system, only that which is offered to the trunked system. When the offered traffic exceeds the maximum capacity of the system, the carried traffic becomes limited due to the limited capacity (i.e. limited number of channels). The maximum possible carried traffic is the total number of channels C in Erlangs.

Load: Traffic intensity across the entire trunked radio system, measured in Erlangs.

Grade of Service (GOS): A measure of congestion which is specified as the probability of a call being blocked (for Erlang B), or the probability of a call being delayed beyond a certain amount of time (for Erlang C).

There are two types of trunked systems which are commonly used. These two types are discussed below.

II.5.A. Trunking with Blocked Calls Cleared

The first type offers no queuing for call requests. That is, for every user who requests service, it is assumed there is no setup time and the user is given immediate access to a channel if one is available. If no channels are available, the requesting user is blocked without access and is free to try again later. This type of trunking is called "blocked calls cleared" and assumes that calls arrive as determined by a Poisson distribution. Furthermore, it is assumed that there are an infinite number of users as well as the following:

(a) there are memoryless arrivals of requests, implying that all users, including blocked users, may request a channel at any time
(b) the probability of a user occupying a channel is exponentially distributed, so that longer calls are less likely to occur as described by an exponential distribution
(c) there are a finite number of channels available in the trunking pool. This is known as an $M/M/m$ queue, and leads to the derivation of the Erlang B formula (also known as the "blocked calls cleared" formula). The Erlang B formula determines the probability that a call

is blocked and is a measure of the GOS for a trunked system which provides no queuing for blocked calls. The Erlang B formula is given by

$$\Pr\left\{\text{Blocking}\right\} = \frac{\dfrac{A^C}{C!}}{\sum_{k=0}^{C} \dfrac{A^k}{k!}} \qquad \text{(II.22)}$$

where C is the number of trunked channels offered by a trunked radio system and A is the total offered traffic. While it is possible to model trunked systems with finite users, the resulting expressions are much more complicated than the Erlang B result, and the added complexity is not warranted for typical trunked systems which have users that outnumber available channels by orders of magnitude. Furthermore, the Erlang B formula provides a conservative estimate of the GOS, as the finite user results always predict a smaller likelihood of blocking. The capacity of a trunked radio system where blocked calls are lost is tabulated for various values of GOS and numbers of channels in Table II.1.

Table II.1: Capacity of an Erlang B System

Number of Channels C	Capacity (Erlangs) for GOS =			
	0.01	0.005	0.002	0.001
2	0.153	0.105	0.065	0.046
5	1.36	1.13	0.900	0.762
10	4.46	3.96	3.43	3.09
20	12.0	11.1	10.1	9.41
40	29.0	27.3	25.7	24.5
100	84.1	80.9	77.4	75.2

Example II-4

How many users can be supported for 0.5% blocking probability for the following number of trunked channels in a blocked calls cleared system? Assume each user generates 0.1 Erlangs of traffic.

(d) 1
(e) 5
(f) 10
(g) 20
(h) 100

Solution

From Table II.1, we can find the total capacity in Erlangs for the 0.5% GOS for different numbers of channels. By using the relation $A = UA_u$, we can obtain the total number of users that can be supported in the system.

a. $C = 1$, $A_u = 0.1$, $GOS = 0.005$, then from Fig. 2.6 (Rappaport, page 49) we get $A = 0.005$. Therefore, total number of users $U = A/A_u = 0.05$ users. But, actually one user could be supported on one channel. So, $U = 1$.

b. $C = 5$. From Fig. 2.6 (Rappaport, page 49) we get $A = 1.13$. Therefore, total number of users $U = A/A_u \approx 11$ users.

c. $C = 10$. From Fig. 2.6 (Rappaport, page 49) we get $A = 3.96$. Therefore, total number of users $U = A/A_u \approx 39$ users.

d. $C = 20$. From Fig. 2.6 (Rappaport, page 49) we get $A = 11.1$. Therefore, total number of users $U = A/A_u \approx 110$ users.

e. $C = 100$. From Fig. 2.6 (Rappaport, page 49) we get $A = 80.9$. Therefore, total number of users $U = A/A_u \approx 809$ users.

Example II-5

An urban area has a population of 2 million residents. Three competing trunked mobile networks (systems A, B, and C) provide eellular 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 probability if each user averages 2 calls per hour at an average call duration of 3 minutes. Assuming that all three trunked systems are operated at maximum capacity, compute the percentage market penetration of each cellular provider.

Solution

Traffic intensity per user $A_u = \lambda H = 2(3/60) = 0.1$ Erlangs.

System A:

From the Erlang B chart, the total carried traffic A is obtained as 12 Erlangs. The number of users that can be supported per cell is $U = A/A_u = 12/0.1 = 120$. Since there are 394 cells. the total number of subscribers that can be supported by System A is equal to $120(394) = 47280$.

System B:

From the Erlang B chart, the total carried traffic A is obtained as 45 Erlangs. The number of users that can be supported per cell is $U = A/A_u = 45/0.1 = 450$. Since there are 98 cells. the total number of subscribers that can be supported by System A is equal to $450(98) = 44100$.

System C:

From the Erlang B chart, the total carried traffic A is obtained as 88 Erlangs. The number of users that can be supported per cell is $U = A/A_u = 88/0.1 = 880$. Since there are 49 cells. the total number of subscribers that can be supported by System A is equal to $880(49) = 43120$.

Therefore, total number of cellular subscribers that can be supported by these three systems are 47280 + 44100 + 43120 = 134500 users.

Since there are 2 million residents in the given urban area, the percentage market penetration is

System A: 47280/2000000 = 2.36 %.

System B: 44100/2000000 = 2.205 %.

System C: 43120/2000000 = 2.156 %.

All Three Systems: 134500/2000000 = 6.725 %.

Example II-6

A certain city has an area of 1,300 square miles and is covered by a cellular system using a 7-cell reuse pattern. Each cell has a radius of 4 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. number of cells in the service area
b. number of channels per cell
c. traffic intensity of each cell
d. maximum carried traffic
e. total number of users that can be served for 2% GOS
f. number of mobiles per channel
g. the theoretical maximum number of users that could be served at one time by the system

Solution

Area covered per cell is $3\sqrt{3}R^2/2 \approx 41.57 \text{ mi}^2$.

a. Total number of cells is $N_c = 1300/41.57 \approx 31$ cells.

b. Total number of channels = $40\times10^6/60\times10^3 \approx 666$. Number of channels per cell $C = 666/7 \approx 95$.

c. From the Erlang B chart, we have traffic intensity per cell $A = 84$ Erlangs/cell.
d. Maximum carried traffic=(number of cells).(traffic intensity per cell)= $31 \times 84 = 2604$ Erlangs.
e. Total number of users= $2604/0.03 \approx 86,800$ users.
f. Number of mobiles per channel $= 86,800/666 \approx 130$ mobiles/channel.
g. The theoretical maximum number of served mobiles is the number of available channels in the system (all channels occupied) $= 95 \times 31 = 2945$.

II.5.B. Trunking with Blocked Calls Delayed

The second kind of trunked system is one in which a queue is provided to hold calls which are blocked. If a channel is not available immediately, the call request may be delayed until a channel becomes available. This type of trunking is called "Blocked Calls Delayed", and its measure of GOS is defined as the probability that a call is blocked after waiting a specific length of time in the queue. To find the GOS, it is first necessary to find the likelihood that a call is initially denied access to the system.

The likelihood of a call not having immediate access to a channel is determined by the Erlang C formula

$$\Pr\{\text{Delay>0}\} = \frac{A^C}{A^C + C!\left(1 - \frac{A}{C}\right)\sum_{k=0}^{C-1}\frac{A^k}{k!}} \qquad \text{(II.23)}$$

If no channels are immediately available the call is delayed, and the probability that the delayed call is forced to wait more than t seconds is given by the probability that a call is delayed, multiplied by the conditional probability that the delay is greater than t seconds. The GOS of a trunked system where blocked calls are delayed is hence given by

$$\begin{aligned}\Pr\{\text{Delay>}t\} &= \Pr\{\text{Delay>0}\}\Pr\{\text{Delay>}t \mid \text{Delay>0}\} \\ &= \Pr\{\text{Delay>0}\}\, e^{-\frac{(C-A)}{H}t}\end{aligned} \qquad \text{(II.24)}$$

The average delay D for all calls in a queued system is given by

$$\begin{aligned}D &= \Pr\{\text{Delay>0}\}\, D_q \\ &= \Pr\{\text{Delay>0}\}\frac{H}{C - A}\end{aligned} \qquad \text{(II.25)}$$

where D_q is the average delay of queued call.

Example II-7

A hexagonal cell within a 4-ce11 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 call:

a. How many users per square kilometer will this 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

$C = 60/4 = 15$ channels/cell. Area covered per cell=$3\sqrt{3}(1.387)^2/2 \approx 5 \text{ km}^2$.

a. From Erlang C chart for 5% probability of delay with $C = 15$, traffic intensity=9.0 Erlangs. Therefore, number of users $= 9/0.029 \approx 310$ users. Users/squared km $= 310/5 = 62$ users/km^2.
b. $H = A_u/\lambda = 0.029$ hour $= 104.4$ seconds. The probability that a delayed call will have to wait for more than 10 seconds is

$$\Pr\{\text{Delay}>t \mid Delay > 0\} = \exp(-(C-A)t/H) = \exp(-(15-9)\times 10/104.4) = 56.29\% = 0.5629$$

c. Probability that a call is delayed more than 10 seconds is

$$\Pr\{\text{Delay}>t \mid Delay > 0\}\ \Pr\{Delay > 0\} = 0.05 \times 0.5629 = 2.81\% = 0.0281$$

Trunking efficiency is a measure of the number of users which can be offered a particular GOS with a particular configuration of fixed channels. The way in which channels are grouped can substantially alter the number of users handled by a trunked system. For example, from Table II.1, 10 trunked channels at a GOS of 0.01 can support 4.46 Erlangs of traffic, whereas 2 groups of 5 trunked channels can support (2)(1.36) Erlangs, or 2.72 Erlangs of traffic. Clearly, 10 channels trunked together support 60% more traffic at a specific GOS than do two 5 channel trunks. It should be clear that the allocation of channels in a trunked radio system has a major impact on overall system capacity.

II.6. Improving Capacity In Cellular Systems

As the demand for wireless service increases, the number of channels assigned to a cell eventually becomes insufficient to support the required number of users. At this point, cellular design techniques are needed to provide more channels per unit coverage area. Techniques such as cell splitting, sectoring, and coverage zone approaches are used in practice to expand the capacity of cellular systems.

- Cell splitting allows an orderly growth of the cellular system.

- Sectoring uses directional antennas to further control the interference and frequency reuse of channels.
- The zone microcell concept distributes the coverage of a cell and extends the cell boundary to hard-to-reach places.

While cell splitting increases the number of base stations in order to increase capacity, sectoring and zone microcells rely on base station antenna placements to improve capacity by reducing co-channel interference. Cell splitting and zone microcell techniques do not suffer the trunking inefficiencies experienced by sectored cells, and enable the base station to oversee all handoff chores related to the microcells, thus reducing the computational load at the MSC.

II.6.A. Cell Splitting

Cell splitting is the process of subdividing a congested cell into smaller cells, each with its own base station and a corresponding reduction in antenna height and transmitter power. Cell splitting increases the capacity of a cellular system since it increases the number of times channels are reused. By defining new cells which have a smaller radius than the original cells and by installing these smaller cells (microcells), capacity increases due to the additional number of channels per unit area.

Imagine if every cell in Figure II.1 were reduced in such a way that the radius of every cell was cut in half. In order to cover the entire service area with smaller cells, approximately four times as many cells would be required. The increased number of cells would increase the number of clusters over the coverage region, which in turn would increase the number of channels, and thus capacity, in the coverage area. Cell splitting allows a system to grow by replacing large cells with smaller cells, while not upsetting the channel allocation scheme required to maintain the minimum co-channel reuse ratio Q between co-channel cells. Note that $Q = \sqrt{3N}$ is not affected by cell splitting.

For the new cells to be 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 found 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.

$$\mathrm{P}_r[\text{at old cell boundary}] \propto P_{t,old} R^{-n} \tag{II.26}$$

$$\mathrm{P}_r[\text{at new cell boundary}] \propto P_{t,new} \left(\frac{R}{2}\right)^{-n} \tag{II.27}$$

where $P_{t,old}$ and $P_{t,new}$ are the transmit powers of the larger and smaller cell base stations, respectively, and n is the path loss exponent. Equating P_r in both (II.26) and (II.27) and dividing the two equations one gets

$$\frac{P_{t,new}\left(\frac{R}{2}\right)^{-n}}{P_{t,old}R^{-n}}=1 \tag{II.28}$$

Simplifying (II.28) yields

$$\frac{P_{t,new}}{P_{t,old}}=\left(\frac{1}{2}\right)^{n} \tag{II.29}$$

If, for example, $n=3$, then the ratio of the new needed transmit power to the old needed transmit power is 1 to 8. This amounts to approximately 9 dB reduction in transmit power.

In practice, not all cells are split at the same time. It is often difficult for service providers to find real estate that is perfectly situated for cell splitting. Therefore, different cell sizes will exist simultaneously. In such situations, special care needs to be taken to keep the distance between co-channel cells at the required minimum, and hence channel assignments become more complicated. Also, handoff issues must be addressed so that high speed and low speed traffic can be simultaneously accommodated. When there are two cell sizes in the same region, equation (II.29) shows that one cannot simply use the original transmit power for all new cells or the new transmit power for all the original cells. If the larger transmit power is used for all cells, some channels used by the smaller cells would not be sufficiently separated from co-channel cells. On the other hand, if the smaller transmit power is used for all the cells, there would be parts of the larger cells left unserved. For this reason, channels in the old cell must be broken down into two channel groups, one that corresponds to the smaller cell reuse requirements and the other that corresponds to the larger cell reuse requirements. The larger cell is usually dedicated to high speed traffic so that handoffs occur less frequently.

II.6.B. Sectoring

Cell splitting achieves capacity improvement by essentially rescaling the system. By decreasing the cell radius R and keeping the co-channel reuse ratio D/R unchanged, cell splitting increases the number of channels per unit area. However, another way to increase capacity is to keep the cell radius unchanged and seek methods to decrease the D/R ratio. In this approach, capacity improvement is achieved by reducing the number of cells in a cluster and thus increasing the frequency reuse. However, in order to do this, it is necessary to reduce the relative interference without decreasing the transmit power.

The co-channel interference in a cellular system may be decreased by replacing a single omni-directional antenna at the base station by several directional antennas, each radiating within a specified sector. By using directional antennas, a given cell will receive interference and transmit with only a fraction of the available co-channel cells. The technique for decreasing co-channel interference and thus increasing system capacity by using directional antennas is called sectoring. The factor by which the co-channel interference is reduced depends on the amount of sectoring used. A cell is normally partitioned into three 120° sectors or six 60° sectors.

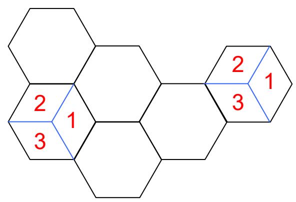

Figure II.6: 120° sectoring

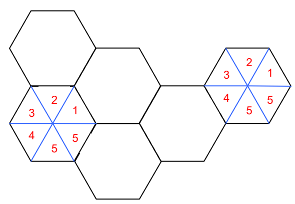

Figure II.7: 60° sectoring

When sectoring is employed, the channels used in a particular cell are broken down into sectored groups and are used only within a particular sector, as illustrated in Figure II.8. Assuming 7-cell reuse, for the case of 120° sectors, the number of interferers in the first tier is reduced from 6 to 2. This is because only 2 of the 6 co-channel cells receive interference with a particular sectored channel group.

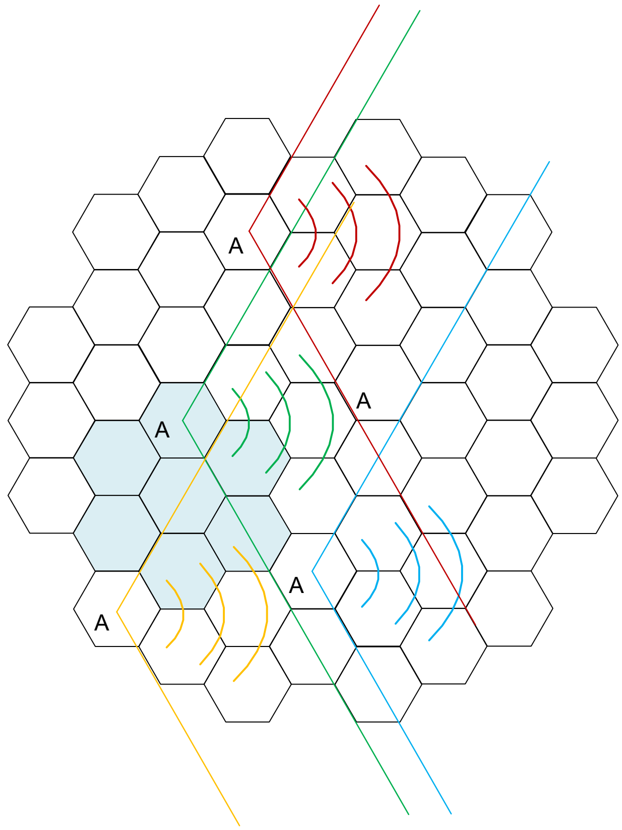

Figure II.8: Illustration of how sectoring reduces interference from co-channel cells

Referring to Figure II.8, consider the interference experienced by a mobile located in the right-most sector in the center cell labeled "A". There are 3 co-channel cell sectors labeled "A" to the right of the center cell, and 3 to the left of the center cell. Out of the 6 co-channel cells, only 2 cells have sectors with antenna patterns which radiate into the center cell, and hence a mobile in the center cell will experience interference on the forward link from only these two sectors. The resulting S/I for this case can be found using equation (II.13) to be 24.2 dB, which is a significant improvement over the omnidirectional case, where the worst case S/I was shown using (II.17) to be 17 dB.

It is the loss of traffic due to decreased trunking efficiency that causes some operators to shy away from the sectoring approach, particularly in dense urban areas where the directional antenna patterns are somewhat ineffective in controlling radio propagation. Because sectoring uses more than one antenna per base station, the available channels in the cel1 must be subdivided and dedicated to a specific antenna. This breaks up the available trunked channel pool into several smaller pools, and decreases trunking efficiency.

Example II-8

Consider a cellular system in which an average call lasts 2 minutes, and the probability of blocking is to be no more than 1%. Assume that every subscriber makes 1 call per hour, on average. If there are a total of 395 traffic channels for a 7-cell reuse system, there will be about 57 traffic channels per cell. Assume that blocked calls are cleared so the blocking is described by the Erlang B distribution. From the Erlang B distribution, it can be found that the unsectored system may handle 44.2 Erlangs or 1326 calls per hour (note that one call is equivalent to 2/60 Erlangs).

Now employing 120° sectoring, there are only 19 channels per antenna sector (57/3 antennas). For the same probability of blocking and average call length, it can be found from the Erlang B distribution that each sector can handle 11.2 Erlangs or 336 calls per hour. Since each cell consists of 3 sectors, this provides a cell capacity of 3 x 336 = 1008 calls per hour, which amounts to a 24% decrease when compared to the unsectored case. Thus, sectoring decreases the trunking efficiency while improving the S/I for each user in the system.

It can be found that using 60° sectors improves the S/I even more. In this case the number of first tier interferers is reduced from 6 to only 1. This results in S/I = 29 dB for a 7-cell system and enables 4-cell reuse. Of course, using 6 seetors per cell reduces the trunking efficiency and increases the number of necessary handoffs even more. If the unsectored system is compared to the 6 sector case, the degradation in trunking efficiency can be shown to be 44%.

***

III. MOBILE RADIO PROPAGATION

The design of spectrally efficient wireless communication systems requires a thorough understanding of the radio propagation channel. The mobile radio channel places fundamental limitations on the performance of wireless communication systems. The transmission path between the transmitter and the receiver can vary from simple line-of-sight to one that is severely obstructed by buildings, mountains, and other objects. Unlike wired channels that are stationary and predictable, radio channels are extremely random and do not offer easy analysis. Even the speed of motion impacts how rapidly the signal level fades as a mobile terminal moves in space. The characteristics of the radio channel will vary greatly with the operating frequency, and the propagation environment, e.g., line-of-sight (LoS) versus non-line-of-sight (NLoS), stationary versus mobile transmitters and receivers, and other factors. Modeling the radio channel has historically been one of the most difficult parts of mobile radio system design, and is typically done in a statistical fashion, based on measurements made specifically for an intended communication system or spectrum allocation.

III.1. Introduction to Radio Wave Propagation

A typical cellular land mobile radio system consists of a collection of fixed base stations (BSs) that define radio coverage areas known as cells. The height and placement of the BS antennas affect the proximity of local scatterers at the BSs. In a macro-cellular environment where the cell radii are large, the BS antennas are well elevated above the local terrain and are free of local scatterers. Mobile stations (MSs), on the other hand, tend to be surrounded by local scatterers due to their low elevation antennas.

The mechanisms behind electromagnetic wave propagation are diverse, but can generally be attributed to reflection, diffraction, and scattering. Most cellular radio systems operate in urban areas where there is no direct line-of-sight path between the transmitter and the receiver, and where the presence of tall buildings causes severe diffraction loss. Due to multiple reflections from various objects, the electromagnetic waves travel along different paths of varying lengths. Multiple plane waves will arrive at the MS (or BS) receiver antenna(s) from different directions, with each having a distinct polarization, amplitude, phase, and delay. This phenomenon is called multipath propagation. The multiple plane waves combine vectorially at each MS (or BS) receiver antenna to produce a composite received signal.

Commercial cellular land mobile radio systems operate at UHF frequencies in bands located at 700/800/900MHz and 1800/1900 MHz. At these frequencies, the carrier wavelength λ_c is approximately 15 cm and 30 cm, respectively, using the relationship

$$c = f_c \lambda_c \tag{III.1}$$

where f_c is the carrier frequency and c is the speed of light. Therefore, small changes in the propagation delays of the individual multipath components due to MS mobility on the order of a few centimeters will cause a large change in the relative phases of the plane wave components arriving at the MS (or BS) receiver antennas. Hence, when the arriving plane waves combine vectorially at the receiver antenna(s), they will experience constructive and destructive addition depending on the physical location of the MS. If the MS is moving or there are changes in the location of the scatterers, then these spatial variations will manifest themselves as time variations in the amplitude and phase of the composite signal received at each MS (or BS) antenna, a phenomenon known as envelope fading. If the propagation environment is such that no individual multipath component is dominant, such as when NLoS conditions exist between the BS and MS, then the composite received envelope under narrowband propagation conditions is often modeled as being Rayleigh distributed at any time. Such a channel is said to exhibit Rayleigh fading. However, if a dominant multipath component exists, such as when a LoS or specular condition exists between the BS and MS, then the envelope is often modeled as being Ricean distributed at any time. Such a channel is said to exhibit Ricean fading.

Propagation models have traditionally focused on predicting the average received signal strength at a given distance from the transmitter, as well as the variability of the signal strength in close spatial proximity to a particular location. Propagation models that predict the mean signal strength for an arbitrary transmitter-receiver (T-R) separation distance are useful in estimating the radio coverage area of a transmitter and are called large-scale propagation models, since they characterize signal strength over large T-R separation distances (several hundreds or thousands of meters). On the other hand, propagation models that characterize the rapid fluctuations of the received signal strength over very short travel distances (a few wavelengths) or short time durations (on the order of seconds) are called small-scale fading models.

As a mobile moves over very small distances, the instantaneous received signal strength may fluctuate rapidly giving rise to small-scale fading. The reason for this is that the received signal is a sum of many contributions coming from different directions. Since the phases are random, the sum of the contributions varies widely; for example, obeys a Rayleigh fading distribution. In small-scale fading, the received signal power may vary by as much as three or four orders of magnitude (30 or 40 dB) when the receiver is moved by only a fraction of a wavelength. As the mobile moves away from the transmitter over much larger distances, the local average received signal will gradually decrease, and it is this local average signal level that is predicted by large-scale propagation models. Typically, the local average received power is computed by averaging signal measurements over a measurement track of 5λ to 40λ. For cellular frequencies in the 1 GHz to 2 GHz band, this corresponds to measuring the local average received power over movements of 1 m to 10 m.

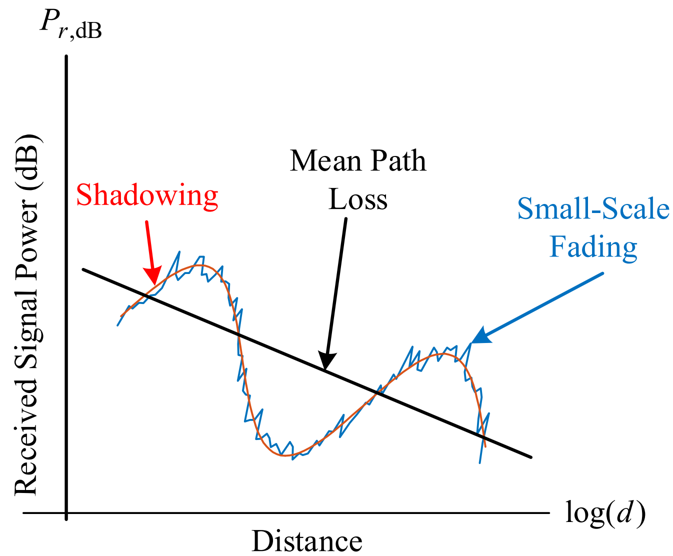

Figure III.1: Large-scale fading vs. small-scale fading

III.2. Free Space Propagation Model

The free space propagation model is used to predict received signal strength when the transmitter and receiver have a clear, unobstructed line-of-sight path between them. Satellite communication systems and microwave line-of-sight radio links typically undergo free space propagation. As with most large-scale radio wave propagation models, the free space model predicts that received power decays as a function of the T-R separation distance raised to some power (i.e. a power law function). The free space power received by a receiver antenna which is separated from a radiating transmitter antenna by a distance d, is given by the Friis free space equation

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

where P_t is the transmitted power, $P_r(d)$ is the received power which is a function of the T-R separation, G_t is the transmitter antenna gain, G_r is the receiver antenna gain, d is the T-R separation distance in meters, L is the system loss factor not related to propagation ($L \geq 1$), and λ is the wavelength in meters. The gain of an antenna is related to its effective aperture A_e by

$$G = \frac{4\pi A_e}{\lambda^2} \tag{III.3}$$

The effective aperture A_e is related to the physical size of the antenna, and λ is related to the carrier frequency by

$$\begin{aligned}\lambda &= \frac{c}{f_c} \\ &= \frac{2\pi c}{\omega_c}\end{aligned} \tag{III.4}$$

where f_c is the carrier frequency in Hz, ω_c is the carrier frequency in radians per second, and c is the speed of light in meters/s. The values for P_t and P_r must be expressed in the same units, and G_t and G_r are dimensionless quantities. The system loss factor L is usually due to transmission line attenuation, filter losses, and antenna losses in the communication system. A value of $L = 1$ indicates no loss in the system hardware.

The Friis free space equation of (III.2) shows that the received power falls off as the square of the T-R separation distance. This implies that the received power decays with distance at a rate of 20 dB/decade.

The path loss, which represents signal attenuation as a positive quantity measured in dB, is defined as the difference (always in dB) between the effective transmitted power and the received power, and may or may not include the effect of the antenna gains. The path loss for the free space model when antenna gains are included is given by

$$\begin{aligned}PL(dB) &= 10\log\left(\frac{P_t}{P_r}\right) \\ &= -10\log\left(\frac{G_t G_r \lambda^2}{(4\pi)^2 d^2}\right)\end{aligned} \tag{III.5}$$

When antenna gains are excluded, the antennas are assumed to have unity gain, and path loss is given by

$$PL(dB) = -10\log\left(\frac{\lambda^2}{(4\pi)^2 d^2}\right) \tag{III.6}$$

The Friis free space model is only a valid predictor for P_r for values of d which are in the far field of the transmitting antenna. The far field, or Fraunhafer region, of a transmitting antenna is defined as the region beyond the far field distance d_f which is related to the largest linear dimension of the transmitter antenna aperture and the carrier wavelength. The Fraunhofer distance is given by

$$d_f = \frac{2D^2}{\lambda} \tag{III.7}$$

where D is the largest physical linear dimension of the antenna. Additionally, to be in the far field region, d_f must satisfy

$$d_f >> D \tag{III.8}$$

and

$$d_f >> \lambda \tag{III.9}$$

Furthermore, it is clear that equation (III.2) does not hold for $d = 0$. For this reason, large-scale propagation models use a close-in distance d_0 as a known received power reference point. The received power $P_r(d)$ at any distance $d > d_0$ may be related to $P_r(d_0)\cdot$ The value $P_r(d_0)$ may be predicted from equation (III.2) or may be measured in the radio environment by taking the average received power at many points located at a close-in radial distance d_0 from the transmitter. The reference distance must be chosen such that it lies in the far field region, that is, $d_0 \ge d_f$, and d_0 is chosen to be smaller than any practical distance used in the mobile communication system. Thus, using equation (III.2), the received power in free space at a distance greater than d_0 is given by

$$P_r(d) = P_r(d_0)\left(\frac{d_0}{d}\right)^2, \quad d \ge d_0 \ge d_f \tag{III.10}$$

In mobile radio systems, it is not uncommon to find that P_r may change by many orders of magnitude over a typical coverage area of several square kilometers. Because of the large dynamic range of received power levels, often dBm or dB units are used to express received power levels. Equation (III.10) may be expressed in units of dBm or dB 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 given by

$$P_{r,dB}(d) = 10\log\left(P_r(d_0)\right) + 20\log\left(\frac{d_0}{d}\right), \quad d \ge d_0 \ge d_f \tag{III.11}$$

$$P_{r,dBm}(d) = 10\log\left(\frac{P_r(d_0)}{0.001\text{ W}}\right) + 20\log\left(\frac{d_0}{d}\right), \quad d \ge d_0 \ge d_f \tag{III.12}$$

where $\mathrm{P}_r(d_0)$ is in units of Watts.

The reference distance d_0 for practical systems using low-gain antennas in the 1-2 GHz region is typically chosen to be 1 m in indoor environments and 100 m or 1 km in outdoor environments, so that the numerator in equations (III.10) and (III.12) is a power of 10. This makes path loss computations easy in dB units.

Example III-1

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

Solution

$$\lambda = \frac{c}{f} = \frac{3\times 10^8}{900\times 10^6} = \frac{1}{3} \text{ m}$$

$$d_f = \frac{2D^2}{\lambda} = \frac{2(1)^2}{1/3} = 6 \text{ m}$$

Example III-2

If a transmitter produces 50 watts of power, express the transmit power in units of

a. dBm
b. dBW

If 50 watts 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 $P_r(10 \text{ km})$? Assume unity gain for the receiver antenna.

Solution

a. $P_{t,dBm} = 10\log\left(\frac{50}{0.001}\right) = 47 \text{ dBm}$

b. $P_{t,dBW} = 10\log(50) = 17 \text{ dBW}$

The received power can be determined using equation (III.2):

$$P_r = \frac{50(1)(1)(1/3)^2}{(4\pi)^2(100)^2(1)} = 3.5\times 10^{-6} \text{ W} = 3.5\times 10^{-3} \text{ mW}$$

$$P_{r,dBm} = 10\log(3.5\times 10^{-3}) = -24.5 \text{ dBm}$$

$$P_{r,dBW} = 10\log(3.5\times 10^{-6}) = -54.5 \text{ dBW}$$

The received power at 10 km can be expressed in terms of dBm using equation (III.12), where $d_0 = 100$ m and $d = 10,000$ m :

$$P_{r,dBm}(10\text{ km}) = P_{r,dBm}(100\text{ m}) + 20\log\left(\frac{100}{10000}\right) = -24.5 - 40 = -64.5\text{ dBm}$$

III.3. Link Budget Design Using Path Loss Models

Most radio propagation models are derived using a combination of analytical and empirical methods. The empirical approach is based on fitting curves or analytical expressions that recreate a set of measured data. This has the advantage of implicitly taking into account all propagation factors, both known and unknown, through actual field measurements. However, the validity of an empirical model at transmission frequencies or environments other than those used to derive the model can only be established by additional measured data in the new environment at the required transmission frequency. Over time, some classical propagation models have emerged, which are now used to predict large-scale coverage for mobile communication systems design. By using path loss models to estimate the received signal level as a function of distance, it becomes possible to predict the SNR for a mobile communication system.

III.3.A. 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 models have 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.

$$\frac{P_r(d_0)}{P_r(d)} = \left(\frac{d}{d_0}\right)^2, \quad d \ge d_0 \ge d_f \tag{III.13}$$

$$\overline{PL}(d) \propto \left(\frac{d}{d_0}\right)^n \tag{III.14}$$

or, in dB scale,

$$\overline{PL}_{dB}(d) = \overline{PL}_{dB}(d_0) + 10n\log\left(\frac{d}{d_0}\right) \tag{III.15}$$

When plotted on a log-log scale, the modeled path loss is a straight line with a slope equal to $10n$ dB per decade. The value of n depends on the specific propagation environment. For example, in free space, n is equal to 2, and when obstructions are present, n will have a larger value.

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 distances are commonly used, whereas in microcellular systems, much smaller distances (such as 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 path loss is calculated using the free space path loss formula given by equation (III.5) or through field measurements at distance d_0.

III.3.B. LOG-NORMAL SHADOWING

The model in equation (III.15) does not consider the fact that the surrounding environmental clutter may be vastly different at two different locations having the same T-R separation. This leads to measured signals which are vastly different than the average value predicted by equation (III.15). Measurements have shown that at any value of d, the path loss $PL(d)$ at a particular location is random and distributed log-normally (normal in dB) about the mean distance-dependent value. That is

$$\begin{aligned} PL_{dB}(d) &= \overline{PL}_{dB}(d) + X_\sigma \\ &= \overline{PL}_{dB}(d_0) + 10n\log\left(\frac{d}{d_0}\right) + X_\sigma \end{aligned} \qquad \text{(III.16)}$$

and

$$P_{r,dB}(d) = P_{t,dB} - PL_{dB}(d) \qquad \text{(III.17)}$$

where X_σ is a zero-mean Gaussian distributed random variable (in dB) with standard deviation σ (also in dB). Let's define the mean values of random variables $PL_{dB}(d)$ and $P_{r,dB}(d)$ as follows:

$$\begin{aligned} \mu_L(d) &= \overline{PL}_{dB}(d) \\ &= \overline{PL}_{dB}(d_0) + 10n\log\left(\frac{d}{d_0}\right) \end{aligned} \qquad \text{(III.18)}$$

$$\begin{aligned} \mu_r(d) &= \overline{P}_{r,dB}(d) \\ &= P_{t,dB} - \overline{PL}_{dB}(d) \end{aligned} \qquad \text{(III.19)}$$

Note that both $PL_{dB}(d)$ and $P_{r,dB}(d)$ have the same variance σ^2. The Gaussian probability density functions of $PL_{dB}(d)$ and $P_{r,dB}(d)$ can be expressed as

$$f_{PL_{dB}(d)}(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu_L(d))^2}{2\sigma^2}} \tag{III.20}$$

$$f_{P_{r,dB}(d)}(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu_r(d))^2}{2\sigma^2}} \tag{III.21}$$

The log-normal distribution describes the random shadowing effects which occur over a large number of measurement locations which have the same T-R separation, but have different levels of clutter on the propagation path. This phenomenon is referred to as log-normal shadowing. Simply put, log-normal shadowing implies that measured signal levels at a specific T-R separation have a Gaussian (normal) distribution about the distance-dependent mean path loss, where the measured signal levels have values in dB units. The standard deviation of the Gaussian distribution that describes the shadowing also has units in dB. Thus, the random effects of shadowing are accounted for using the Gaussian distribution.

The reference distance d_0, the path loss exponent n , and the standard deviation σ, statistically describe the path loss model for an arbitrary location having a specific T-R separation, and this model may be used in computer simulation to provide received power levels for random locations in communication system design and analysis.

In practice, the values of n and σ are computed from measured data, using linear regression such that the difference between the measured and estimated path losses is minimized in a mean square error sense over a wide range of measurement locations and T-R separations. The value of $\overline{PL}_{dB}(d_0)$ is based on either measurements or on a free space assumption from the transmitter to d_0.

The probability that the received signal level will exceed a certain value γ can be calculated from (III.21) as

$$\begin{aligned}\Pr\left\{P_{r,dB}(d)>\gamma\right\}&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{\gamma}^{\infty}e^{-\frac{(x-\mu_r(d))^2}{2\sigma^2}}dx\\&=Q\left(\frac{\gamma-\mu_r(d)}{\sigma}\right)\end{aligned} \tag{III.22}$$

where the Q-function is defined as

$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-\frac{x^2}{2}} dx \qquad \text{(III.23)}$$

Similarly, the probability that the received signal level will be below γ is given by

$$\Pr\left\{P_{r,dB}(d) < \gamma\right\} = Q\left(\frac{\mu_r(d) - \gamma}{\sigma}\right) \qquad \text{(III.24)}$$

Note that the following Q-function identity has been used in (III.24):

$$Q(-z) = 1 - Q(z) \qquad \text{(III.25)}$$

Example III-3

Suppose the standard deviation of the received power due to shadowing is equal to 8 dB.

1. What is the probability that the path loss will exceed the mean path loss by at least 5 dB?
2. What is the probability that the path loss will exceed the mean path loss by at least 10 dB?

Solution

$$\Pr\left\{PL_{dB} - \overline{PL}_{dB} > \xi\right\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\xi}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= Q\left(\frac{\xi}{\sigma}\right)$$

1. $\Pr\left\{PL_{dB} - \overline{PL}_{dB} > 5\right\} = Q\left(\frac{5}{8}\right) = 0.266$.
2. $\Pr\left\{PL_{dB} - \overline{PL}_{dB} > 10\right\} = Q\left(\frac{10}{8}\right) = 0.1056$.

III.3.C. Percentage of Coverage Area

It is clear that due to random effects of shadowing, some locations within a coverage area will be below a particular desired received signal threshold. It is often useful to compute how the boundary coverage relates to the percent of area covered within the boundary. For a circular coverage area having radius R from a base station, let there be some desired received signal threshold γ. We are interested in computing $U(\gamma)$, the percentage of useful service area (i.e. the percentage of area with a received signal that is equal or greater than γ), given a known likelihood of coverage at the cell boundary. Letting $d = r$ represent the radial distance from the transmitter, it can be shown that if $\Pr\left\{P_{r,dB}(r) > \gamma\right\}$ is the probability that the random received signal at $d = r$ exceeds the threshold γ within an incremental area dA, then $U(\gamma)$ can be found by

$$\begin{aligned} U(\gamma) &= \frac{1}{\pi R^2} \int \Pr\left\{P_{r,dB}(r) > \gamma\right\} dA \\ &= \frac{1}{\pi R^2} \int_0^{2\pi} \int_0^{R} \Pr\left\{P_{r,dB}(r) > \gamma\right\} r dr d\theta \end{aligned} \tag{III.26}$$

Using (III.22),

$$\Pr\left\{P_{r,dB}(r) > \gamma\right\} = Q\left(\frac{\gamma - \mu_r(r)}{\sigma}\right) \tag{III.27}$$

From (III.17), we have

$$\mu_r(r) = P_{t,dB} - \overline{PL}_{dB}(r) \tag{III.28}$$

From (III.16), we have

$$\overline{PL}_{dB}(r) = \overline{PL}_{dB}(d_0) + 10n\log\left(\frac{r}{d_0}\right) \tag{III.29}$$

Hence, we have

$$\begin{aligned} \mu_r(r) &= P_{t,dB} - \left(\overline{PL}_{dB}(d_0) - 10n\log\left(\frac{r}{d_0}\right)\right) \\ &= P_{t,dB} - \left(\overline{PL}_{dB}(d_0) + 10n\log\left(\frac{r}{R}\right) + 10n\log\left(\frac{R}{d_0}\right)\right) \end{aligned} \tag{III.30}$$

Hence, (III.27) can be rewritten in the form

$$\Pr\left\{P_{r,dB}(r) > \gamma\right\} = Q\left(\frac{\gamma - \left[P_{t,dB} - \left(\overline{PL}_{dB}(d_0) + 10n\log\left(\frac{r}{R}\right) + 10n\log\left(\frac{R}{d_0}\right)\right)\right]}{\sigma}\right) \tag{III.31}$$

Let

$$a = \frac{\gamma - P_{t,dB} + \overline{PL}_{dB}(d_0) + 10n\log\left(R/d_0\right)}{\sigma} \tag{III.32}$$

$$\begin{aligned} b &= \frac{10n\log_{10}(e)}{\sigma} \\ &= \frac{10\log_{10}(e)}{\sigma/n} \end{aligned} \tag{III.33}$$

Then,

$$\Pr\left\{P_{r,dB}(r) > \gamma\right\} = Q\left(a + \frac{10\log\left(\frac{r}{R}\right)}{\sigma/n}\right) \tag{III.34}$$

Using (III.31), (III.32) and (III.33) in (III.26) yields

$$U(\gamma) = \frac{1}{2} - \frac{1}{R^2}\int_0^R r\left(1 - 2Q\left(\sqrt{2}\left[a + b\ln\left(\frac{r}{R}\right)\right]\right)\right)dr \tag{III.35}$$

Introducing the change of variables

$$z = \sqrt{2}\left[a + b\ln\left(\frac{r}{R}\right)\right] \tag{III.36}$$

equation (III.35) simplifies to

$$U(\gamma) = Q\left(a\sqrt{2}\right) + Q\left(\frac{1-ab}{b}\sqrt{2}\right)e^{\frac{1-2ab}{b^2}} \tag{III.37}$$

By choosing the signal level such that $\overline{P}_{r,dB}(R) = \gamma$, $U(\gamma)$ can be shown to be

$$U(\gamma) = \frac{1}{2}\left[1 + 2Q\left(\sqrt{\frac{2}{b^2}}\right)e^{\frac{1}{b^2}}\right] \tag{III.38}$$

Example III-4

Four received power measurements were taken at distances of 100 m, 200 m, 1 km, and 3 km from a transmitter. These measured values are given in the following table. It is assumed that the path loss for these measurements follows the model in equation (III.16), where $d_0 = 100\text{ m}$.

Distance From Transmitter (m)	Received Power (dBm)
100	0
200	-20
1000	-35
3000	-70

a. Find the minimum mean square error (MMSE) estimate for the path loss exponent n.
b. Calculate the standard deviation about the mean value.
c. Estimate the received power at $d = 2\text{ km}$ using the resulting model.
d. Predict the likelihood that the received signal level at 2 km will be greater than -60 dBm.

Solution

The MMSE estimate may be found using the following method. Let p_i be the received power at a distance d_i and let $\hat{p}_i$ be the estimate for p_i using the $(d/d_0)^n$ path loss model of equation (III.14). The sum of squared errors between the measured and estimated values is given by

$$J(n) = \sum_{i=1}^{4} (p_i - \hat{p}_i)^2$$

The value of n which minimizes the mean square error can be obtained by equating the derivative of $J(n)$ to zero, and then solving for n.

a. Using equation (III.15), $\hat{p}_i = p_i(d_0) - 10n\log_{10}(d_i/100)$. Note that $P_{r,dBm}(d_0) = 0$ dBm. Then

$$\hat{p}_1 = 0 - 10n\log_{10}(100/100) = 0$$

$$\hat{p}_2 = 0 - 10n\log_{10}(200/100) = -3n$$

$$\hat{p}_3 = 0 - 10n\log_{10}(1000/100) = -10n$$

$$\hat{p}_4 = 0 - 10n\log_{10}(3000/100) = -14.77n$$

The sum of squared errors is then given by:

$$J(n) = (0)^2 + (3n-20)^2 + (10n-35)^2 + (14.77n-70)^2$$
$$= 6525 - 2887.8n + 327.153n^2$$

$$\frac{dJ}{dn} = 654.306n - 2887.8$$

Setting this equal to zero, the value of n is obtained as $n \approx 4.4$.

b. The sample variance $\sigma^2 = J(n)/4 = 152.36$. Therefore, $\sigma = 6.17$ dB.

c. The estimate of the received power at $d = 2$ km is given by:

$$\hat{p}(2000\text{ m}) = 0 - 10(4.4)\log_{10}(2000/100) = -57.24\text{ dBm}$$

A Gaussian random variable having zero mean and $\sigma = 6.17$ could be added to this value to simulate random shadowing effects at $d = 2$ km.

d. The probability that the received signal level will be greater than -60 dBm is given by:

$$\Pr\left\{P_{r,dBm}(2000\text{ m}) > -60\text{ dBm}\right\} = Q\left(\frac{-60+57.24}{6.17}\right) = 0.674$$

III.4. Outdoor Propagation Models

Radio transmission in a mobile communications system often takes place over irregular terrain. The terrain profile of a particular area needs to be taken into account for estimating the path loss. The terrain profile may vary from a simple curved earth profile to a highly mountainous profile. The presence of trees, buildings, and other obstacles also must be taken into account. A number of propagation models are available 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, the methods vary 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.

III.5. Indoor Propagation Models

There is a great deal of interest in characterizing radio propagation inside buildings. The indoor radio channel differs from the traditional mobile radio channel in two aspects - the distances covered are much smaller, and the variability of the environment is much greater for a much smaller range of T-R separation distances. It has been observed that propagation within buildings is strongly influenced by specific features such as the layout of the building, the construction materials, and the building type.

Indoor radio propagation is dominated by the same mechanisms as outdoor: reflection, diffraction, and scattering. However, conditions are much more variable. For example, signal levels vary greatly depending on whether interior doors are open or closed inside a building. Where antennas are mounted also impacts large-scale propagation. Antennas mounted at desk level in a partitioned office receive vastly different signals than those mounted on the ceiling. Also, the smaller propagation distances make it more difficult to insure far-field radiation for all receiver locations and types of antennas.

***

IV. SMALL-SCALE FADING AND MULTIPATH

Small-scale fading, or simply fading, is used to describe the rapid fluctuation of the amplitude of a radio signal over a short period of time or travel distance, so that large-scale path loss effects may be ignored. Fading is caused by interference between two or more versions of the transmitted signal which arrive at the receiver at slightly different times. These waves, called multipath waves, combine at the receiver antenna to give a resultant signal which can vary widely in amplitude and phase, depending on the distribution of the intensity and relative propagation time of the waves and the bandwidth of the transmitted signal.

IV.1. Small-Scale Multipath Propagation

Multipath in the radio channel creates small-scale fading effects. The three most important effects are:

- Rapid changes in signal strength over a small travel distance or time interval.
- Random frequency modulation due to varying Doppler shifts on different multipath signals.
- Time dispersion (echoes) caused by multipath propagation delays.

In built-up urban areas, fading occurs because the height of the mobile antennas are well below the height of surrounding structures, so there is no single line-or-sight path to the base station. Even when a line-of-sight exists, multipath still occurs due to reflections from the ground and surrounding structures. The incoming radio waves arrive from different directions with different propagation delays. The signal received by the mobile at any point in space may consist of a large number of plane waves having randomly distributed amplitudes, phases, and angles of arrival. These multipath components combine vectorially at the receiver antenna, and can cause the signal received by the mobile to distort or fade. Even when a mobile receiver is stationary, the received signal may fade due to movement of surrounding objects in the radio channel.

If objects in the radio channel are static, and motion is considered to be only due to that of the mobile, then fading is purely a spatial phenomenon. The spatial variations of the resulting signal are seen as temporal variations by the receiver as it moves through the multipath field. Due to the constructive and destructive effects of multipath waves summing at various points in space, a receiver moving at high speed can pass through several fades in a small period of time. In a more serious case, a receiver may stop at a particular location at which the received signal is in a deep fade. Maintaining good communications can then become very difficult, although passing vehicles or people walking in the vicinity of the mobile can often disturb the field pattern, thereby diminishing the likelihood of the received signal remaining in a deep null for a long period of time. Antenna space diversity can prevent deep fading nulls.

Due to the relative motion between the mobile and the base station, each multipath wave experiences an apparent shift in frequency. The shift in received signal frequency due to motion is called the Doppler shift, and is directly proportional to the velocity and direction of motion of the mobile with respect to the direction of arrival of the received multi path wave.

IV.1.A. FACTORS INFLUENCING SMALL-SCALE FADING

Many physical factors in the radio propagation channel influence small scale fading. These include the following:

- ***Multipath propagation***: Multipath propagation often lengthens the time required for the baseband portion of the signal to reach the receiver which can cause signal smearing due to intersymbol interference.
- ***Speed of the mobile***: The relative motion between the base station and the mobile results in random frequency modulation due to different Doppler shifts on each of the multi path components.
- ***Speed of surrounding objects***: If objects in the radio channel are in motion, they induce a time varying Doppler shift on multipath components. If the surrounding objects move at a greater rate than the mobile, then this effect dominates the small-scale fading. Otherwise, motion of surrounding objects may be ignored, and only the speed of the mobile needs to be considered.
- ***Transmission bandwidth of the signal***: If the transmitted radio signal bandwidth is greater than the "bandwidth" of the multipath channel, the received signal will be distorted.

IV.1.B. DOPPLER SHIFT

Consider a mobile moving at a constant velocity v, along a path segment having length d between points X and Y, while it receives signals from a remote source S. See Figure IV.1.

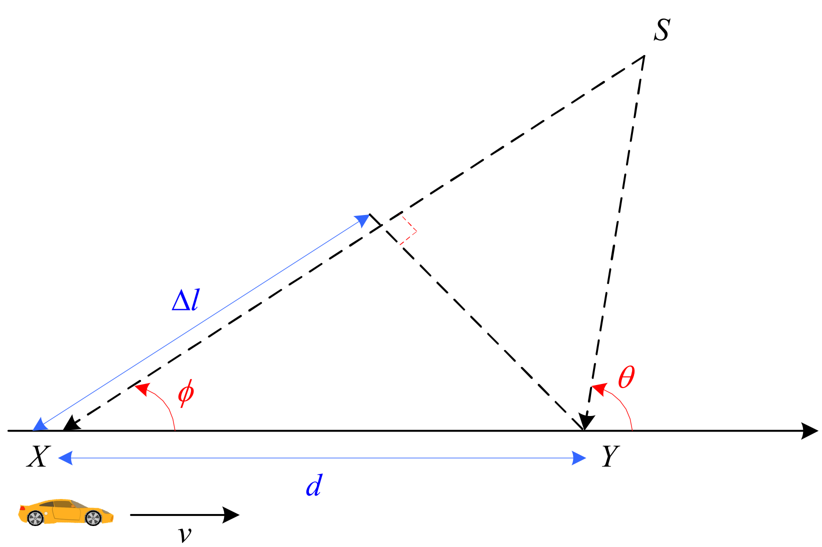

Figure IV.1

The difference in path lengths traveled by the wave from source S to the mobile at points X and Y is equal to

$$\begin{aligned} \Delta l &= d\cos\phi \\ &\approx d\cos\theta \\ &= v\Delta t\cos\theta \end{aligned} \tag{IV.1}$$

where Δt is the time required for the mobile to travel from X to Y, and θ is assumed to be the same at points X and Y since the source is assumed to be very far away. The phase change in the received signal due to the difference in path lengths is therefore

$$\begin{aligned} \Delta\phi &= \frac{2\pi\Delta l}{\lambda} \\ &= \frac{2\pi v\Delta t}{\lambda}\cos\theta \end{aligned} \tag{IV.2}$$

The apparent change in frequency, or Doppler shift, is given by

$$\begin{aligned} f_d &= \frac{1}{2\pi}\frac{\Delta\phi}{\Delta t} \\ &= \frac{v}{\lambda}\cos\theta \end{aligned} \tag{IV.3}$$

Equation (IV.3) relates the Doppler shift to the mobile velocity and the spatial angle between the direction of motion of the mobile and the direction of arrival of the wave. It can be seen from equation (IV.3) that if the mobile is moving toward the direction of arrival of the wave, the Doppler shift is positive (i.e., the apparent received frequency is increased), and if the mobile is moving away from the direction of arrival of the wave, the Doppler shift is negative (i.e., the apparent received frequency is decreased). Multipath components from a CW signal which arrive from different directions contribute to Doppler spreading of the received signal, thus increasing the signal bandwidth.

Example IV-1

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

a. directly towards 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

$$\lambda = \frac{c}{f_c} = \frac{3\times 10^8}{1850\times 10^6} = 0.162 \text{ m}$$

$$v = 60\frac{\text{mi}}{h}\frac{\text{h}}{3600 \text{ s}}\frac{1609.344 \text{ m}}{1 \text{ mi}} = 26.822 \text{ m/s}$$

a. $\cos\theta = 1$. $f_d = \frac{26.822}{0.162} = 165.568$ Hz. $f = f_c + f_d = 1850.000166$ MHz.

b. $\cos\theta = -1$. $f = f_c - f_d = 1849.999834$ MHz.

c. $\cos\theta = 0$. $f = f_c = 1850$ MHz.

IV.2. Impulse Response Model of a Multipath Channel

The small-scale variations of a mobile radio signal can be directly related to the impulse response of the mobile radio channel. The impulse response is a channel characterization and contains all information necessary to simulate or analyze any type of radio transmission through the channel. This stems from the fact that a mobile radio channel may be modeled as a linear filter with a time varying impulse response, where the time variation is due to receiver motion in space. The filtering nature of the channel is caused by the summation of amplitudes and delays of the multiple arriving waves at any instant of time. The impulse response is a useful characterization of the channel, since it may be used to predict and compare the performance of many different mobile communication systems and transmission bandwidths for a particular mobile channel condition.

To show that a mobile radio channel may be modeled as a linear filter with a time varying impulse response, consider the case where time variation is due strictly to receiver motion in space. This is shown in Figure IV.2.

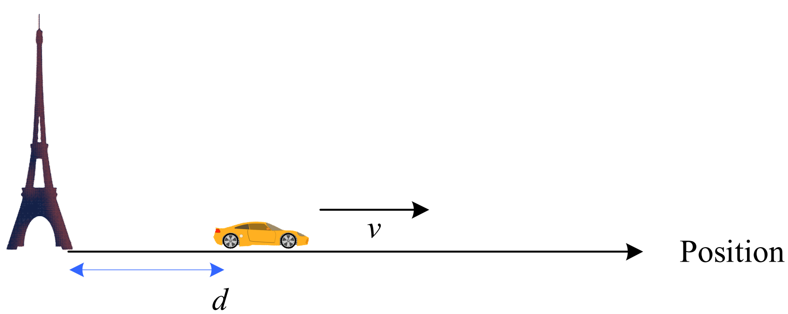

Figure IV.2

In Figure IV.2, the receiver moves along the ground at some constant velocity v. For a fixed position d, the channel between the transmitter and the receiver can be modeled as a linear time invariant system. However, due to the different multipath waves which have propagation delays that vary over different spatial locations of the receiver, the impulse response of the channel should be a function of the position of the receiver. That is, the channel impulse response can be expressed as $h(d,t)$. Let $x(t)$ represent the transmitted signal, then the received signal $y(d,t)$ at position d can be expressed as a convolution of $x(t)$ with $h(d,t)$ as follows

$$\begin{aligned} y(d,t) &= x(t) * h(d,t) \\ &= \int_{-\infty}^{\infty} x(\tau)h(d,t-\tau)d\tau \end{aligned} \tag{IV.4}$$

For a causal system, $h(d,t)=0$ for $t<0$, thus equation (IV.4) reduces to

$$y(d,t) = \int_{-\infty}^{t} x(\tau)h(d,t-\tau)d\tau \tag{IV.5}$$

Since the receiver moves along the ground at a constant velocity v, the position of the receiver can by expressed as

$$d = vt \tag{IV.6}$$

Substituting (IV.6) in (IV.5), we obtain

$$y(vt,t) = \int_{-\infty}^{t} x(\tau)h(vt,t-\tau)d\tau \tag{IV.7}$$

Since v is a constant, $y(vt,t)$ is just a function of t. Therefore, equation (IV.7) can be expressed as

$$y(t) = \int_{-\infty}^{t} x(\tau)h(vt,t-\tau)d\tau \tag{IV.8}$$

Since v may be assumed constant over a short time (or distance) interval, we may let $x(t)$ represent the transmitted bandpass waveform, $y(t)$ the received waveform, and $h(t,\tau)$ the impulse response of the time varying multipath radio channel. The impulse response $h(t,\tau)$ completely characterizes the channel and is a function of both t and τ. The variable t represents the time variations due to motion, whereas τ represents the channel multipath delay for a fixed value of the variable t. The received signal $y(t)$ can be expressed as a convolution of the transmitted signal $x(t)$ with the channel impulse response.

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t,\tau) d\tau \tag{IV.9}$$

If the multipath channel is assumed to be a bandlimited bandpass channel, which is reasonable, then $h(t,\tau)$ may be equivalently described by a complex baseband impulse response $h_b(t,\tau)$ with the input and output being the complex envelope representations of the transmitted and received signals, respectively. That is

$$r(t) = \frac{1}{2} c(t) * h_b(t,\tau) \tag{IV.10}$$

where

$$x(t) = \mathrm{Re}\left\{ c(t) e^{j2\pi f_c t} \right\} \tag{IV.11}$$

$$y(t) = \mathrm{Re}\left\{ r(t) e^{j2\pi f_c t} \right\} \tag{IV.12}$$

Note that based on the above, we have

$$\overline{x^2(t)} = \frac{1}{2} \overline{|c(t)|^2} \tag{IV.13}$$

It is useful to discretize the multipath delay axis τ of the impulse response into equal time delay segments called excess delay bins, where each bin has a time delay width equal to $\tau_{i+1} - \tau_i$, where $\tau_0 = 0$ represents the first arriving signal at the receiver. Let, for $i = 0,1,\cdots,N-1$,

$$\tau_i = i\Delta\tau \tag{IV.14}$$

where N represents the total number of possible equally-spaced multipath components, including the first arriving component. Any number of multipath signals received within the ith bin are represented by a single resolvable multipath component having delay τ_i. This technique of quantizing the delay bins determines the time delay resolution of the channel model, and the useful frequency span of the model can be shown to be $1/(2\Delta\tau)$. Excess delay is the relative delay of the ith multipath component as compared to the first arriving component and is given by τ_i. The maximum excess delay of the channel is given by $N\Delta\tau$.

Since the received signal in a multipath channel consists of a series of attenuated, time-delayed, phase shifted replicas of the transmitted signal, the baseband impulse response of a multipath channel can be expressed as

$$h_b(t,\tau)=\sum_{i=0}^{N-1} a_i(t,\tau)e^{j2\pi f_c\tau_i(t)}e^{j\phi_i(t,\tau)}\delta(\tau-\tau_i(t)) \tag{IV.15}$$

where $a_i(t,\tau)$ and $\tau_i(t)$ are the real amplitudes and excess delays, respectively, of ith multipath component at time t. The phase term $2\pi f_c\tau_i(t)$ in (IV.15) represents the phase shift due to free space propagation of the ith multipath component. The phase term $\phi_i(t,\tau)$ in (IV.15) represents additional phase shifts which are encountered in the channel. In general, the phase term is simply represented by a single variable $\theta_i(t,\tau)$ which lumps together all the mechanisms for phase shifts of a single multipath component within the ith excess delay bin.

If the channel impulse response is assumed to be time invariant, at least over a small-scale time or distance interval, then the channel impulse response may be simplified as

$$h_b(\tau)=\sum_{i=0}^{N-1} a_i e^{-j\theta_i}\delta(\tau-\tau_i) \tag{IV.16}$$

For small-scale channel modeling, the power delay profile of the channel is found by taking the spatial average of $|h_b(t,\tau)|^2$ over the area. By making several measurements of $|h_b(t,\tau)|^2$ in different locations, it is possible to build an ensemble of power delay profiles, each one representing a possible small-scale multipath channel state.

In actual wireless communication systems, the impulse response of a multipath channel is measured in the field using channel sounding techniques. We now consider two extreme channel sounding cases as a means of demonstrating how small scale fading behaves quite differently for two signals with different bandwidths in the identical multipath channel.

Consider a pulsed, transmitted RF signal of the form

$$x(t)=\mathrm{Re}\left\{p(t)e^{j2\pi f_c t}\right\} \tag{IV.17}$$

where $p(t)$ is a repetitive baseband pulse train with a very narrow pulse width T_{bb} and a repetition period T_{REP} which is much greater than the maximum measured excess delay $\tau_{\max}$ in the channel. Let, for $0\le t<T_{REP}$ and $T_{bb}<<\tau_{\max}$,

$$p(t)=\begin{cases}2\sqrt{\dfrac{\tau_{\max}}{T_{bb}}}, & 0\le t<T_{bb}\\ 0, & \text{otherwise}\end{cases} \tag{IV.18}$$

Then,

$$r(t) = \frac{1}{2}\sum_{i=0}^{N-1} a_i e^{-j\theta_i} p(t-\tau_i) \tag{IV.19}$$

The quantity $|r(t_0)|^2$ is called the instantaneous multipath power delay profile of the channel, and is equal to the energy received over the time duration of the multipath delay divided by $\tau_{\max}$. It can be shown that

$$|r(t_0)|^2 = \sum_{i=0}^{N-1} a_k^2(t_0) \tag{IV.20}$$

Assuming that the received power from the multipath components forms a random process where each component has a random amplitude and phase at any time t, the average small-scale received power for the wideband probe is found as

$$\overline{P}_{WB} = \sum_{i=0}^{N-1} \overline{a_k^2(t_0)} \tag{IV.21}$$

It can be seen from equations (IV.20) and (IV.21) that the small-scale received power is simply the sum of the powers received in each multipath component. In practice, the amplitudes of individual multipath components do not fluctuate widely in a local area. Thus, the received power of a wideband signal such as $p(t)$ does not fluctuate significantly when a receiver is moved about a local area.

Now, instead of a pulse, consider a CW signal which is transmitted into the exact same channel, and let the complex envelope be given by $c(t) = 2$. Then, the instantaneous complex envelope of the received signal is given by

$$r(t) = \sum_{i=0}^{N-1} a_i e^{j\theta_i(t,\tau)} \tag{IV.22}$$

and the instantaneous power is given by

$$|r(t)|^2 = \left|\sum_{i=0}^{N-1} a_i e^{j\theta_i(t,\tau)}\right|^2 \tag{IV.23}$$

As the receiver is moved over a local area, the channel changes, and the received signal strength will vary at a rate governed by the fluctuations of a_i and θ_i. As mentioned earlier, a_i varies little over local areas, but θ_i will vary greatly due to changes in propagation distance over space, resulting in large fluctuations of $r(t)$ as the receiver is moved over small distances (on the order of a wavelength). That is, since $r(t)$ is the sum of the individual multipath components, the instantaneous phases of the multipath components cause the large fluctuations which typify small-scale fading for CW signals. The average received power over a local area is then given by

$$\begin{aligned}\bar{P}_{CW} &= \mathrm{E}_{\underline{a},\underline{\theta}}\left[\left|\sum_{i=0}^{N-1} a_i e^{j\theta_i}\right|^2\right] \\ &\approx \sum_{i=0}^{N-1} \overline{a_i^2} + 2\sum_{i=0}^{N-1}\sum_{\substack{j=0 \\ j\neq i}}^{N-1} r_{ij}\,\overline{\cos(\theta_i - \theta_j)}\end{aligned} \tag{IV.24}$$

where r_{ij} is the path amplitude correlation coefficient defined to be

$$r_{ij} = \mathrm{E}\left[a_i a_j\right] \tag{IV.25}$$

If $\overline{\cos(\theta_i - \theta_j)} = 0\ \forall i \neq j$ or $r_{ij} = 0\ \forall i \neq j$, then the average power for a CW signal is equivalent to the average received power for a wideband signal in a small-scale region. This can occur when either the multipath phases are identically and independently distributed (i.i.d uniform) over $\left[0, 2\pi\right]$ or when the path amplitudes are uncorrelated. The i.i.d uniform distribution of θ is a valid assumption since multipath components traverse differential path lengths that measure hundreds of wavelengths and are likely to arrive with random phases. If for some reason it is believed that the phases are not independent, the average wideband power and average CW power will still be equal if the paths have uncorrelated amplitudes. However, if the phases of the paths are dependent upon each other, then the amplitudes are likely to be correlated, since the same mechanism which affects the path phases is likely to also affect the amplitudes. This situation is highly unlikely at transmission frequencies used in wireless mobile systems.

Thus it is seen that the received local ensemble average power of wideband and narrowband signals are equivalent. When the transmitted signal has a bandwidth much greater than the bandwidth of the channel, then the multipath structure is completely resolved by the received signal at any time, and the received power varies very little since the individual multipath amplitudes do not change rapidly over a local area. However, if the transmitted signal has a very narrow bandwidth (e.g., the baseband signal has a duration greater than the excess delay of the channel), then multipath is not resolved by the received signal, and large signal fluctuations (fading) occur at the receiver due to the phase shifts of the many unresolved multipath components.

Example IV-2

Assume a discrete channel impulse response is used to model urban 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 multi path bins is fixed at 64, find

(a) $\Delta\tau$
(b) the maximum bandwidth which the two models can accurately represent.

Repeat for an indoor channel model with excess delays as large as 500 ns.

Solution

(c) For urban radio channels, $\Delta\tau = \frac{\tau_N}{N} = \frac{100}{64} = 1.5625\ \mu\text{s}$. For microcellular channels, $\Delta\tau = \frac{4000}{64} = 62.5\ \text{ns}$. For indoor channels, $\Delta\tau = \frac{500}{64} = 7.8125\ \text{ns}$.

(d) For urban radio channels, $\frac{1}{2\Delta\tau} = \frac{1}{1.5625\times 10^{-6}} = 0.32\ \text{MHz}$. For microcellular channels, $\frac{1}{2\Delta\tau} = \frac{1}{62.5\times 10^{-9}} = 8\ \text{MHz}$. For indoor channels, $\frac{1}{2\Delta\tau} = \frac{1}{7.8125\times 10^{-9}} = 64\ \text{MHz}$.

Example IV-3

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. The second component is 3 dB weaker than the first component and is assumed to arrive at , $\tau = 1\ \mu\text{s}$ with an initial phase of 0°. The mobile moves directly towards the direction of arrival of the first component and directly away from the direction of arrival of the second component.

(e) Compute the narrowband instantaneous power at time intervals of 0.1 s from 0 to 0.5 s.
(f) Compute the average narrowband power received over the observation interval.
(g) Compare the average widehand received power over the observation interval.

Solution

$$\lambda = \frac{3\times 10^8}{1000\times 10^6} = 0.3\ \text{m}.$$

$-70\ \text{dBm} \to 100\ \text{pW}$. $-73\ \text{dBm} \to 50\ \text{pW}$.

(h) $$|r(0)|^2 = \left|\sum_{i=0}^{1} a_i e^{j\theta_i(0)}\right|^2 = \left|\sqrt{100}e^{j0} + \sqrt{50}e^{j0}\right|^2 = 291\ \text{pW}.$$

Every 0.1 s, the phase changes by $\frac{2\pi vt}{\lambda} = \frac{2\pi(10)(0.1)}{0.3} = \frac{10}{3}(2\pi) \to \frac{1}{3}(2\pi) \to 120^\circ$.

Since the mobile moves towards 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.

$$|r(0.1)|^2 = \left|\sum_{i=0}^{1} a_i e^{j\theta_i(0.1)}\right|^2 = \left|\sqrt{100}e^{j\frac{2\pi}{3}} + \sqrt{50}e^{-j\frac{2\pi}{3}}\right|^2 = 78.2 \text{ pW}.$$

$$|r(0.2)|^2 = \left|\sum_{i=0}^{1} a_i e^{j\theta_i(0.2)}\right|^2 = \left|\sqrt{100}e^{j(2)\frac{2\pi}{3}} + \sqrt{50}e^{-j(2)\frac{2\pi}{3}}\right|^2 = 81.5 \text{ pW}.$$

$$|r(0.3)|^2 = \left|\sum_{i=0}^{1} a_i e^{j\theta_i(0.3)}\right|^2 = \left|\sqrt{100}e^{j(3)\frac{2\pi}{3}} + \sqrt{50}e^{-j(3)\frac{2\pi}{3}}\right|^2 = 291 \text{ pW}.$$

$$|r(0.4)|^2 = \left|\sum_{i=0}^{1} a_i e^{j\theta_i(0.4)}\right|^2 = \left|\sqrt{100}e^{j(4)\frac{2\pi}{3}} + \sqrt{50}e^{-j(4)\frac{2\pi}{3}}\right|^2 = 78.2 \text{ pW}.$$

$$|r(0.5)|^2 = \left|\sum_{i=0}^{1} a_i e^{j\theta_i(0.5)}\right|^2 = \left|\sqrt{100}e^{j(5)\frac{2\pi}{3}} + \sqrt{50}e^{-j(5)\frac{2\pi}{3}}\right|^2 = 81.5 \text{ pW}.$$

(i) Average narrowband power received over the observation interval= $\frac{291(2)+78.2(2)+81.5(2)}{6} = 150.233 \text{ pW}$.

(j) $\overline{P}_{WB} = \sum_{i=0}^{1} \overline{a_i^2} = 100 + 50 = 150 \text{ pW}$.

As can be seen, the narrowband and wideband received power are virtually identical when averaged over 0.5 s (or 5 m). While the CW signal fades over the observation interval, the wideband signal power remains constant.

***