LECTURE NOTES

ON

WIRELESS COMMUNICATION AND NETWORKS

ELECTRONICS & COMMUNICATION ENGINEERING

(BTEC-601-180

GGS COLLEGE OF MODERN TECHNOLOGY, KHARAR

SYLLABUS

UNIT-I

The Cellular Concept-System Design Fundamentals: Introduction, Frequency Reuse, Channel Assignment Strategies, Handoff Strategies- Prioritizing Handoffs, Practical Handoff Considerations, Interference and system capacity — Co channel Interference and system capacity, Channel planning for Wireless Systems, Adjacent Channel interference , Power Control for Reducing interference, Trunking and Grade of Service, Improving Coverage & Capacity in Cellular Systems- Cell Splitting, Sectoring.

UNIT—II

Mobile Radio Propagation: Large-Scale Path Loss: Introduction to Radio Wave Propagation, Free Space Propagation Model, Relating Power to Electric Field, The Three Basic Propagation Mechanisms, Reflection- Reflection from Dielectrics, Brewster Angle, Reflection from prefect conductors, Ground Reflection (Two-Ray) Model, Diffraction-Fresnel Zone Geometry, Knife-edge Diffraction Model, Multiple knife-edge Diffraction, Scattering, Outdoor Propagation Models- Longley-Ryce Model, Okumura Model, Hata Model, PCS Extension to Hata Model, Walfisch and Bertoni Model, Wideband PCS Microcell Model, lndoor Propagation Models-Partition losses (Same Floor), Partition losses between Floors, Log-distance path loss model, Ericsson Multiple Breakpoint Model, Attenuation Factor Model, Signal penetration into buildings, Ray Tracing and Site Specific Modeling.

.

UNIT —III

Mobile Radio Propagation: Small —Scale Fading and Multipath: Small Scale Multipath propagation-Factors influencing small scale fading, Doppler shift, Impulse Response Model of a multipath channel-Relationship between Bandwidth and Received power, Small-Scale Multipath Measurements-Direct RF Pulse System, Spread Spectrum Sliding Correlator Channel Sounding, Frequency Domain Channels Sounding, Parameters of Mobile Multipath Channels-Time Dispersion Parameters, Coherence Bandwidth, Doppler Spread and Coherence Time, Types of Small-Scale Fading-Fading effects Due to Multipath Time Delay Spread, Flat fading, Frequency selective fading, Fading effects Due to Doppler Spread-Fast fading, slow fading, Statistical Models for multipath Fading Channels-Clarke's model for flat fading, spectral shape due to Doppler spread in Clarke's model, Simulation of Clarke and Gans Fading Model, Level crossing and fading statistics, Two-ray Rayleigh Fading Model.

UNIT -IV

Equalization and Diversity: Introduction, Fundamentals of Equalization, Training A Generic Adaptive Equalizer, Equalizers in a communication Receiver, Linear Equalizers, Non-linear Equalization-Decision Feedback Equalization (DFE), Maximum Likelihood Sequence Estimation (MLSE) Equalizer, Algorithms for adaptive equalization-Zero Forcing Algorithm, Least Mean Square Algorithm, Recursive least squares algorithm. Diversity Techniques-Derivation of selection Diversity improvement, Derivation of Maximal Ratio Combining improvement, Practical Space Diversity Consideration-Selection Diversity, Feedback or Scanning Diversity, Maximal Ratio Combining, Equal Gain Combining, Polarization Diversity, Frequency Diversity, Time Diversity, RAKE Receiver.

UNIT -V

Wireless Networks: Introduction to wireless Networks, Advantages and disadvantages of Wireless Local Area Networks, WLAN Topologies, WLAN Standard IEEE 802.11 ,IEEE 802.11 Medium Access Control, Comparision

of IEEE 802.11 a,b,g and n standards, IEEE 802.16 and its enhancements, Wireless PANs, Hiper Lan, WLL.

TEXT BOOKS

1. Wireless Communications, Principles, Practice — Theodore, S.Rappaport, 2nd Ed., 2002, PHI.
2. Wireless Communications-Andrea Goldsmith, 2005 Cambridge University Press.
3. Mobile Cellular Communication — Gottapu Sasibhushana Rao, Pearson Education, 2012.

REFERENCE BOOKS

1. Principles of Wireless Networks — Kaveh Pah Laven and P. Krishna Murthy, 2002, PE
2. Wireless Digital Communications — Kamilo Feher, 1999, PHI.
3. Wireless Communication and Networking — William Stallings, 2003,PHI.
4. Wireless Communication — Upen Dalal, Oxford Univ. Press
5. Wireless Communications and Networking — Vijay K. Gary, Elsevier.

UNIT 1

Cellular System

Introduction

In the older mobile radio systems, single high power transmitter was used to provide coverage in the entire area. Although this technique provided a good coverage, but it was virtually impossible in this technique to re-use the same radio channels in the system, and any effort to re-use the radio channels would result in interference. Therefore, in order to improve the performance of a wireless system with the rise in the demand for the services, a cellular concept was later proposed. This chapter will examine several parameters related with the cellular concept.

The Cellular Concept

The design aim of early mobile wireless communication systems was to get a huge coverage area with a single, high-power transmitter and an antenna installed on a giant tower, transmitting a data on a single frequency. Although this method accomplished a good coverage, but it also means that it was practically not possible to reuse the same frequency all over the system, because any effort to reuse the same frequency would result in interference.

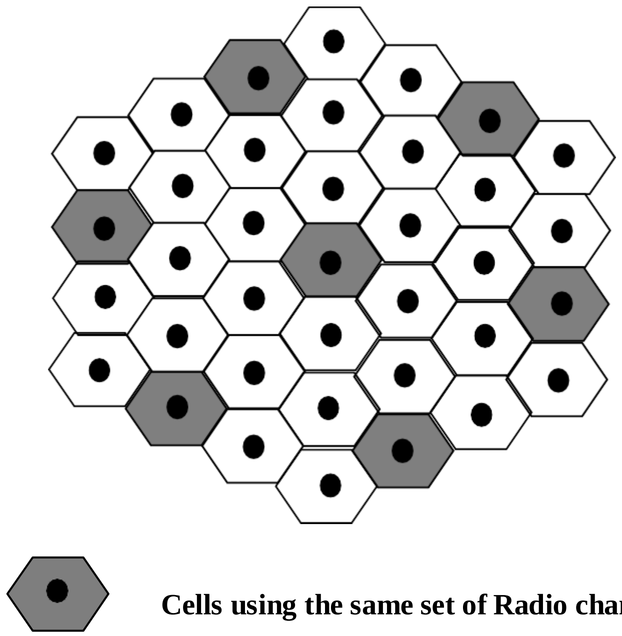
The cellular concept was a major breakthrough in order to solve the problems of limited user capacity and spectral congestion. Cellular system provides high capacity with a limited frequency spectrum without making any major technological changes [1]. It is a system-level idea in which a single high-power transmitter is replaced with multiple low-power transmitters, and small segment of the service area is being covered by each transmitter, which is referred to as a cell. Each base station (transmitter) is allocated a part of the total number of channels present in the whole system, and different groups of radio channels are allocated to the neighboring base stations so that all the channels present in the system are allocated to a moderately small number of neighboring base stations.

The mobile transceivers (also called mobile phones, handsets, mobile terminals or mobile stations) exchange radio signals with any number of base stations. Mobile phones are not linked to a specific base station, but can utilize any one of the base stations put up by the company. Multiple base stations covers the entire region in such a way that the user can move around and phone call can be carried on without interruption, possibly using more than one base station. The procedure of changing a base station at cell boundaries is called *handover*. Communication from the Mobile Station (MS) or mobile phones to the Base Station (BS) happens on an uplink channel also called reverse link, and downlink channel or forward link is used for communication from BS to MS. To maintain a bidirectional communication between a MS and BS, transmission resources must be offered in both the uplink and downlink directions. This can take place either using Frequency-Division Duplex (FDD), in which separate frequencies are used for both uplink and downlink channels, or through Time-Division Duplex (TDD), where uplink and downlink communications take place on the same frequency, but vary in time.

FDD is the most efficient technique if traffic is symmetric, and FDD has also made the task of radio planning more efficient and easier, because no interference takes place between base stations as they transmit and receive data on different frequencies. In case of an asymmetry in the uplink and downlink data speed, the TDD performs better than FDD. As the uplink data rate increases, extra bandwidth is dynamically allocated to that, and as the data rate decreases, the allotted bandwidth is taken away.

Some of the important cellular concepts are:

- Frequency reuse
- Channel Allocation
- Handoff
- Interference and system capacity
- Trunking and grade of service
- Improving coverage and capacity



Cells using the same set of Radio channels

Fig. 1.1: Cellular Network

Frequency Reuse

Conventional communication systems faced the problems of limited service area capability and ineffective radio spectrum utilization. This is because these systems are generally designed to provide service in an autonomous geographic region and by selecting radio channels from a particular frequency band. On the other hand, the present mobile communication systems are designed to offer a wide coverage area and high grade of service. These systems are also expected to provide a continuous communication through an efficient utilization of available radio spectrum. Therefore, the design of mobile radio network must satisfy the following objectives i.e., providing continuous service, and wide service area, while efficiently using the radio spectrum.

In order to achieve these objectives, the present mobile systems use cellular networks which depend more on an intelligent channel allocation and reuse of channels throughout the region . Each base station is allocated a set of radio channels, which are to be used in a geographic area called a *cell*. Base stations in the neighboring cells are allocated radio channel sets, which are entirely different. The antennas of base station antennas are designed to get the required coverage within the specific cell. By restricting the coverage area of a base station to within the cell boundaries, the same set of radio channels can be used in the different cells that are separated from each other by distances which are large enough in order to maintain interference levels within limits. The procedure of radio sets selection and allocation to all the base stations present within a network is called *frequency reuse* .

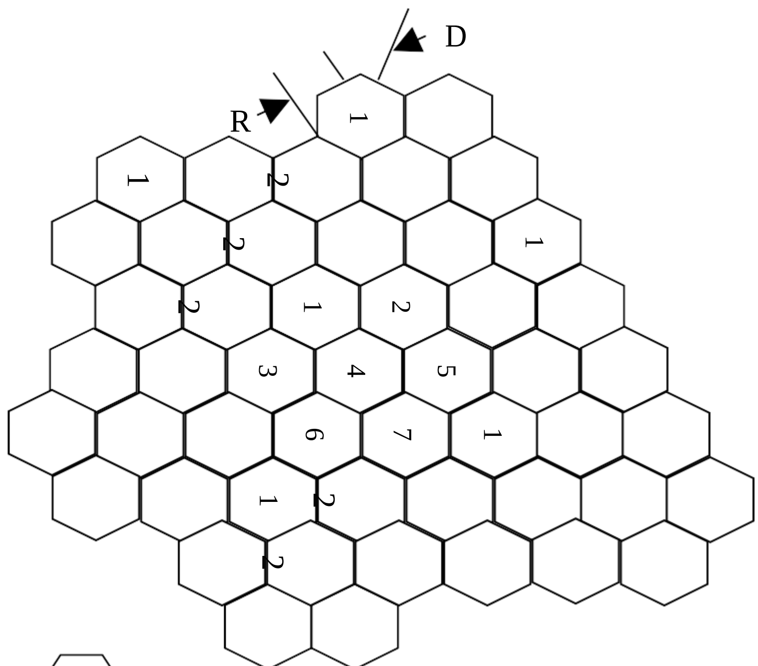
Fig. 1.1 shows the frequency reuse concept in a cell in a cellular network, in which cells utilize the same set of radio channels. The frequency reuse plan indicates where different radio channels are used. The hexagonal shape of cell is purely theoretical and is a simple model of radio coverage for each base station, although it has been globally adopted as the hexagon permits the easy analysis of a cellular system. The radio coverage of a cell can be calculated from field measurements. Although the actual radio coverage is very amorphous, a natural shape of a cell is required for an organized system design. While a circle is generally chosen to represent the coverage area of BS, but the circles present in the neighborhood cannot cover the entire region without leaving gaps or overlapping regions. Therefore, when selecting the cell shapes which can cover the entire geographical region without overlapping, there are three choices possible: a hexagon, square, and triangle. A particular design of the cell chosen in order to serve the weakest mobiles within the coverage area, and these are generally present at the cell boundaries of the cell. As hexagon covers the largest area from the center of a polygon to its farthest point, therefore, hexagon geometry can cover the entire geographic region to the fullest with minimum number of cells. When hexagon geometry is used to cover the entire geographic area, the base stations are either put up at the center of the cell, these cells are also called center excited cells or at the three of the six vertices (edge excited cells). Generally, center excited cells use omni-directional antennas and corner excited cells use directional antennas, but practically considerations for placing base stations are not exactly the same as they are shown in the hexagonal layouts.

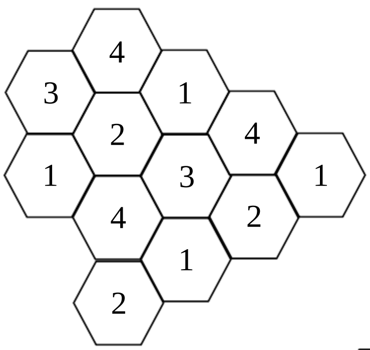
Channel Reuse Schemes

The radio channel reuse model can be used in the time and space domain. Channel reuse in the time domain turns out to be occupation of same frequency in different time slots and is also called Time Division Multiplexing. Channel reuse in the space domain is categorized into:

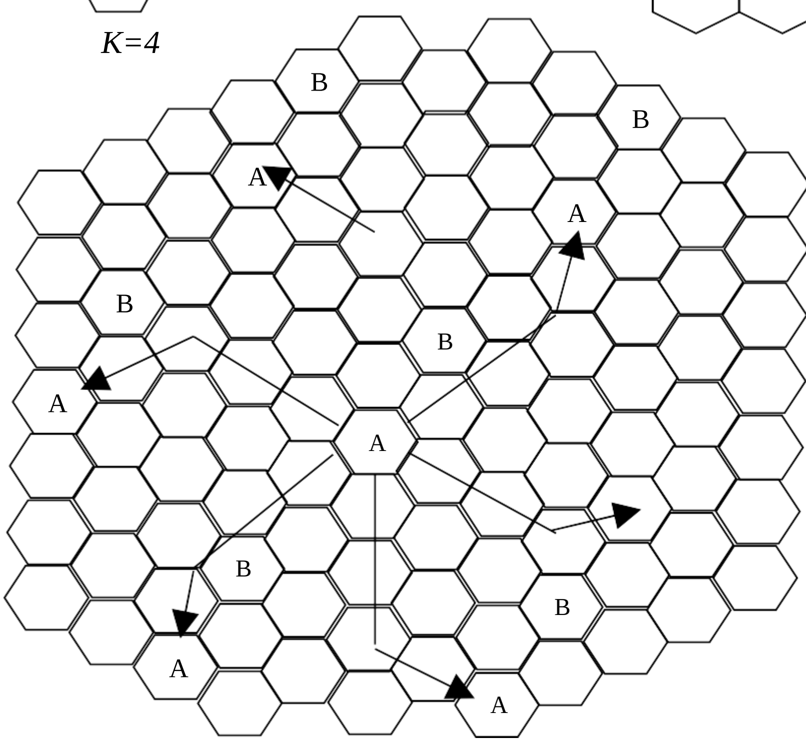
a) Same channel is allocated in two different areas, e.g. AM and FM radio stations using same channels in two different cities.

b) Same channel is frequently used in same area and in one system the scheme used is cellular systems. The entire spectrum is then divided into K reuse sets.





K=4



q=D/R=4
K=7

Fig. 1.2: K-Cell Reuse Pattern

Locating Co-channel Cells in a Cellular Network

Cells, which use the same set of channels, are called co-channels cells. For determining the location of co-channel cell present in the neighborhood, two shift parameters *i* and *j* are used where i and j are separated by 60^0, as shown in Fig. 1.3 below. The shift parameters can have any value 0, 1, 2 ...,n.

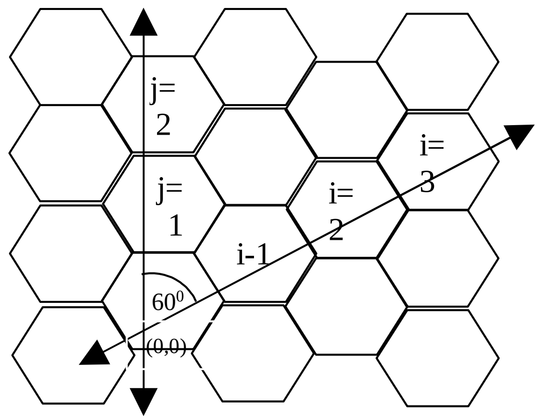


Fig. 1.3: Shift Parameters *i* and *j* in Hexagonal Network

To find the location of nearest co-channel cell, mark the center of the cell as (0, 0) for which co-channel cells are to be located. Define the unit distance as the distance of centre of two adjacent cells, and follow the two steps given below:

Step 1**:** Move *i* number of cells along *i* axis

Step 2**:** Turn 60^0 anti-clockwise and move *j* number of cells

The technique of locating co-channel cells using the preceding procedure is shown in Fig. 2.4 for *i*=3 and *j*=2. The shift parameters *i* and *j* measures the number of neighboring cells between co-channel cells.

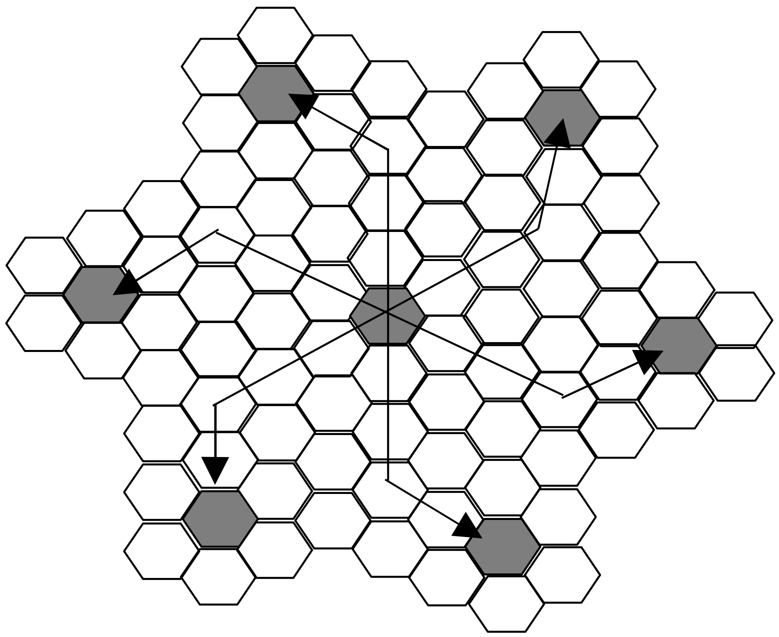


Fig. 1.4: Locating Co-channel Cells when $i=3$ & $j=2$

The relationship between cluster size K and shift parameters i & j, is given below:

Let *'R'* be the distance between the center of a regular hexagon to any of its vertex. A regular hexagon is one whose all sides are also of equal length i.e. *'R'*. Let *'d'* be the distance between the centre of two neighboring hexagons, and following steps are followed while calculating the size of a cluster *'K'*.

Step 1: To show that $d = \sqrt{3}R$

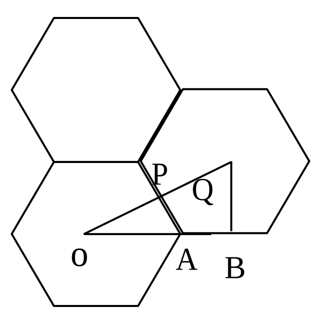


Fig. 1.5: Distance Between two adjacent cells

From the geometry of the Fig. 2.5, OA = R and AB = R/2 (2.1)

$$\text{Then, } OB = OA + AB = R + R/2 = 3R/2 \tag{2.2}$$

Then, in right-angled Δ OAP

$$OP = OA \sin 60^0 = \left(\sqrt{3}/2\right)R \tag{1.3}$$

Let the distance between the centers of two neighboring hexagonal cells, OQ, be denoted by 'd', then,

$$OQ = OP + PQ \text{ (where } OP = PQ)$$

Therefore,

$$d = \left[\left(\sqrt{3}/2\right)R\right] + \left[\left(\sqrt{3}/2\right)R\right]$$

$$\text{Hence, } d = \sqrt{3}R \tag{1.4}$$

Step 2: Area of a small hexagon, $A_{small\ hexagon}$ W

The area of a hexagonal cell with radius R is given as

$$A_{small\ hexagon} = \left(3\sqrt{3}/2\right) \times R^2 \tag{1.5}$$

Step 3: To find the relation between D, d and shift parameters

Let 'D' be the distance between the center of a particular cell under consideration to the centre of the nearest co-channel cell.

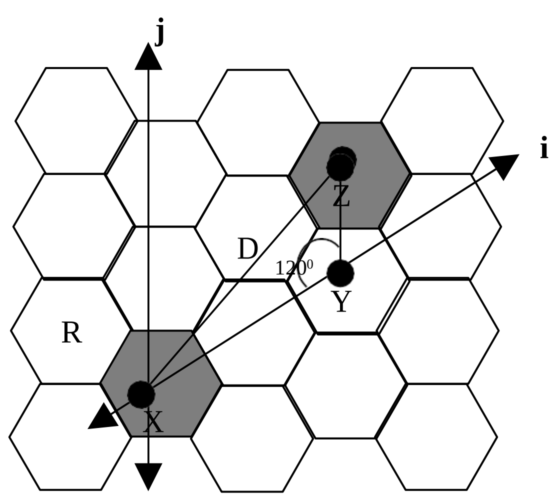


Fig. 1.6: Relationship Between K and Shift Parameters *(i & j)*

Using cosine formula Δ XYZ in Fig. 2.6, we have

$$XZ^2 = XY^2 + YZ^2 - 2 \times XY \times YZ \cos 120^0$$

$$\text{or, } D^2 = (i \times d)^2 + (j \times d)^2 - 2 \times (i \times d) \times (j \times d) \cos 120^0$$

$$D^2 = (i \times d)^2 + (j \times d)^2 - 2 \times (i \times d) \times (j \times d) \times (-1/2)$$

$$D^2 = (i \times d)^2 + (j \times d)^2 + (i \times d) \times (j \times d)$$

$$D^2 = d^2 (i^2 + j^2 + i \times j) \tag{1.6}$$

$$D^2 = 3 \times R^2 \times (i^2 + j^2 + i \times j) \tag{1.7}$$

Step 4: To find the area of a large hexagon, $A_{large\ hexagon}$

By joining the centers of the six nearest neighboring co-channel cells, a large hexagon is formed with radius equal to D, which is also the co-channel cell separation. Refer Fig. 1.7.

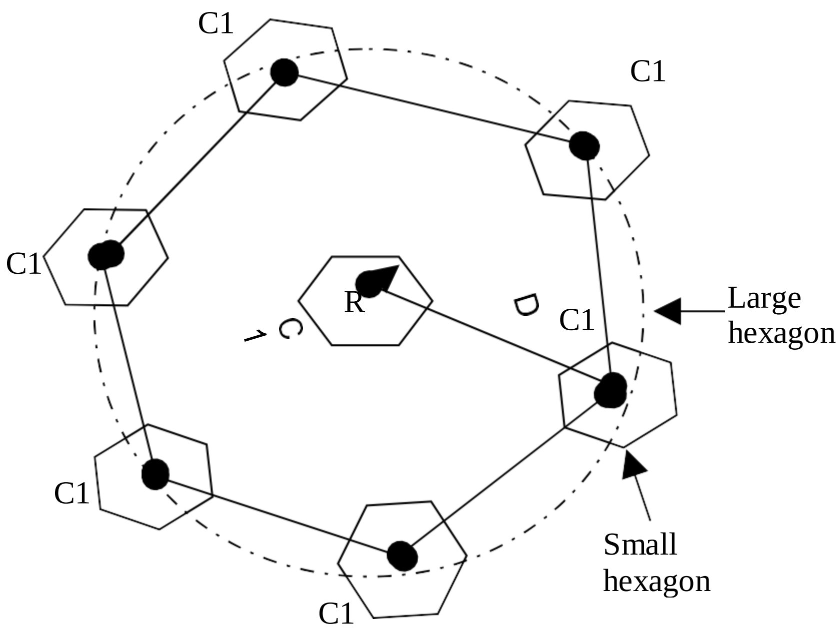


Fig. 1.7: Larger Hexagon in the First Tier

The area of the large hexagon having a radius D can be given as

$$A_{large\ hexagon} = \left(3\sqrt{3}/2\right) \times D^2 \tag{1.8}$$

Using equation 1.7

$$A_{large\ hexagon} = \left(3\sqrt{3}/2\right) \times 3 \times R^2 \times \left(i^2 + j^2 + i \times j\right) \qquad (1.9)$$

Step 7: To find the number of cells in the large hexagon (L)

Number of cells in large hexagon

$$L = A_{large\ hexagon} / A_{small\ hexagon} \qquad (1.10)$$

Using equations 2.9, 2.5 & 2.10, we get

$$L = 3 \times (i^2 + j^2 + i \times j) \qquad (1.11)$$

Step 8: Find the correlation between L and cluster size K

It can be seen from Fig. 2.8, that the larger hexagon is created by joining the centers of co-channel cells present in the first tier contains 7 cells of the central cluster plus 1/3rd of the number of 7 cells of all the neighboring six clusters. Therefore, it can be calculated that the larger hexagon consisting of the central cluster of *K* cells plus 1/3rd the number of the cells connected with six neighboring clusters present in the first tier.

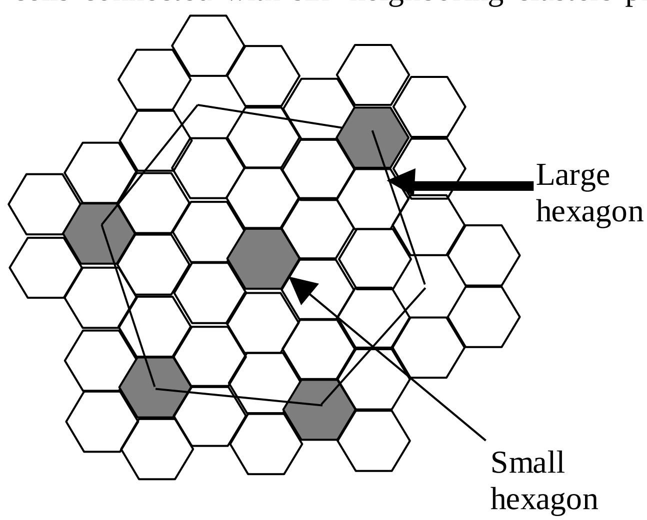


Fig. 2.8: Number of Clusters in the First Tier for N=7

Hence, the total number of cells enclosed by the larger hexagon is

$$L = K + 6 \times [(1 / 3) \times K)$$

$$L = 3 \times K \qquad (1.12)$$

Step 9: To establish relation between K and shift parameters

From equation 1.11 and 1.12, we get

$$3 \times K = 3 \times (i^2 + j^2 + i \times j)$$

$$K = (i^2 + j^2 + i \times j) \qquad (1.13)$$

The Table 1.1 shows the frequency reuse patterns along with the cluster sizes

Table 1.1: Frequency Reuse Pattern and Cluster Size

Frequency Reuse Pattern ***(I, j)***	**Cluster Size** $\boldsymbol{K = (i^2 + j^2 + i \times j)}$
(1, 1)	3
(2, 0)	4
(2, 1)	7
(3, 0)	9
(2, 2)	12
(3, 1)	13
(4, 0)	16
(2, 3)	19
(4, 1)	21
(5, 0)	25

Frequency Reuse Distance

To reuse the same set of radio channels in another cell, it must be separated by a distance called frequency reuse distance, which is generally represented by D.

Reusing the same frequency channel in different cells is restricted by co-channel interference between cells. So, it is necessary to find the minimum frequency reuse distance D in order to minimize the co-channel interference. Fig. 2.9 illustrates the separation of cells by frequency reuse distance in a cluster of 7 cells. In order to derive a formula to compute D, necessary properties of regular hexagon cell geometry are first discussed.

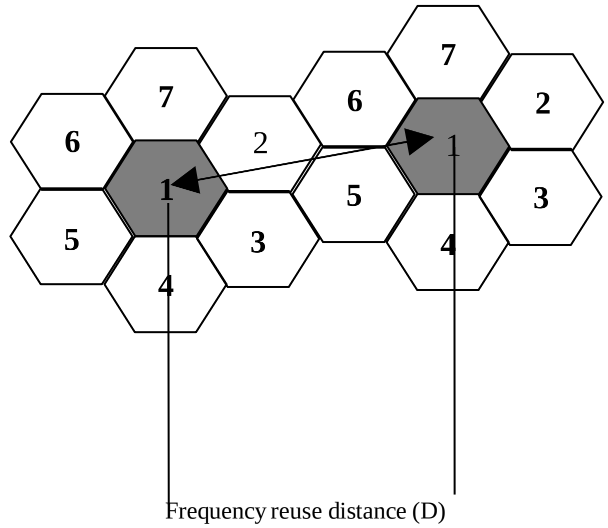


Fig. 1.9: Frequency Reuse Distance

The frequency reuse distance (D), which allows the same radio channel to be reused in co-channel cells, depends on many factors:

- the number of co-channel cells in the neighborhood of the central cell
- the type of geographical terrain
- the antenna height
- the transmitted signal strength by each cell-site

Suppose the size of all the cells in a cellular is approximately same, and it is usually calculated by the coverage area of the proper signal strength in every cell. The co-channel interference does not depend on transmitted power of each, if the cell size is fixed, i.e., the threshold level of received signal at the mobile unit is tuned to the size of the cell.
The co-channel interference depends upon the frequency reuse ratio, q, and is defined as

$$q = D / R$$

Where D is the distance between the two neighboring co-channel cells, and R is the radius of the cells. The parameter q is also referred to as the frequency reuse ratio or co-channel reuse ratio. The following steps are used to find the relationship between frequency reuse ratio q and cluster size K

Fig. 2.10 shows an array of regular hexagonal cells, where R is the cell radius. Due to the hexagonal geometry each hexagon has exactly six equidistant neighbors.

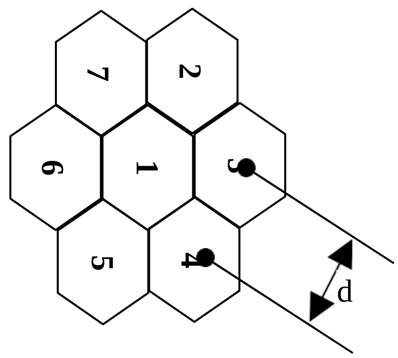


Fig. 1.10: Distance Between Two Adjacent Cells (d)

Let d be the distance between two cell centers of neighboring cells. Therefore,

$$d = \sqrt{3}R$$

The relationship between D, d, and shift parameters is

$$D^2 = 3 \times R^2 \times (i^2 + j^2 + i \times j)$$

As $K = i^2 + j^2 + i \times j$

$$D^2 = 3 \times R^2 \times K$$

$$\frac{D^2}{R^2} = 3 \times K$$

$$\frac{D}{R} = \sqrt{3K}$$

As q = D/R

$$q = \sqrt{3K}$$

Thus, the frequency reuse ratio q can be computed from the cluster size K. Table 2.2 shows the frequency reuse ratios for different cluster sizes, K

Table 1.2: Frequency Reuse Ratio and Cluster Size

Cluster Size ***K***	**Frequency Reuse Ratio** **q =** $\sqrt{3K}$
3	3.00
4	3.46
7	4.58
9	5.20
12	6.00
13	6.24
19	7.55
21	7.94
27	9.00

As the D/R measurement is a ratio, if the cell radius is decreased, then the distance between co-channel cells must also be decreased by the same amount, for keeping co-channel interference reduction factor same. On the other hand, if a cell has a large radius, then the distance between frequency reusing cells must be increased proportionally in order to have the same D/R ratio.

As frequency reuse ratio (q) increases with the increase in cluster size (K), the smaller value of K largely increase the capacity of the cellular system. But it will also increase the co-channel interference. Therefore, the particular value of q (or K) is selected in order to keep the signal-to-cochannel interference ratio at an acceptable level. If all the antennas transmit the same power, then with the increase in K, the frequency reuse distance (D) increases, and reduce the likelihood that co-channel interference may occur. Therefore, the challenge is to get the optimal value of K so that the desired system performance can be achieved in terms of increased system capacity, efficient radio spectrum utilization and signal quality.

Channel Allocation Schemes

For effective utilization of the radio spectrum, a channel reuse scheme is required which must be able to increase the capacity and reduce interference. Several channel allocation schemes have been proposed to address these objectives. Channel allocation schemes are classified into *fixed, dynamic, and hybrid.* The selection of a particular channel allocation scheme influences the performance of the system, mainly how to manage the calls when a call is handed-over from one cell to another [190], [117], [186], [163].

In a fixed channel allocation scheme, a set of nominal channels are permanently allocated to each cell. Any call generated from within the cell can only be served by the idle radio channels present in that cell. If all the radio channels present in that cell are occupied, then the call is *blocked*. However, there exist a several variations of the fixed allocation. In one of the variation, a cell can borrow channels from neighboring cells if its own channels are already busy, and this scheme is called channel borrowing strategy. Such a borrowing procedure is being managed by mobile switching center (MSC) and it try to make sure that the borrowing of a radio channel form neighboring cells does not interfere with any of the existing calls present in the donor cell.

In a dynamic channel allocation scheme, cells are not allocated radio channels permanently. Instead, every time when a call is received, the serving base station (BS) enquires a channel from the MSC. The MSC allocates a channel to the cell after taking into consideration the possibility of future blocking rate of the candidate cell, the re-use distance of the channel, and several other parameters.

Therefore, the MSC then allocates a particular channel if that radio channel is currently not in use in the candidate cell as well in any other neighboring cell which falls inside the minimum channel reuse distance in order to avoid co-channel interference. The Dynamic channel allocation minimizes the possibility of blocking, thereby increasing the trunking capacity of the system, as all the available channels are accessible to all the cells. In Dynamic channel allocation schemes MSC gather information on traffic distribution, channel occupancy of all channels on a regular basis. This results in increased channel utilization with decreased probability of dropped and blocked calls, but at the same time the computational load on the system also increases.

Handoff Strategies

When a mobile moves from one cell to another cell when a call is in progress, the MSC automatically shifts the call to a new channel present in the new cell. This handoff operation requires the identification of a new base station, and channels that are associated with the new base station.

In any cellular network, managing handoff is very important job. Many handoff schemes give high priority to handover requests over new call requests while allocating free channels, and it must be performed successfully and as infrequently as possible. Therefore, in order to satisfy these requirements, optimum signal at which to begin a handoff level must be specified by system designers. When an optimal signal level for acceptable voice quality is specified, a somewhat stronger signal level is used as a threshold at which a handoff is made. This margin is given by A = $P_{r\ handoff}$-$P_{r\ minimam\_usable}$, and it should not be too large or too small. If A is very large, needless handoffs which can burden the MSC may take place, and if A is very small, there may not be a sufficient time to complete a handoff process, before a call is vanished due to weak signal. Therefore, A should be carefully selected to meet these contradictory requirements. Fig. 1.11 shows a handoff situation. Fig. 1.11(a) presents a case in which a handoff does not take place and the signal strength falls below the minimum acceptable level in order to keep the channel active. This call dropping occurs when there is tremendous delay by the MSC in allocating a handoff, or when the threshold A is too small. During high traffic loads unnecessary delays may take place and this happens either due to computational overloading at the MSC or no free channels are available in any of the neighboring cells and thereby MSC has to wait until a free channel is found in a neighboring cell.
While deciding about handoff initiation time, it is important to make sure that the drop in the signal level is not due to temporary fading but the mobile is in fact moving away from its base station. Therefore, base station observes the signal strength for a definite period

of time before a handoff begins. This signal strength measurement must be optimized in order to avoid unwanted handoffs, while ensuring that unwanted handoffs are completed before a call gets dropped. The time required to come to a decision if a handoff is needed, depends on the speed of the vehicle at which it is moving. Information about the speed of vehicle can also be calculated from the fading signal received at the base station.

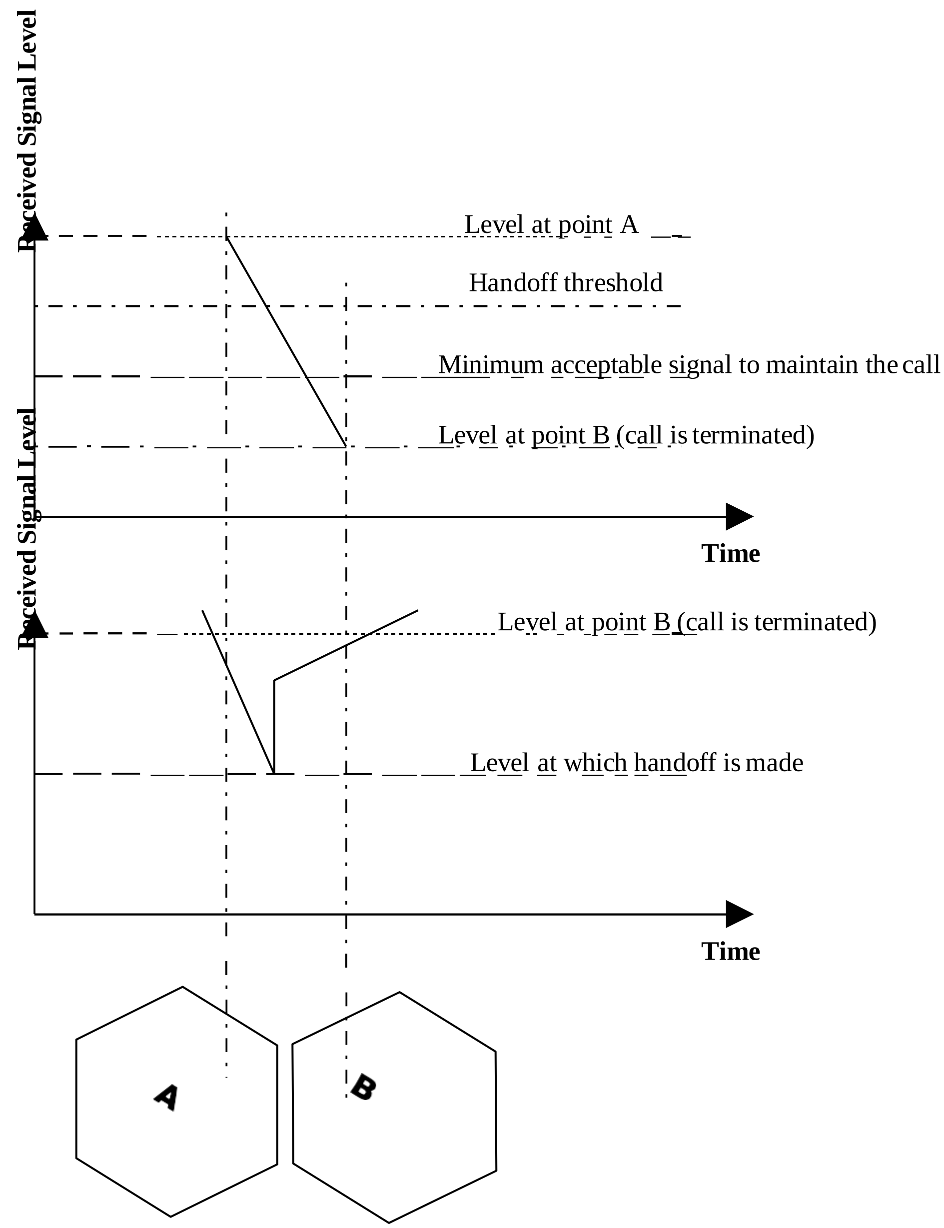


Fig. 1.11: Handoff Situation

The time during which a caller remains within a cell, without any handoff to the neighboring cells, is called the *dwell time* . The dwell time of a call depends upon a number of factors i.e. propagation, interference, distance between the caller and the base station, and several other time varying factors. It has been analyzed that variation of dwell time depends on the speed of the caller and the radio coverage type. e.g., a cell in which radio coverage is provided to highway callers (using vehicles), a large number of callers have a moderately steady speed and they follow fixed paths with good radio coverage. For such instances, the dwell time for random caller is a random variable having distribution that is highly concentrated on the mean dwell time. Whereas, for callers present in dense, micro-cellular environments, there is normally a huge deviation of dwell time about the mean, and the dwell times in general are shorter than the cell geometry. It is clear that the information of dwell time is very important while designing handoff algorithms.

In first generation cellular systems, signal strength computations are done by the base stations and monitored by the MSC. All the base stations regularly observe the signal strengths of its reverse channels to find out the relative location of each mobile user with respect to the base station. In addition to calculating the radio signal strength indication (RSSI) of ongoing calls in the cell, an extra receiver in each base station, is used to find out signal strengths of mobile users present in the neighboring cells. The extra receiver is controlled by the MSC and is used to examine the signal strength of callers in the neighboring cells, and informs RSSI to the MSC. Based on the RSSI values received from each extra receiver, the MSC determines whether handoff is required or not.

In second generation cellular systems using digital TDMA technology, handoff decisions are *mobile assisted*. In *mobile assisted handoff (MAHO),* each mobile station measures the received power from the neighboring base stations and informs these results to the serving base station. A handoff starts when the power received from the base station of a neighboring cell go above the power received from the present base station. In MAHO scheme, the call to be handed off between different base stations at a lot faster speed than in first generation systems because the handoff computations are done by each mobile and by keeping the MSC out of these computations. MAHO is suitable for micro-cellular network architectures where handoffs are more frequent.

When a call is in progress, if a mobile shifts from one cellular system to an another cellular system managed by a different MSC, an *intersystem handoff* is required. An MSC performs an intersystem handoff when a signal goes weak in a particular cell and the MSC fails to find another cell inside its system to which it can move the ongoing call, and several issues should be addressed while intersystem handoff is implemented. e.g. a local call might automatically turn into a long-distance call when the caller shifts out of its home network and enters into a neighboring system.

Various systems have different methods for dealing with hand-off requests. Several systems manage handoff requests in the same way as they manage new call requests. In such systems, the possibility that a handoff call will not be served by a new base station is equivalent to the blocking probability of new calls. However, if a call is terminated unexpectedly while in progress is more frustrating than being blocked occasionally on a new call. Therefore, to improve the quality of service, various methods have been created to give priority to handoff call requests over new call requests while allocating channels.

Prioritizing Handoffs

One scheme for prioritizing handoffs call requests is called the *guard channel concept,* in which a part of the existing channels in a cell is reserved entirely for handoff call requests. The major drawback of this scheme is that it reduces the total carried traffic, as smaller number of channels is allocated to new calls. However, guard channels scheme present efficient spectrum utilization when dynamic channel allocation strategies are used.

Queuing of handoff calls is another way to minimize the forced call terminations due to unavailability of channels in the cell. There is actually a tradeoff between the minimization in the possibility of forced call termination of handoff calls and total carried traffic. Handoff call queuing is possible as there is a fixed time interval between the time the received signal strength falls below the handoff threshold and the time the call is terminated due to unavailability of signal strength. The queue size and delay time is calculated from the traffic pattern of the service area. It should be noted that queuing of handoff calls does not promise a zero forced call terminations, because large delays will force the received signal strength to fall below the minimum level required to maintain communication and therefore, lead to forced handoff call termination.

Interference and System Capacity

Interference is one of the major factors affecting the performance of cellular radio systems. Sources of interference consist of another mobile inside the same cell, an ongoing call in a neighboring cell, other base stations transmitting signal in the same frequency band, or any non-cellular system which accidentally transmits energy into the cellular frequency band. Interference on voice signals could give rise to cross talk, where the caller hears interference in the background due to the presence of an unwanted transmission. The presence of interference in control channels, gives rise to missed and blocked calls. Interference is very dangerous in urban areas, due to the presence of larger base stations and mobile with greater RF noise. Interference has been accepted as a major obstruction in increasing the capacity of a system and is largely responsible for dropped calls in a network. The two major types of interferences that are taken consideration while allocating channels to the calls are *co-channel* and *adjacent channel interference.* While interfering signals are generated inside the cellular system by cellular transmitters, but they are difficult to control. The interference due to out-of-band users is very difficult to control, which happens without any word of warning, because of front end overload of subscriber equipment or intermittent inter-modulation products.

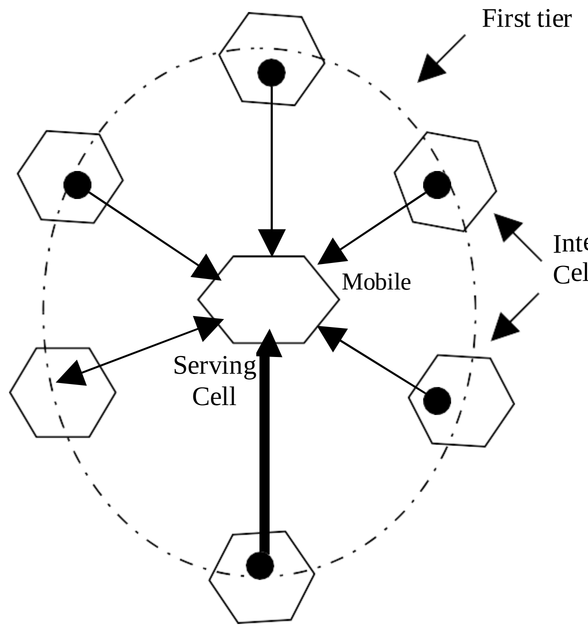
Co-channel Interference and System Capacity

The channel reuse approach is very useful for increasing the efficiency of radio spectrum utilization but it results in co-channel interference because the same radio channel is repeatedly used in different co-channel cells in a network. In this case, the quality of a received signal is very much affected both by the amount of radio coverage area and the co-channel interference.

Co-channel interference takes place when two or more transmitters located within a wireless system, or even a neighboring wireless system, which are transmitting on the same radio channel. Co-channel interference happens when the same carrier frequency (base station) reaches the same receiver (mobile phone) from two different transmitters.

This type of interference is generally generated because channel sets have been allocated to two different cells that are nor far enough geographically, and their signals are strong enough to cause interference to each other. Thus, co-channel interference can either modify the receiver or mask the particular signal. It may also merge with the particular signal to cause severe distortions in the output signal.

The co-channel interference can be evaluated by picking any particular channel and transmitting data on that channel at all co-channel sites. In a cellular system with hexagonal shaped cells, there are six co-channel interfering cells in the first tier. Fig. 2.12 shows a Test 1 which is set-up to calculate the co-channel interference at the mobile unit, in this test mobile unit is not stationary but is continuously moving in its serving cell.



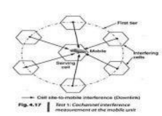


Fig. 1.12: Co-channel Interference Measurement at the Mobile Unit

In a small cell system, interference will be the major dominating factor and thermal noise can be neglected. Thus the S/I can also be written as:

$$\frac{S}{I} = \frac{1}{\sum_{k=1}^{6} \left(\frac{D_k}{R} \right)^{-\gamma}} \quad (1.14)$$

where

S/I = Signal to interference ratio at the desired mobile receiver,

S = desired signal power,

I = Interference power,

$2 \leq \gamma \leq 5$ is the propagation path-loss slope and γ depends on the terrain environment.

If we assume, for simplification, that D_k is the same for the six interfering cells, i.e., $D = D_k$, then the formula above becomes:

$$\frac{S}{I} = \frac{1}{6(q)^{-\gamma}} = \frac{q^{\gamma}}{6} \tag{1.15}$$

1

For analog systems using frequency modulation, normal cellular practice is to specify an S/I ratio to be 18 dB or higher based on subjective tests. An S/I of 18 dB is the measured value for the accepted voice quality from the present-day cellular mobile receivers.

Using an S/I ratio equal to 18dB ($10^{18/10} = 63.1$) and γ =4 in the Eq. (1.15), then

$$q = [6 \times 63.1]^{0.25} = 4.41 \tag{1.17}$$

Substituting q from Eq. (2.17) into Eq. (2.12) yields

$$N = \frac{(4.41)^2}{3} = 6.49 \approx 7\,. \tag{1.18}$$

Eq. (1.18) indicates that a 7-cell reuse pattern is needed for an S/I ratio of 18 dB.

Therefore, the performance of interference-limited cellular mobile system can be calculated from the following results.

a) If the signal-to-interference ratio (S/I) is greater than 18 dB, then the system is said to be correctly designed.

b) If S/I is less than 18 dB and signal-to-noise ratio (S/N) is greater than 18 dB, then the system is said to be experiencing with a co-channel interference problem.

c) If both S/I and S/N are less than 18 dB and S/I is approximately same as S/N in a cell, then the system has a radio coverage problem.

d) If both S/I and S/N are less than 18 dB and S/I is less than S/N, the system has both co-channel interference and radio coverage problem.

Therefore, the reciprocity theorem can be used to study the radio coverage problem, but it does not give accurate results when used for the study of co-channel interference problem. Therefore, it is suggested to perform Test 2 in order to measure co-channel interference at the cell-site. In Test 2 shown in Fig. 2.13, both the mobile unit present in the serving cell and six other mobile units present in the neighboring cells are transmitting simultaneously at the same channel.

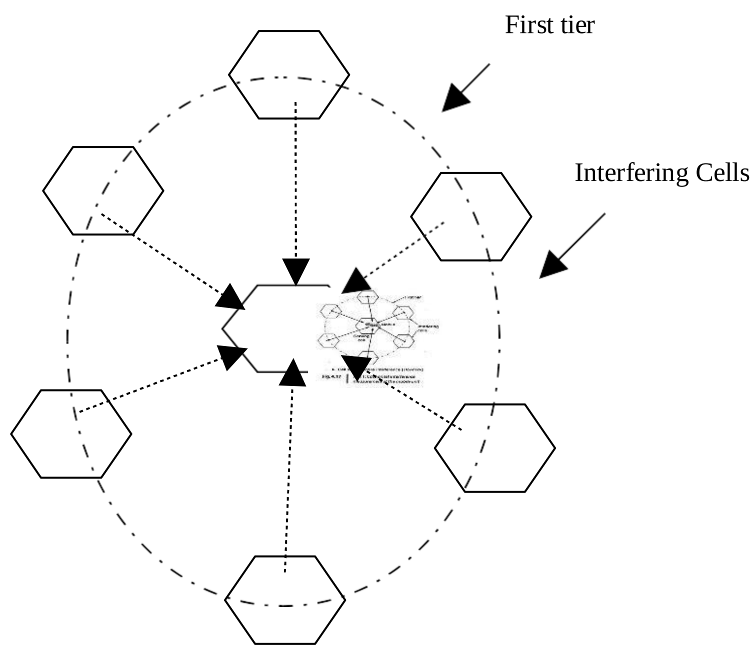


Fig. 1.13: Co-channel Interference Measurement at the Cell-site

The received signal strength measurements are done at the serving cell, under the following conditions:

- When only the mobile unit present in the serving cell transmits (signal measured as S)
- Up to six interference levels are measured at the serving cell-site due to presence of six mobile units in the neighboring cells (the average signal measured as I)
- Noise from sources other than mobile unit (signal measured as N)

Then the received S/I and S/N is computed at the serving cell site. The test results are compared with the Test 1, and from the it can be easily found whether the cellular system has a radio coverage or a co-channel interference problem or both.

Co-channel Interference Reduction Methods

Interference is major factor affecting the performance of cellular communication systems. Sources of interference may consist of a different mobile working in the same or in the neighboring cells, which are operating in the same frequency band that may leak energy into the cellular band.

Cells that use same set of radio channels are called co-channel cells, and the interference caused by the received signals coming from these cells is called co-channel interference. If the different cells in the cellular network use different radio channels then the inter-cell interference should be kept at a minimum level. When the number of mobile users increase and the radio channels available in the system are limited, then, in order to satisfy this high demand, the radio channels have to be reused in various cells. That is why for increasing the capacity, there exist many co-channel cells which can simultaneously serve the large number of users.

In fact, Deployment of radio channel reuse is required to improve the capacity of a system. But, the reuse mechanism brings in co-channel interference from neighboring cells using the same set of radio channels. Therefore, the quality of received signal gets affected by the amount of co-channel interference and the extent of radio coverage. Therefore, frequency reuse should be planned very carefully in order to keep the co-channel interference at an acceptable level.

The co-channel interference can be reduced by the following methods:

a. *Increasing the distance(D) between two co-channel cells, D*

   As D increases, the strength of interfering signal from co-channel interfering cells decreases significantly. But it is not wise to increase D because as D is increased, K must also be increased. High value of K means fewer number of radio channels are available per cell for a given spectrum. This results into decrease of the system capacity in terms of channels that are available per cell.

b. *Reducing the antenna heights*

   Reducing antenna height is a good method to minimize the co-channel interference in some environment, e.g., on a high hill. In the cellular system design effective antenna height is considered rather than the actual antenna height. Therefore, the effective antenna height changes according to the present location of the mobile unit in such a difficult terrain.

   When the antenna is put up on top of the hill, the effective antenna height gets more than the actual antenna height. So, in order to minimize the co-channel interference, antenna with lower height should be used without decreasing the received signal strength either at the cell-site or at the mobile device. Similarly, lower antenna height in a valley is very useful in minimizing the radiated power in a far-off high-elevation area where the mobile user is believed to be present.

   However, reducing the antenna height does not always minimize the co-channel interference, e.g., in forests, the larger antenna height clears the tops of the longest trees in the surrounding area, particularly when they are located very close to the antenna. But reducing the antenna height would not be appropriate for minimizing co-channel interference because unnecessary attenuation of the signal would occur in the vicinity of the antenna as well as in the cell boundary if the height of the antenna is below the treetop level.

c. *Using directional antennas.*

   The use of directional antennas in every cell can minimize the co-channel interference if the co-channel interference cannot be avoided by a fixed division of co-channel cells. This will also improve the system capacity even if the traffic increases. The co-channel interference can be further minimized by smartly setting up the directional antenna.

d. *Use of diversity schemes at the receiver.*

The diversity scheme used at the receiving end of the antenna is an efficient technique for minimizing the co-channel interference because any unwanted action performed at the receiving end to increase the signal interference would not cause further interference. For example, the division of two receiving antennas installed at the cell-site meeting the condition of h/s=11, (where h is the antenna height and s is the division between two antennas), would produce the correlation coefficient of 0.7 for a two-branch diversity system. The two correlated signals can be combined with the use of selective combiner. The mobile transmitter could suffer up to 7 dB minimization in power and the same performance at the cell-site can be achieved as a non-diversity receiver. Therefore, interference from the mobile transmitters to the receivers can be significantly reduced.

Adjacent Channel Interference

Signals from neighboring radio channels, also called adjacent channel, leak into the particular channel, thus causing adjacent channel interference. Adjacent channel interference takes place due to the inability of a mobile phone to separate out the signals of adjacent channels allocated to neighboring cell sites (e.g., channel 101 in cell A, and channel 102 in cell B), where both A and D cells are present in the same reuse cluster. The problem of adjacent channel interference can become more serious if a user transmitting on a channel, which is extremely close to a subscriber's receiver channel, while the receiver tries to receive a signal from base station on the desired channel. This is called the *near and far* effect, where a neighboring transmitter catches the receiver of the user. Otherwise, the near-far effect occurs when a mobile near to a base station transmits on a channel which is close to the one being used by a weak mobile. The base station may find some trouble in separating out a particular user from the one using adjacent channel

Adjacent channel interference can be reduced through careful and thorough filtering and efficient channel allocations. As each cell is allocated only a portion of the total channels, a cell must not be allocated channels which are located adjacent in frequency. By maintaining the channel separation as large as possible in a given cell, the adjacent channel interference may well be minimized significantly. Hence, instead of allocating contiguous band of channels to each cell, channels are allocated such a way that the frequency separation between channels in a given cell should be maximized. With sequentially allocating consecutive channels to various cells, several channel allocation schemes are capable enough to keep apart adjacent channels present in a cell with bandwidth of N channels, where N is the size of a cluster. However, some channel allocation schemes also avoid a secondary source of adjacent channel interference by not using the adjacent channels in neighboring cells.

Trunking and Grade of Service

In cellular mobile communication, the two major aspects that have to be considered with extra care are: trunking, and grade of service. These aspects are to be planned very well in order to get a better system performance. The grade of service is a standard which is used to define the performance of a cellular mobile communication system by specifying a desired probability of a mobile user acquiring channel access, when a definite number of radio channels are present in the system. The cellular communication network depends on a trunking system to fit large number of mobile users in a limited radio band. The statistical behavior of mobile users is being exploited by trunking so that a fixed number of channels can be allocated to large mobile users. In trunking, large number of mobile users is being accommodated to share the limited radio channels available in a cell.

In trunked cellular communication systems, each mobile user present in network is allocated a channel on the basis of a request. After the call is terminated, the occupied channels immediately go back to the pool of available channels. When a mobile user made a request for channels and if all of the radio channels are occupied, then the incoming call is blocked. In few communication systems, a queue is generally used to keep the requesting mobile users until a channel becomes free. The grade of service (GOS) is used to determine the capability of a user to get access to trunked radio systems during busy hours. The busy hour is generally based on customer’s request for channels during peak load.

It is, therefore, necessary to approximate the maximum required capacity in terms of number of available channels and to allocate the appropriate number of channels in order to meet the GOS. GOS is generally defined as the probability that a call is blocked. A call which cannot get completed after the call request is made by a user is called a blocked or lost call, and it may happen either due to channel congestion or due to the non-availability of a free channel. Therefore, GOS can be computed from channel congestion which is defined as the call blocking probability, or being delayed beyond a certain time.

The traffic intensity (A_u Erlangs) generated by each user is

$$A_u = \lambda H$$

where λ is the average number of calls generated per unit time and H is the average duration of each call. If A is having U users and number of channels are not mentioned, then the total offered traffic intensity A is

$$\mathrm{A} = UA_u$$

Additionally, if a trunked system is having C channels, and the traffic is equally divided between the channels, then the intensity of traffic (A_c) for each radio channel is

$$A_c = UA_u/C \qquad (1.19)$$

Note that when the offered traffic goes past the maximum capacity of the system, the total carried traffic gets very limited due to the limited number of channels. The maximum possible carried traffic is the total number of channels, C, in Erlangs. The AMPS system is generally developed for a GOS of 2% blocking and it shows that 2 out of 100 calls will be blocked because channels are occupied during the busiest hour.
Different types of trunked radio systems commonly used in the networks are:

1. In the first type, no queuing is offered for call requests i.e., for each user who requests service, there exists no setup time and if free radio channel is available, it is immediately allocated to the user. If all the channels are busy, then the requesting user is blocked. In this trunking system, it is assumed that call arrival follows a Poisson distribution and the trunking is also called *blocked calls cleared*. Moreover, it

is also assumed that there are unlimited users in the network and having the following additional features:

(a) The channel request can be made at any time by all the mobile users (both new and blocked users); (b) the probability of a user being allocated a channel is exponentially distributed, therefore, occurrence of longer call duration is very unlikely as explained by an exponential distribution; and (c) there are a fixed number of channels present in the trunking pool, and it is known as an M/M/m queue, which leads us to the derivation of the Erlang B formula. The Erlang B formula helps in finding the probability that a call is blocked and also measures the GOS for a trunked radio system which does not provide queuing for blocked calls. The Erlang B formula is

$$Pr\ [blocking] = \frac{\frac{A^C}{C!}}{\sum_{k=0}^{C} \frac{A^k}{k!}} = \text{GOS} \qquad (1.20)$$

where C is the number of trunked channels present in the trunked radio system and A is the offered traffic. It is possible to design a trunked systems with fixed number of users, but the final expressions are found to be very complex than the Erlang B, and the added complexity is not acceptable for typical trunked radio systems in which number of users are more than the channels present in the system.. The capacity of a trunked radio system in which blocked calls are lost is shown in Table 2.3.

Table 2.3: 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
4	0.869	0.701	0.535	0.439
5	1.36	1.13	0.9OO	0.762
10	4.46	3.96	3.43	3.09
20	12.0	11.1	10.1	9.41
24	15.3	14.2	13.0	12.2
40	29.0	27.3	25.7	24.5

70	56.1	53.7	51.0	49.2
100	84.1	80.9	77.4	75.2

2. In a second form of trunked networks, a queue is used to keep the blocked calls. If all the channels are presently busy, then the call can be postponed until a free channel is found, and this whole process is *Blocked Calls Delayed,* and the GOS for this type of trunking is the probability that a new call is not allocated to a channel even after waiting a certain time in the queue. The probability that a new call is not allocated a channel immediately is calculated by the Erlang C formula

$$Pr\,[delay > 0] = \frac{A^C}{A^C + C!\left(1 - \frac{A}{C}\right)\sum_{k=0}^{C-1}\frac{A^k}{k!}} \qquad (1.21)$$

If no channel is currently found free then the call is delayed, and the GOS of a trunked system in which the blocked calls are delayed is given by

$$\begin{aligned} Pr[delay>t] &= Pr\,[delay>0]\,Pr\,[delay > t \mid delay > 0] \qquad (1.22) \\ &= Pr\,[delay>0]\exp(-(C-A)t/H) \end{aligned}$$

For all the calls n a queued system the average delay D is given by

$$D = Pr\,[delay > 0]\,\frac{H}{C-A} \qquad (1.23)$$

Improving Capacity In Cellular Systems

With the rise in the demand for wireless services, the number of radio channels allocated to each cell could become inadequate in order to satisfy this increase in the demand. Therefore, to increase the capacity (i.e. a cellular system can take up more calls) of a cellular system, it is very important to allocate more number of radio channels to each cell in order to meet the requirements of mobile traffic. Various techniques that are proposed for increasing the capacity of a cellular system is as follows:

i. Cell splitting
ii. Cell sectoring
iii. Repeaters for extending range

iv. Micro zone method

Cell Splitting

Cell splitting is a method in which congested (heavy traffic) cell is subdivided into smaller cells, and each smaller cell is having its own base station with reduction in antenna height and transmitter power. The original congested bigger cell is called macrocell and the smaller cells are called microcells. Capacity of cellular network can be increased by creating micro-cells within the original cells which are having smaller radius than macro-cells, therefore, the capacity of a system increases because more channels per unit area are now available in a network.

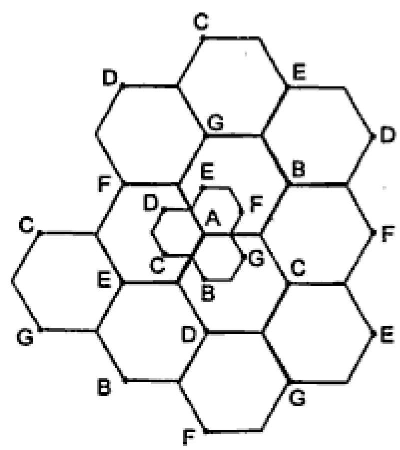


Fig. 1.14: Cell Splitting

Fig. 2.14 shows a cell splitting in which a congested cell, divided into smaller micro-cells, and the base stations are put up at corners of the cells. The micro-cells are to be added in such a way in order to the frequency reuse plan of the system should be preserved. For micro-cells, the transmit power of transmitter should be reduced, and each micro-cell is having half the radius to that of macro-cell. Therefore, transmit power of the new cells can be calculated by analyzing the received power at the cell boundaries. This is required in order to make sure that frequency reuse plan for the micro-cells is also working the same way as it was working for the macro-cells.

$$P_{r-O} \propto P_{tp} R^{-n}$$

$$P_{r\text{-}N} \propto P_{tN}\left(\frac{R}{2}\right)^{-n}$$

Where P_{tp} is the transmit power of macro-cell

P_{tN} is the transmit power of macro-cell

n is the path loss exponent

$R, \left(\frac{R}{2}\right)$ is the radius of macro and micro-cells

In cell splitting, following factors should be carefully monitored;

1. In cell splitting, allocation of channels to the new cells (micro-cells) must be done very cautiously. So, in order to avoid co-channel interference, cells must follow the minimum reuse distance principle.

2. Power levels of the transmitters for new and old cells must be redesigned. If the transmitter of the old cell has the same power as that of new cells, then the channels in old cell interfere with the channels of new cell. But, if the power level of transmitter is too low then it may result into in sufficient area coverage.

3. In order to overcome the problem of point (2); the channels of macro-cell is divided into two parts. The channels in the first part are for the new cell and other part consists of channel for the old cell. Splitting of cells is done according to the number of subscribers present in the areas, and the power levels of the transmitters must be redesigned according to the allocated channels to old and new cells.

4. Antennas of different heights and power levels are used for smooth and easy handoff, and this technique is called Umbrella cell approach. Using this approach large coverage area is provided for high speed users and small coverage area to low speed users. Therefore, the number of call handoffs is maximized for high speed users and provides more channels for slow speed users.

5. The main idea behind cell splitting is the rescaling of entire system. In cell splitting, reuse factor (D/R) is kept constant because by decreasing the radius of cell (R) and, at the same time, the separation between co-channel (D) is also decreased. So, high capacity can be achieved without changing the (D/R) ratio of system.

Sectoring

Another way of improving the channel capacity of a cellular system is to decrease the D/R ratio while keeping the same cell radius. Improvement in the capacity can be accomplished by reducing the number of cells in a cluster, hence increasing the frequency reuse. To achieve this, the relative interference must be minimized without decreasing the transmit power.

For minimizing co-channel interference in a cellular network, a single omni-directional antenna is replaced with multiple directional antennas, with each transmitting within a smaller region. These smaller regions are called sectors and minimizing co-channel interference while improving the capacity of a system by using multiple directional antennas is called sectoring. The amount up to which co-channel interference is minimized depends on the amount of sectoring used. A cell is generally divided either into three 120 degree or six 60 degree sectors. In the three-sector arrangement, three antennas are generally located in each sector with one transmit and two receive antennas. The placement of two receive antennas provide antenna diversity, which is also known as space diversity. Space diversity greatly improves the reception of a signal by efficiently providing a big target for signals transmitted from mobile units. The division between the two receive antenna depends on the height of the antennas above ground.

When sectoring technique is used in cellular systems, the channels used in a particular sector are actually broken down into sectored groups, which are only used inside a particular sector. With 7-cell reuse pattern and 120 degree sectors, the number of interfering cells in the neighboring tier is brought down from six to two. Cell sectoring also improves the signal-to-interference ratio, thereby increasing the capacity of a cellular system. This method of cell sectoring is very efficient, because it utilized the existing system structures. Cell sectoring also minimized the co-channel interference, with the use of directional antennas, a particular cell will get interference and transmit only a fraction of the available co-channel cells.

It is seen that the reuse ratio $q = (N_I \times S/I)^{1/n}$, where N_I depends on the type of antenna used. For an omni-directional antenna with only first-tier of co-channel interferer, the number of co-channel interfering cells $N_I = 6$, but for a 120 degree directional antenna, it is 2

So, the increase in S/I ratio is

$$\frac{(N_I \times S/I)_{120^0}}{(N_I \times S/I)_{omni}} = \frac{q^n_{120^0}}{q^n_{omni}}$$

$$\frac{(S/I)_{120^0}}{(S/I)_{omni}} = 3$$

n = path loss exponent N_I = Number of co-channel interfering cells

q = frequency reuse ratio = D/R

Thus, S/I ratio increases with the increase in number of sectors, but at the cost of additional handoff that might be required for the movement of a user from one sector to another.

Microcell Zone Concept

The micro-cell zone concept is associated with sharing the same radio equipment by different micro-cells. It results in decreasing of cluster size and, therefore, increase in system capacity. The micro-cell zone concept is used in practice to improve the capacity of cellular systems.

To improve both capacity and signal quality of a cellular system, cell sectoring depends upon correct setting up of directional antennas at the cell-site. But it also gives rise to increase in the number of handoffs and trunking inefficiencies. In a 3-sector or 6-sector cellular system, each sector acts like a new cell with a different shape and cell. Channels allocated to the un-sectored cell are divided between the different sectors present in a cell, thereby decreasing number of channels available in each sector. Furthermore, handoff takes place every time a mobile user moves from one sector to another sector of the same cell. This results in significant increase of network load on BSC and MSC of the cellular system. The problem of channel partitioning and increase in network load become very hard if all the 3 or 6-sectored directional antennas are placed at the centre of the cell.

As shown in the Fig. 1.15, three directional antennas are put at a point, Z1, also called zone-site, where three adjacent cells C1, C2, and C3 meet with each other. Z1, Z2 and Z3 are three zone-sites of the cell C1, and each zone-site is using three 135 degree directional antennas. All the three zone-sites also behave as receivers, which also receive signals transmitted by a mobile user present anywhere in the cell. All the three zone-sites are linked to one common base station, as shown in Fig. 2.16. This arrangement is known as Lee's micro-cell zone concept.

In order to avoid delay, these zone-sites are connected through a high-speed fiber link to the base station. The base station first finds out, which of the three zone-sites has the better received signal strength from the mobile user and then that particular zone-site is used to transmit the signal to the mobile user. Therefore, only one zone-site is active at a time for communicating with the user and it also minimizes the co-channel interference experienced by the mobile user.

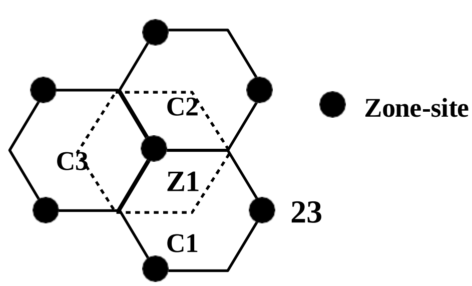


Fig. 1.15: Location of Zone-sites in Sectored Cells

Therefore, micro-cell zone architecture minimizes the co-channel interference, improves system capacity, demands less handoffs, and the system is easy to implement. The system capacity for a system with cluster size k=3 is 2.33 times greater than the present analog cellular system with k=7 for the C/I requirement of 18 dB. This micro-cell system gives improved voice quality than the AMPS cellular system at 850 MHz. The micro-cell zone concept can be used with both digital communication systems and personal communication systems, and is best suited for indoor applications. It is also very useful to provide services along highways or in crowded urban areas.

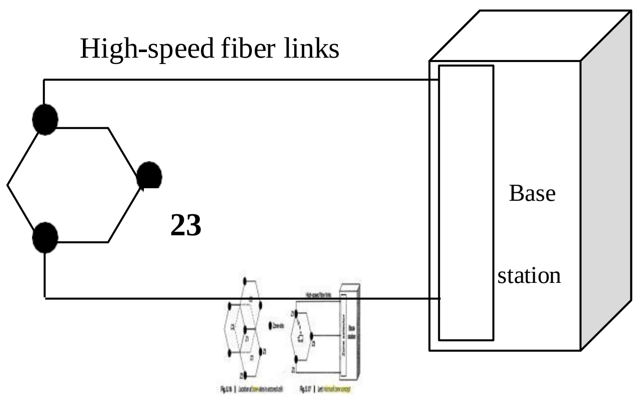


Fig. 1.16: Lee's Microcell Zone Concept

Advantages of micro zone concept:

1. When the mobile user moves from one zone to another within the same cell, the mobile user can keep the same channel for the call progress.
2. The effect of interference is very low due to the installation of low power transmitters.
3. Better signal quality is possible.
4. Fewer handoffs when a call is in progress.

Repeaters for Range Extensions

Wireless operators want to provide dedicated coverage for users located within buildings, or in valleys or tunnels as these areas are sometimes very hard to reach. Radio re-transmitters, also known as repeaters, are frequently used to provide coverage in such areas where range extension capabilities are required. Repeaters are bidirectional devices, as the signals can be concurrently transmitted to and received from a base station.

Repeaters may be installed anywhere as they function using over the air signals, and are able repeat entire frequency band. After receiving signals from base station, the repeater amplifies the signals before it forwards them to the coverage area. As repeaters can also reradiate the received noise, so repeaters must be installed very carefully. Directional antennas or distributed antenna system (DAS) are linked practically to the repeater inputs or outputs for spot coverage, mainly in tunnels or buildings.

A service provider dedicates some amount of cell site's traffic for the areas covered by the repeater by modifying the coverage of cell. As the repeater do not add more channels to the system, it only reradiates the base station signal into specific locations. Repeaters are generally used to provide coverage into those areas, where signal reception has been very weak. Signal penetration inside the building is generally provided by installing micro-cells outside the big building, and installing many repeaters inside the buildings. This technique provides better coverage into targeted areas, but does not increase the capacity that is required with the rise in the indoor and outdoor traffic. Therefore, dedicated base stations inside buildings are required to meet the service demands of large number of cellular users present inside the building. Finding a proper location for repeaters and distributed antennas inside the building needs a very careful planning, mainly due to the interference signals reradiated into the building. Also, repeaters must be able to match the available capacity from the base station. Software SitePlanner helps the engineers to decide the best location for putting up the repeaters and DAS network.

Conclusion

The fundamental concepts of frequency reuse, frequency planning, handoff, and trunking efficiency are presented in this chapter. The performance determining parameters such as grade of service, spectrum efficiency, and radio capacity under diverse situations are also discussed. Handoffs are essential to pass mobile traffic from one cell to another, and there is variety of different ways to implement handoffs. The capacity of a cellular system depends upon several variables. The *S/I* influence the frequency reuse factor of a cellular system, which restricts the number of radio channels within the coverage area. The number of users in a particular area is greatly influenced by trunking efficiency.

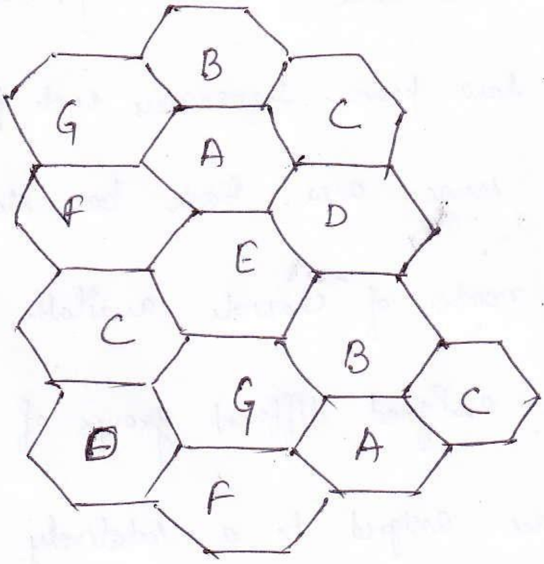
Cellular Concept :→

A cellular system provides a wireless connection to PSTN for any user location within the radio range of the system. cellular system accommodate a large number of user over a large geographic area, within a limited frequency spectrum. 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 (to small) by another base station located some distance away. 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 change.

The cellular concept is a system-level idea which call for replacing a single, high power transmitter with many low power transmitters 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 basestation are assigned different groups of channels so that all the available channels are assigned to a relatively small number of neighboring basestation. Neighbouring base station are assigned different groups of channels so that the interference between basestation is minimized.

Frequency Reuse Basic theory of Hexagonal cell layout :→

Cellular radio system rely on an intelligent allocation and reuse of channels throughout a coverage region. Each cellular base station is allocated a group of radio channel to be used within a small geographic area called a cell. Base station in adjacent cell are assigned channel groups which contain completely different channel than neighbouring cell. The design process of selecting and allocating channel group for all of the cellular base stations within a system is called frequency reuse or frequency planning.



The following fig shows the concept of cellular frequency reuse, where cells labeled with the same letter use the same group of channel. The hexagonal cell shape is a simplistic model of the radio coverage for each base station, but it has been universally adopted since hexagon permits easy and manageable analysis of a cellular system. The actual radio coverage of a cell is known as the footprint. The hexagonal geometry is chosen because of the following reasons (i) A cell must be designed to serve the weakest mobile within the footprint and these are typically located at the edge of the cell. (ii) For a given distance between the center of a polygon and its farthest perimeter points, the hexagon has the largest area of the three (square, triangle, polygon).

When using hexagons to model coverage area, base station transmitter are depicted as either in the center of the cell, or on three of the six cell vertices

To understand the frequency reuse concept, consider a cellular system which has a total of S duplex channels available for use. If each cell is allocated a group of K channels and if 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

$$S = KN$$

The N cells which collectively use the complete set of available frequencies is called cluster.

$$C = MKN = MS$$

The factor N is called cluster size and is typically equal to 4, 7 or 12.

The geometry of hexagon is such that the number of cells per cluster N, can only have values which satisfy.

$$N = i^2 + ij + j^2$$

Channel Allocation Schemes :→

For efficient utilization of the radio spectrum, a frequency reuse scheme that is consistent with the objective of increasing capacity and minimizing interference is required.

Channel assignment strategies can be classified as either fixed or dynamic.

Fixed Channel Assignment :→

Each cell is allocated a predetermined set of voice channels. Any call attempt within the cell can only be served by the unused channel in that particular cell. If all the channels in that cell are occupied, the call is blocked and the subscriber doesnot receive service. Several strategies exist one of them is called Borrowing strategy, in which a cell is allowed to borrow channel from neighbouring cell if all of its own channels are already occupied.

Dynamic channel Assignment :→

In this voice channels are not allocated to different cell 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.

Handover Analysis (Handoff Strategies) :→

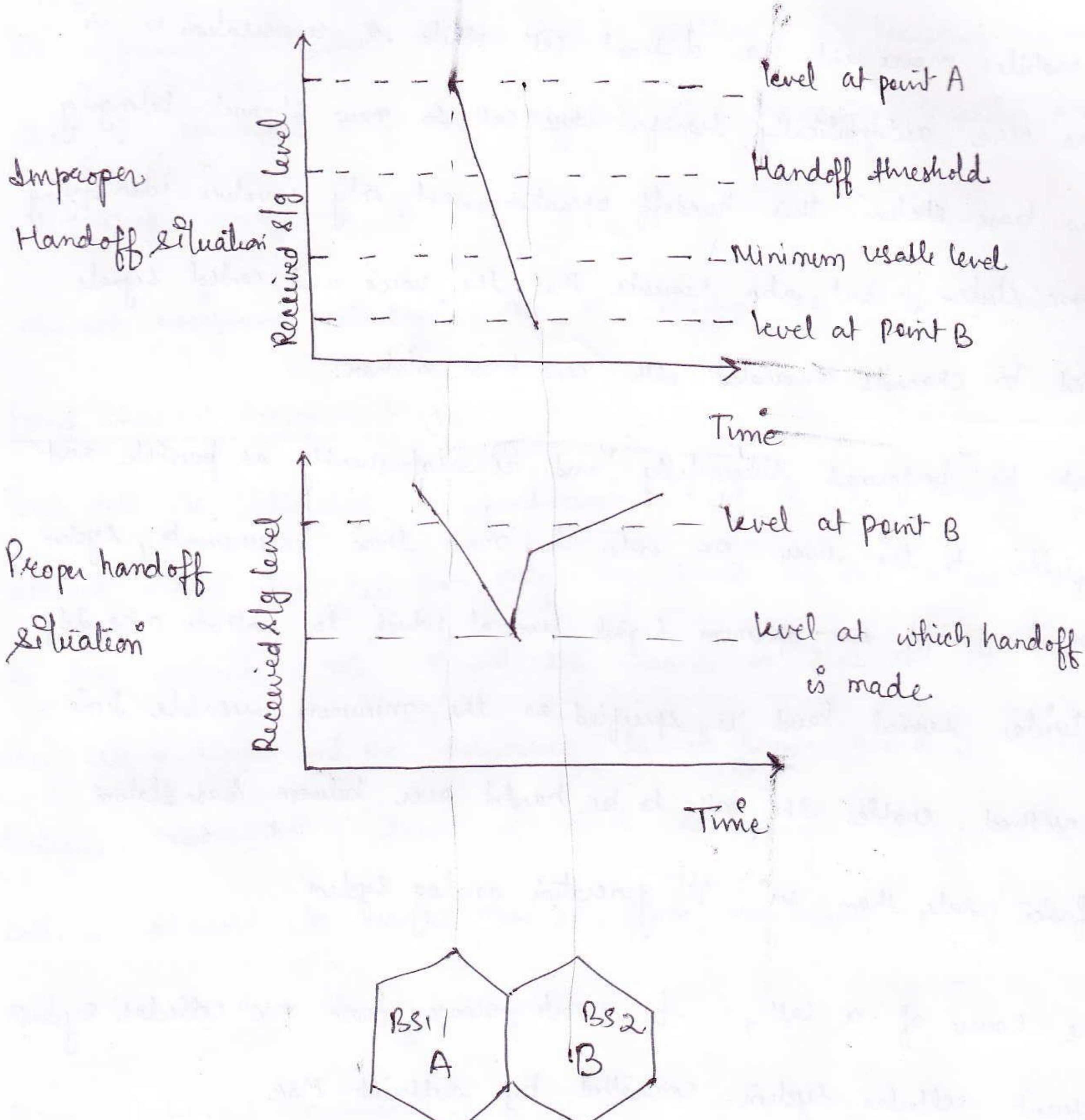
when a mobile moves into a different cell while a conversation is in progress , the MSC automatically transfer the call to new channel belonging to the new base station. This handoff operation not only involves identifying a new base station, but also requires that the voice and control signals be allocated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

Handoff must be performed successfully and as infrequently as possible and be imperceptible to the user. In order to meet these requirements, system designer must specify an optimum signal level at which to initiate a handoff. once a particular signal level is specified as the minimum usable time. the MAHO method enables the call to be handed over between base station at much faster rate than in 1st generation analog system

During the course of a call, if mobile moves from one cellular system to a different cellular system controlled by different MSC.

once a particular signal level is specified as the minimum usable signal for acceptable voice quality at the BS, a slightly stronger signal level is used as threshold at which handoff is made.

This margin given by $\Delta = P_{r\ handoff} - P_{r\ minimum\ usable}$ cannot be too large or too small.



If Δ is too large, unnecessary handoff which burden the MSC may occur, if Δ is too small, there may be insufficient time to complete a handoff before a call is lost due to weak s/g condition. The above fig shows the case where a handoff is not made and the signal drop 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 handoff or when the threshold Δ is set too small for the handoff time in the system. Excessive delay 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 BS. The time over which a call may be maintained within a cell, without handoff is called the dwell time.

There are different type of Handoffs.

Intersystem Handoff → If mobile moves from one cellular system to a different cellular system controlled by a different MSC, an intersystem handoff become necessary.

Hard Handoff → This Handoff is characterized by a mobile having a radio link with only one Associate Point (AP) at any time. Thus the old connection is terminated before a new connection is activated. This mode of operation is reffered to as break-before-make.

Soft Handoff →

In soft handoff, the mobile can simultaneously communicate with more than one AP during the handoff. This is, a new connection is made before breaking the old connection and is reffered to as make before break. CDMA system uses soft handoff technique.

Prioritizing Handoff :→

One method of giving priority to handoff is called the Guard channel concept, whereby a fraction of the total available channels in a cell is reserved exclusively for handoff request from ongoing call which may be handed off into the cell.
This method has the disadvantage of reducing the total carried traffic. 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.

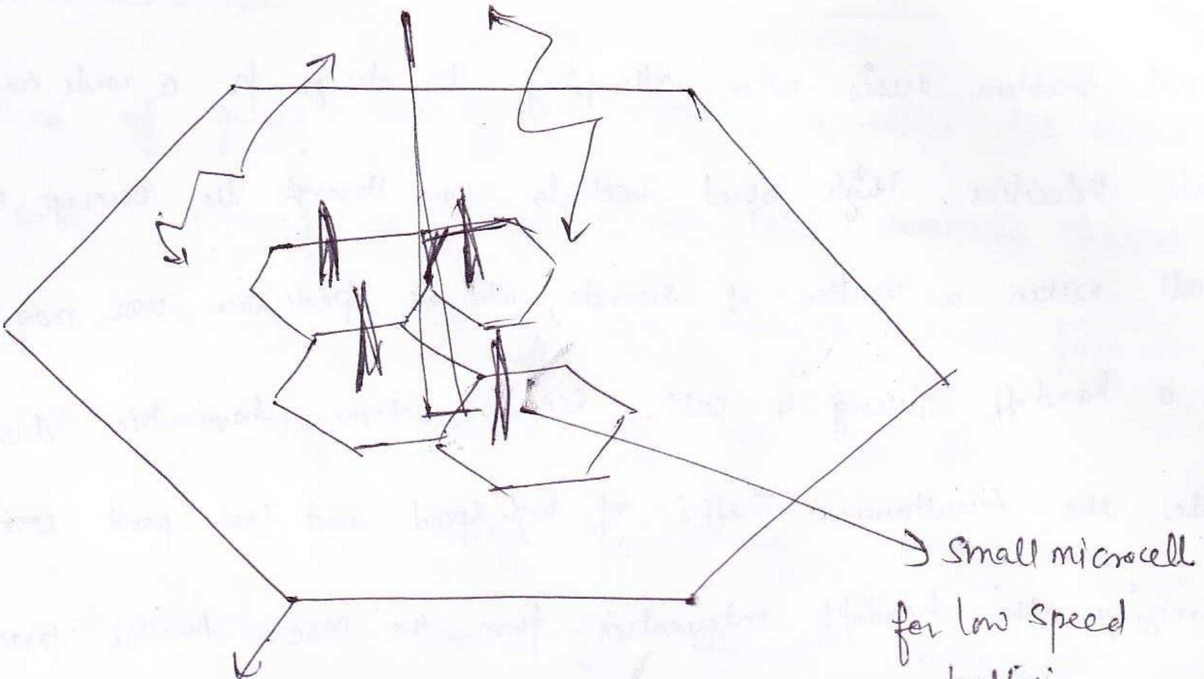
Queuing of handoff requests is another method to decrease the probability of forced termination of a call due to lack of available channels (there is tradeoff between the decrease in probability)
Queuing of handoff is possible due to the fact that there is a finite time interval between the time the received signal level drop below the handoff threshold and the time the call is terminated due to insufficient signal level.

Practical Handoff considerations :→

Several problem arises when attempting to design for a wide range of mobile velocities. High speed vechicles pass through the coverage region of a cell within a matter of seconds, whereas pedestrian user may never need a handoff during a call. Several schemes have been devised to handle the simultaneous traffic of high speed and low speed user while minimizing the handoff intervention from the MSC. Another practical limitation is the ability to obtain new cell sites.

The cellular concept celarly provides additional capacity through the add' of cellsites, in practise it is difficult for cellular service provider to obtain new physical cell site location in urban area. By using different antenna height and different power level, it is possible to provide large or small cell which are co-located at single location. This technique is named as Umbrella Approach.

The Umbrella cell approach ensures that the number of handoff is minimized for high speed user and provides additional microcell channel for pedestrian handoff is minimized for high speed user and provides additional microcell channel for pedestrian user.



Another practical handoff problem in microcell system is known as cell dragging. cell dragging results from pedestrian user that provide a very strong signal to the BS. such situation occur in an urban environment where there is trade-off or LOS radio path between the subscriber and the BS.

In 1st generation analog cellular system, the typical time to make a handoff, once the signal level is deemed to be below the handoff threshold is about 10sec. This requires that the value for Δ be on the order of 6dB to 12dB.

in digital cellular system 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 require only 1 or 2 seconds. Δ usually between 0dB and 6dB in modern cellular system. The fastest handoff process support a much greater range of options for handling high speed and low speed users and provide the MSC with substantial time to rescue a call that is in need of handoff.

<u>Interference and system capacity</u> :→

Interference is the major limiting factor in the performance of cellular radio system. Two major types of system-generated cellular interference are Co-channel interference and adjacent channel interference.

<u>Co-channel Interference and System capacity</u> 1→

Frequency reuse implies that in a given coverage area there are several cells that use the same set of frequencies. These are called Co-channel cells, and the interference between signals from these cells is called Co-channel interference. Co-channel 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 cell. To reduce Co-channel interference, Co-channel cell must be physically separated by a minimum distance to provide sufficient isolation due to propagation.

When the size 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 become a function of the radius of 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 cell is increased. The parameter Q, called the co-channel reuse ratio,

$$Q = \frac{D}{R} = \sqrt{3N}$$

Q provide larger capacity since the cluster size N is small.

Let i_0 be the number of co-channel interfering cells. Then signal to interference ratio for a mobile receiver which monitors a forward channel can be expressed as

$$\frac{S}{I} = \frac{S}{\sum_{i=1}^{i_0} I_i}$$

S = Desired signal power from the desired B.S

I_i = Interference power caused by the ith interfering co-channel base station.

The average received power P_r at a distance d from transmitting antenna is approximated by

$$P_r = P_0 \left(\frac{d}{d_0}\right)^{-n}$$

$$P_r(dBm) = P_0(dBm) - 10n\log(d/d_0)$$

Also $$\frac{S}{I} = \frac{R^{-n}}{\sum_{i=1}^{i_0} (D_i)^{-n}}$$

$$\frac{S}{I} = \frac{(D/R)^{-n}}{i_0} = \frac{(\sqrt{3N})^{n}}{i_0}$$

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

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

Adjacent channel Interference :→

Interference resulting from signals which are adjacent in frequency to the desired signal is called adjacent channel Interference. Adjacent channel Interference results from imperfect receiver filter which allow nearby frequencies to leak into the passband. The problem can be particularly serious if an adjacent channel user is transmitting in very close range to a subscriber's receiver while the receiver attempts to receive a base station on the desired channel. This is reffered to as near-far effect, where a nearby transmitter capture the receiver of the subscriber. The near-far effect occurs when a mobile close to a base station transmits on a channel close to one being used by a weak mobile. The base station may have difficulty in discriminating desired mobile user from the bleedover caused by the close adjacent channel mobile.

By keeping the frequency separation between each channel in a given cell as large as possible, the adjacent channel interference may be considerably reduced. Thus instead of assigning channels which form a contigous band of frequencies within a particular cell, channels are allocated such that the frequency separation between channels in a given cell is minimized.

If the frequency reuse factor is large, the separation between adjacent channels at the base station may not be sufficient to keep the adjacent channel interfence level within tolerable limit. Eg, if a close-in mobile is 20 times as close as to the basestation as another mobile and has energy spillout of its passband, the signal-to-Interfence ratio at the base station for the weak mobile

$$\frac{S}{I} = (20)^{-n}$$

In practice, base station receiver are preceded by a high Q cavity filter in order to reject adjacent channel interference.

Erlang capacity comparison :→

In order to make a comparison of Erlang capacity firstly we must know about Trunking and Grade of service and based on that we made comparison.

Trunking determines the number of telephone circuits that need to be allocated for office buildings with hundreds of telephones and this same principle is used in designing the cellular radio system.

To design trunked radio system 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 was developed by Erlang. One Erlang represents the amount of traffic intensity (bear his name) carried by a channel that is completely occupied.

The Grade of service is a measure of the ability of a user to access a trunked system during the busiest hour. 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.

The traffic intensity offered by each user is equal to the call requests (per unit time for each user) rate multiplied by the holding Time. That is, each user generates a traffic intensity of Au Erlang given by

$$A_u = \lambda H$$

H = average duration of call

λ = average no of call request per unit time for each user.

For a system containing U users and an unspecified number of channels the total offered traffic intensity A

$$A = UA_U$$

C channel trunked system

$$A_C = UA_U/C$$

The Blocking capacity is given by

$$Pr[\text{blocking}] = \frac{\frac{A^C}{C!}}{\sum_{k=0}^{C} \frac{A^k}{k!}} = GOS \rightarrow \text{Erlang B formula}$$

Erlang C $\rightarrow$

$$Pr[\text{delay} > 0] = \frac{A^C}{A^C + C!\left(1-\frac{A}{C}\right)\sum_{k=0}^{C-1}\frac{A^k}{k!}}$$

Cellular CDMA :→

1.) Many users of a CDMA system share the same frequency.

2.) CDMA has a soft capacity limits.

3.) Multipath fading may be substantially reduced because the signal is spread over a large spectrum.

4.) Channel data rates are very high in CDMA system.

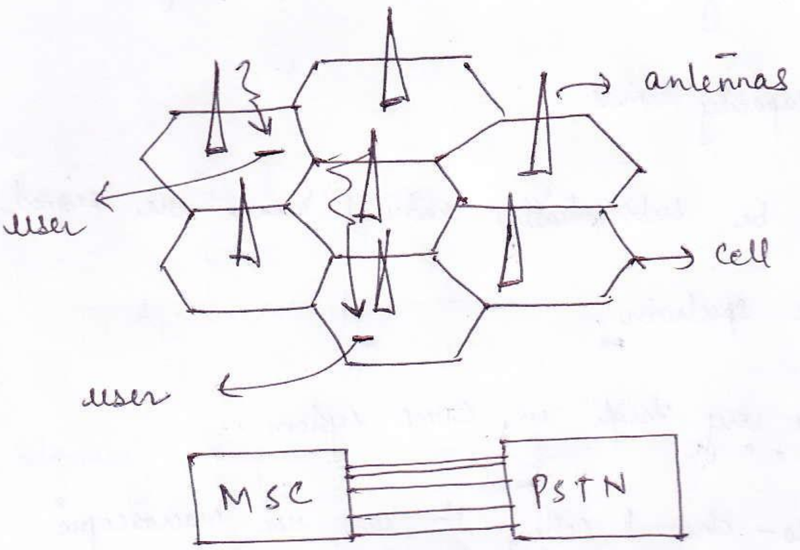
5.) Since CDMA uses co-channel cells, it can use macroscopic spatial diversity to provide soft handoff.

6.) Self-jamming is a problem in CDMA system. Self-jamming arises from the fact that the spreading sequence of different users are not exactly orthogonal.

7.) The near-far problem occurs at a CDMA receiver if an undesired user has a high detected power as compared to the desired user.

Cellular system :→

The following fig shows a basic cellular system which consists of mobile station, base station and 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 mobile station contains a transceiver, an antenna and control circuitry and may be mounted in a vehicle or used as a portable hand-held unit. The BS consists of several transmitter and receivers which simultaneously handle full duplex communications and generally have tower which support several transmitting and receiving antenna. The BS serve as a bridge between all mobile users in the cell and connects the simultaneous mobile call via telephone lines or microwave links to the MSC.

The MSC coordinates the activities of all the BS and connects the entire cellular system to the PSTN. A typical MSC handles 100,000 cellular subscibers and 5000 simullaneous conversations at a time and accomodates all biling and system maintenance functions as well. In large cities, several MSCs are used by single user.

Communication between the BS and the mobile is defined by a standard Common air Interface (CAI) that specifies four different channels. The channel used for voice transmission from the base station to mobile are called forward voice channel (FVC) and the channels used for voice transmission from mobile to the base station are called Reverse voice channel (RVC). Two channels responsible for initating mobile calls are forward control channel (FCC) and Reverse control channel (RCC). Control channels are also named as setup channel because they are only involved in setting up a call and moving it to an unused voice channel. Control channels transmit and receive data message that carry call initiation and service request and are monitored by mobile when they donot have a call in progress. Forward control channels also serve as beacons which continually broadcast all of the traffic request for all the mobile in the system.

Spectrum efficiency :→

Because the frequency spectrum is a limited resource, we should utilize it very effectively. In order to approach this goal, spectrum efficiency should be clearly defined from either a total system point of view or a fixed point-to-point link perspective. For most radio systems, spectrum efficiency is the same as channel efficiency, the maximum number of channels that can be provided in a given frequency band. An appropriate definition of spectrum efficiency for cellular mobile radio is the number of channels per cell. i.e. in cellular radio system

Spectrum efficiency = channel Efficiency.

FDM (Frequency Division Multiplexing) :→

1.) Frequency Division Multiplexing is the oldest and probably still the most widely used channel allocation scheme.

2.) FDM also have some drawbacks. First Guard bands are needed between the channels to keep the stations separated.

3.) Second, the stations must be carefully power controlled.

4.) If the number of stations is small and fixed, the frequency channels can be allocated statically in advance.

5) The common signaling channel was divided into units of 50msec.

A unit contained 50 slots of 1 msec. Each slot was owned by one of 50 ground stations. When a ground station had data to send, it picked a currently unused channel at random and wrote the number of that channel in its next 128-bit slot.

TDM (Time Division Multiplexing) :→

TDM is well understood and widely used in practice. It requires time synchronization for the slots, but this can be provided by a reference station as described for slotted ALOHA.

Advanced Communication Technology satellite (ACTS), which was designed for a few dozen stations. ACTS was launched in 1992 and has four independent 110-Mbps TDM channels, two uplink and two downlink. Each channel is organized as a sequence of 1-msec frames, each frame containing 1728-time slots. Each time slot has a 64-bit payload, allowing each one to hold a 64-kbps voice channel. time slot management is done by one of the ground station the MCS (Master control station). The basic operation of ACTS is a continous three step process, each step lasting 1msec.

UNIT II

Mobile Radio Propagation: Link Calculation and Antenna System

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,

$$P_r(d)= P_tG_tG_r\lambda^2/(4\pi d)^2$$

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 and λ is the wavelength in meters. The gain of an antenna is related to its effective aperture, A_e by,

$$G=4\pi A_e/\lambda^2$$

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

$$\lambda = c/f= 2\pi c/\omega_c$$

where f is the carrier frequency in Hertz, ω_c , is the carrier frequency in radians per second, and c is the speed of light given in meters/s.

An isotropic radiator is an ideal antenna which radiates power with unit gain uniformly in all directions, and is often used to reference antenna gains in wireless systems. The e/7'ective isotropic radiated power (EIRP) is defined as

$$EIRP = P_tG_t$$

and represents the maximum radiated power available from a transmitter in the direction of maximum antenna gain, as compared to an isotropic radiator. In practice, effective radiated power (ERP) is used instead of EIRP to denote the maximum radiated power as compared to a half-wave dipole antenna (instead of an isotropic antenna).

The path loss, which represents signal attenuation as a positive quantity measured in dB, is defined as the difference (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

$$PL\ (dB) = 10\log(P_t/P_r) = -10\log[P_tG_tG_r\lambda^2/(4\pi d)^2]$$

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

$$PL\ (dB) = 10\log(P_t/P_r) = -10\log[\lambda^2/(4\pi d)^2]$$

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 Fraunhofer 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

$$d_f = 2D^2/\lambda$$

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

$$d_f >> D$$

 # If a transmitter produces 50 watts of power, express the transmit power in units of (a) dBm, and (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 Pr(10 km) ? Assume unity gain for the receiver antenna.

<u>The Three Basic Propagation Mechanisms:</u>

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

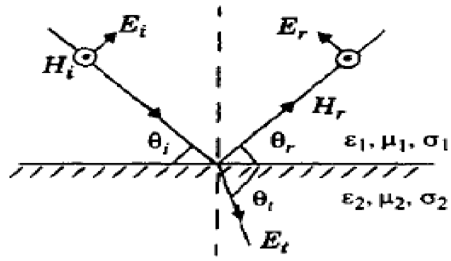
Reflection occurs when a propagating electromagnetic wave impinges upon an object which has very large dimensions when compared to the wavelength of the propagating wave. Reflections occur from the surface of the earth and from buildings and walls.

Diffraction occurs when the radio path between the transmitter and receiver is obstructed by a surface that has sharp irregularities (edges). The secondary waves resulting from the obstructing surface are present throughout the space and even behind the obstacle, giving rise to a bending of waves around the obstacle, even when a line-of-sight path does not exist between transmitter and receiver. At high frequencies, diffraction, like reflection depends on the geometry of the object, as well as the amplitude, phase, and polarization of the incident wave at the point of diffraction.

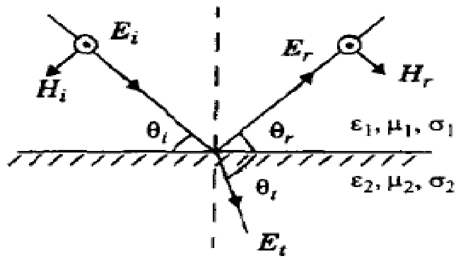
Scattering occurs when the medium through which the wave travels consists of objects with dimensions that are small compared to the wavelength, and where the number of obstacles per unit volume is large. Scattered waves are produced by rough surfaces, small objects, or by other irregularities in the channel. In practice, foliage, street signs, and lamp posts induce scattering in a mobile communications system.

Reflection:

When a radio wave propagating in one medium impinges upon another medium having different electrical properties, the wave is partially reflected and partially transmitted. If the plane wave is incident on a perfect dielectric, part of the energy is transmitted into the second medium and part of the energy is reflected back into the first medium, and there is no loss of energy in absorption. lf the second medium is a perfect conductor, then all incident energy is reflected back into the first medium without loss of energy. The electric field intensity of the reflected and transmitted waves may be related to the incident wave in the medium of origin through the Fresnel reflection coefficient (Γ). The reflection coefficient is a function of the material properties, and generally depends on the wave polarization, angle of incidence, and the frequency of the propagating wave.



(a) E-field in the plane of incidence



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

Figure 3.4
Geometry for calculating the reflection coefficients between two dielectrics.

Reflection from dielectrics:

Figure 3.4 shows an electromagnetic wave incident at an angle θ_i with the plane of the boundary between two dielectric media. As shown in the figure, part of the energy is reflected back to the first media at an angle θ_r, and part of the energy is transmitted (refracted) into the second media at an angle θ_t. The nature of reflection varies with the direction of polarization of the E-field. The behavior for arbitrary directions of polarization can be studied by considering the two distinct cases shown in Figure

The plane of incidence is defined as the plane containing the incident, reflected, and transmitted rays. In Figure 3.4a, the E—field polarization is parallel with the plane of incidence (that is, the E-field has a vertical polarization, or normal component, with respect to the reflecting surface) and in Figure 3.4b, the E-field polarization is perpendicular to the plane of incidence (that is, the incident E-field is pointing out of the page towards the reader, and is perpendicular to the page and parallel to the reflecting surface).

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

$$\Gamma_{\parallel} = \frac{E_r}{E_i} = \frac{\eta_2 \sin\theta_t - \eta_1 \sin\theta_i}{\eta_2 \sin\theta_t + \eta_1 \sin\theta_i} \qquad \textbf{(E-field in plane of incidence)}$$

$$\Gamma_{\perp} = \frac{E_r}{E_i} = \frac{\eta_2 \sin\theta_i - \eta_1 \sin\theta_t}{\eta_2 \sin\theta_i + \eta_1 \sin\theta_t} \qquad \textbf{(E-field not in plane of incidence)}$$

Where η is the intrinsic impedance of the respective medium.

Or,

$$\Gamma_{\parallel} = \frac{-\varepsilon_r \sin\theta_i + \sqrt{\varepsilon_r - \cos^2\theta_i}}{\varepsilon_r \sin\theta_i + \sqrt{\varepsilon_r - \cos^2\theta_i}}$$

$$\Gamma_{\perp} = \frac{\sin\theta_i - \sqrt{\varepsilon_r - \cos^2\theta_i}}{\sin\theta_i + \sqrt{\varepsilon_r - \cos^2\theta_i}}$$

Where ε is the permittivity of the respective medium.

Brewster Angle:

The Brewster angle is the angle at which no reflection occurs in the medium of origin. It occurs when the incident angle BB is such that the reflection coefficient $\Gamma_{\parallel}$ is equal to zero (see Figure 3.6), The Brewster angle is given by the value of θ_B which satisfies

$\mathrm{Sin}(\theta_B) = \sqrt{(\varepsilon_1)}/\sqrt{(\varepsilon_1+\varepsilon_2)}$

For the case when the first medium is free space and the second medium has a relative permittivity ε_r, above equation can be expressed as

$\mathrm{Sin}(\theta_B) = \sqrt{(\varepsilon_r-1)}/\sqrt{(\varepsilon_r^{\,2}-1)}$

Note that the Brewster angle occurs only for vertical (i.e. parallel) polarization.

Calculate the Brewster angle for a wave impinging on ground having a permittivity of $\varepsilon_r = 4$.

Reflection from Perfect Conductors:

Since electromagnetic energy cannot pass through a perfect conductor a plane wave incident on a conductor has all of its energy reflected. As the electric field at the surface of the conductor must be equal to zero at all times in order to obey Maxwel1's equations, the reflected wave must be equal in magnitude to the incident wave. For the case when E-field polarization is in the plane of incidence, the boundary conditions require that

$\theta_i = \theta_r$

and $E_i = E_r$ (E-field in plane of incidence)

Similarly, for the case when the E-field is horizontally polarized, the boundary conditions require that
$\theta_i = \theta_r$
and $E_i = - E_r$ (E-field not in plane of incidence)

<u>Ground Reflection (2-ray) Model:</u>

In a mobile radio channel, a single direct path between the base station and a mobile is seldom the only physical means for propagation, and hence the free space propagation model is in most cases inaccurate when used alone. The 2-ray ground reflection model shown in Figure 3.7 is a useful propagation model that is based on geometric optics, and considers both the direct path and a ground reflected propagation path between transmitter and receiver. This model has been found to be reasonably accurate for predicting the large-scale signal strength over distances of several kilometers for mobile radio systems that use tall towers (heights which exceed 50 m), as well as for line of-sight, microcell channels in urban environments.

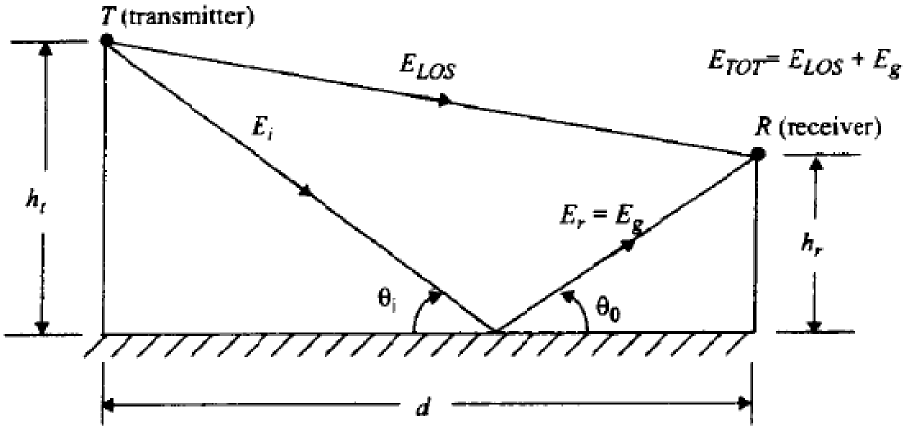


Figure 3.7
Two-ray ground reflection model.

Referring to Figure 3.7, ht is the height of the transmitter and hr is the height of the receiver. lf Eo is the free space E-field (in units of V/m) at a reference distance do from the transmitter, then for d > do , the free space propagating E—field is given by

$$E(d, t) = \frac{E_0 d_0}{d} \cos\left(\omega_c\left(t - \frac{d}{c}\right)\right) \qquad (d > d_0)$$

Two propagating waves arrive at the receiver: the direct wave that travels a distance d'; and the reflected wave that travels a distance d”.

The electric field $E_{TOT}(d, t)$ can be expressed as the sum of equations for distances d’ and d” (i.e. direct wave and reflected wave.

$$E_{TOT}(d, t) = \frac{E_0 d_0}{d'} \cos\left(\omega_c\left(t - \frac{d'}{c}\right)\right) + (-1)\frac{E_0 d_0}{d''} \cos\left(\omega_c\left(t - \frac{d''}{c}\right)\right)$$

Diffraction:

Diffraction allows radio signals to propagate around the curved surface of the earth, beyond the horizon, and to propagate behind obstructions. Although the received field strength decreases rapidly as a receiver moves deeper into the obstructed (shadowed) region, the diffraction field still exists and often has sufficient strength to produce a useful signal.

The phenomenon of diffraction can be explained by Huygen`s principle, which states that all points on a wavefront can be considered as point sources for the production of secondary wavelets, and that these wavelets combine to produce a new wavefront in the direction of` propagation. Diffraction is caused by the propagation of secondary wavelets into a shadowed region. The field strength of a diffracted wave in the shadowed region is the vector sum of the electric field components of all the secondary wavelets in the space around the obstacle.

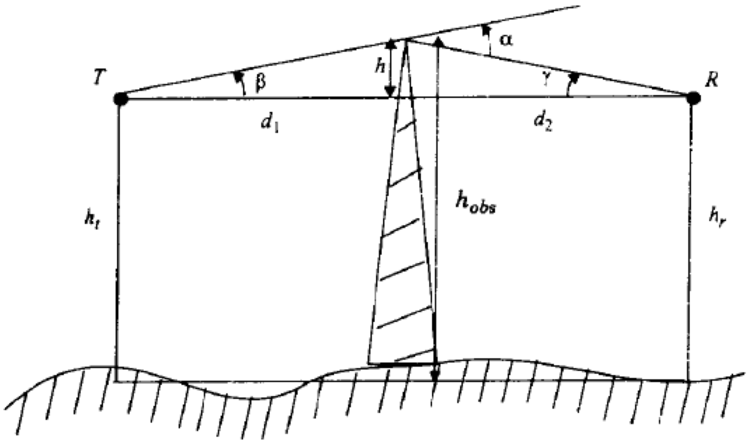
Fresnel Zone Geometry:

Consider a transmitter and receiver separated in free space as shown in Figure 3.10a. Let an obstructing screen of effective height h with infinite width (going into and out of the paper,) be placed between them at a distance d_1 from the transmitter and d_2 from the receiver. It is apparent that the wave propagating from the transmitter to the receiver via the top of the screen travels a longer distance than if a direct line-of-sight path (through the screen) existed. Assuming $h \ll d_1, d_2$ and $h \gg \lambda$, then the difference between the direct path and the diffracted path, called the excess path length (Δ) , can be obtained from the geometry of Figure as

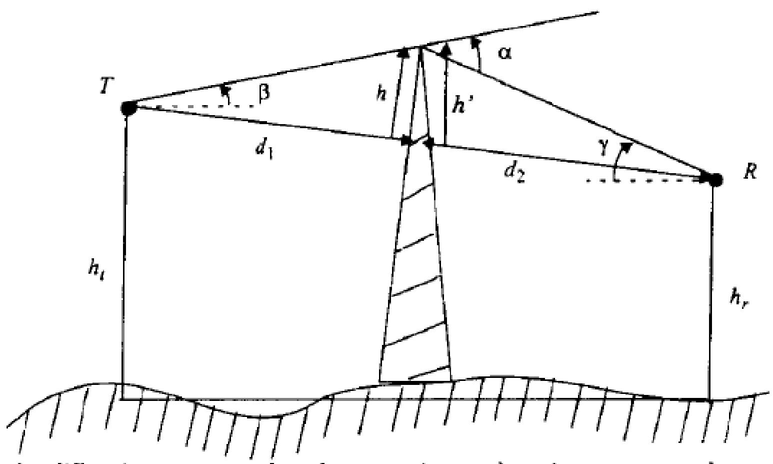
$$\Delta \approx \frac{h^2}{2}\frac{(d_1+d_2)}{d_1 d_2}$$

The corresponding phase difference is given by

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



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



(b) Knife-edge diffraction geometry when the transmitter and receiver are not at the same height. Note that if α and β are small and $h << d_1$ and d_2, then h and h' are virtually identical and the geometry may be redrawn as shown in Figure 3.10c.

Knife-edge Diffraction Model:

Estimating the signal attenuation caused by diffraction of radio waves over hills and buildings is essential in predicting the field strength in a given service area. Generally, it is impossible to make very precise estimates of the diffraction losses, and in practice prediction is a process of theoretical approximation modified by necessary empirical corrections. Though the calculation of diffraction losses over complex and irregular terrain is a mathematically difficult problem, expressions for diffraction losses for many simple cases have been derived. As a starting point, the limiting case of propagation over a knife-edge gives good insight into the order of magnitude of diffraction loss.

When shadowing is caused by a single object such as a hill or mountain, the attenuation caused by diffraction can be estimated by treating the obstruction as a diffracting knife edge. This is the simplest of diffraction models, and the diffraction loss in this case can be readily estimated using the classical Fresnel solution for the field behind a knife edge (also called a half-plane).

Multiple Knife-edge Diffraction:

ln many practical situations, especially in hilly terrain, the propagation path may consist of more than one obstruction, in which case the total diffraction loss due to all of the obstacles must be computed. Bullington suggested that the series of obstacles be replaced by a single equivalent obstacle so that the path loss can be obtained using single knife-edge diffraction models. This method, illustrated in Figure 3.15, oversimplifies the calculations and often provides very optimistic estimates of the received signal strength. In a more rigorous treatment, Millington et. al. gave a wave-theory solution for the field behind two knife edges in series. This solution is very useful and can be applied easily for predicting diffraction losses due to two knife edges. However, extending this to more than two knife edges becomes a formidable mathematical problem. Many models that are mathematically less complicated have been developed to estimate the diffraction losses due to multiple obstructions.

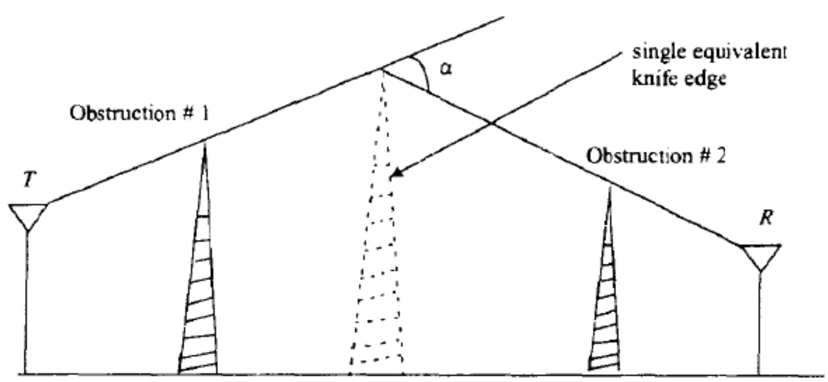


Figure 3.15
Bullington's construction of an equivalent knife edge [From [Bul47] © IEEE]

<u>Scattering:</u>

The actual received signal in a mobile radio environment is often stronger than what is predicted by reflection and diffraction models alone. This is because when a radio wave impinges on a rough surface, the reflected energy is spread out (diffused) in all directions due to scattering. Objects such as lamp posts and trees tend to scatter energy in all directions, thereby providing additional radio energy at a receiver. Flat surfaces that have much larger dimension than a wavelength may be modeled as reflective surfaces. However, the roughness of such surfaces often induces propagation effects different from the specular reflection described earlier in this chapter. Surface roughness is often tested using the Rayleigh criterion which defines a critical height (hc) of surface protuberances for a given angle of incidence i.e. given by

$$\mathbf{hc = \lambda/(8\sin\theta_i)}$$

A surface is considered smooth if its minimum to maximum protuberance h is less than hc, and is considered rough if the protuberance is greater than hc. For rough surfaces, the flat surface reflection coefficient needs to be multiplied by a scattering loss factor, ρs, to account for the diminished reflected field.

Outdoor propagation model:

Longley-Rice Model:

The Longley-Rice model is applicable to point-to-point communication systems in the frequency range from 40 MHz to 100 GHz, over different kinds of terrain. The median transmission loss is predicted using the path geometry of the terrain profile and the refractivity of the troposphere. Geometric optics techniques (primarily the 2-ray ground reflection model) are used to predict signal strengths within the radio horizon. Diffraction losses over isolated obstacles are estimated using the Fresnel-Kirchoff knife-edge models. Forward scatter theory is used to make troposcatter predictions over long distances.

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

Okumura Model:

Okumura`s model is one of the most widely used models for signal prediction in urban areas. This model is applicable for frequencies in the range 150 MHz to 1920 MHz (although it is typically extrapolated up to 3000 MHz) and distances of 1 km to 100 km. It can be used for base station antenna heights ranging from 30 m to 1000 m. Okumura developed a set of curves giving the median attenuation relative to free space (Amu), in an urban area over a quasi-smooth terrain with a base station effective antenna height

(h_{te}) of 200 m and a mobile antenna height (h_{re}) of 3 m. These curves were developed from extensive measurements using vertical omni-directional antennas at both the base and mobile, and are plotted as a function of frequency in the range 100 MHz to 1920 MHz and as a function of distance from the base station in the range 1 km to 100 km. To determine path loss using Okumura`s model, the free space path loss between the points of interest is first determined, and then the value of $A_{mu}(f, d)$ (as read from the curves) is added to it along with correction factors to account for the type of terrain. The model can be expressed as

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

where L50 is the 50th percentile (i.e., median) value of propagation path loss, L_F is the free space propagation loss, A_{mu} is the median attenuation relative to free space, $G(h_{te})$ is the base station antenna height gain factor, $G(h_{re})$ is the mobile antenna height gain factor, and G_{AREA} is the gain due to the type of environment. Note that the antenna height gains are strictly a function of height and have nothing to do with antenna patterns.

$$G(h_{te}) = 20\log\left(\frac{h_{te}}{200}\right) \qquad 1000\text{ m} > h_{te} > 30\text{ m}$$

$$G(h_{re}) = 10\log\left(\frac{h_{re}}{3}\right) \qquad h_{re} \le 3\text{ m}$$

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

Hata Model:

The Hata model [Hat90] is an empirical formulation of the graphical path loss data provided by Okumura, and is valid from 150 MHz to 1500 MHz. Hata presented the urban area propagation loss as a standard formula and supplied correction equations for application to other situations. The standard formula for median path loss in urban areas is given by

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

where fc is the frequency (in MHz) from 150 MHz to 1500 MHz, h_{te} is the effective transmitter (base station) antenna height (in meters) ranging from 30 m to 200 m, h_{re} is the effective receiver (mobile) antenna height (in meters) ranging from 1 m to 10 m, d is the T-R separation distance (in km), and $a(h_{re})$ is the correction factor for effective mobile antenna height which is a function of the size of the coverage area. For a small to medium sized city, the mobile antenna correction factor is given by

$$a(h_{re}) = (1.11ogf_c - 0.7)h_{re} - (1.56\log f_c - 0.8)\text{ dB}$$

and for a large city, it is given by

$a(h_{re}) = 8.29(\log 1.54 h_{re})^2$ -1.1 dB for $f_c <= 300$ MHz

$a(h_{re}) = 3.2(\log 11.75 h_{re})2$ —4.97 dB for $f_c >= 300$ MHz

To obtain the path loss in a suburban area the standard Hata formula in equations are modified as

$$L_{50}(dB) = L_{50}(urban)-2[\log(f_c/28)]^2-5.4$$

and for path loss in open rural areas, the formula is modified as

$$L_{50}(dB) = L_{50}(urban) – 4.78(\log f_c)^2 + 18.33\log f_c - 40.94$$

Indoor Propagation Models:

With the advent of Personal Communication Systems (PCS), 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. This section outlines models for path loss within buildings.

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.

Partition Losses (same floor):

Buildings have a wide variety of partitions and obstacles which form the internal and external structure. Houses typically use a wood frame partition with plaster board to form internal walls and have wood or non-reinforced concrete between floors. Office buildings, on the other hand, often have large open areas (open plan) which are constructed by using moveable office partitions so that the space may be reconfigured easily, and use metal reinforced concrete between floors. Partitions that are formed as part of the building structure are called hard partitions, and partitions that may be moved and which do not span to the ceiling are called soft partitions. Partitions vary widely in their physical and electrical characteristics, making it difficult to apply general models to specific indoor installations.

Partition Losses between Floors:

The losses between floors of a building are determined by the external dimensions and materials of the building, as well as the type of construction used to create the floors and the external surroundings. Even the number of windows in a building and the presence of tinting (which attenuates radio energy) can impact the loss between floors. It can be seen that for all three buildings, the attenuation between one floors of the building is greater than the incremental attenuation caused by each additional floor. After about five or six floor separations, very little additional path loss is experienced.

Propagation Model (Practical Link Budget design)

Free Space Propagation Model

The reduction of the average signal level as the mobile station moves away from the base station is called propagation path loss.

A good path loss prediction model should be able to distinguish between rural, suburban, and urban areas, since there are important differences in radio propagation in these different area types.

Typical Free Space Propagation model is given by

Received Power (P_r) = $P_t + G_t + G_r$ - (32.44+20log (d) +20 log(f)) Where

P_r = Received power in dBm P_t = Transmit power in dBm

G_t = Transmitting antenna gain in dBi.

G_r = Receiving antenna gain in dBi. d = Distance in Km

f = Frequency in MHz

Hata Model

Hata has developed three empirical path loss models based on the measurements by Okumura in Tokyo area of Japan. The Hata model for typical path loss for urban area is given as

$$\mathbf{L_u = 69.55 + 26.16 \log(f) - 13.82 \log(h_B) - a(h_M) + [44.9 - 6.55 \log(h_B)] \log(d)}$$

Where

f=propagation frequency in MHz.
h_B=height of the base station antenna in m
h_M=height of the mobile station antenna in m

d=distance between base station and the mobile station in km.

The term $a(h_M)$ is a correction factor, the value of which depends upon the terrain type. The value of $a(h_M)$ for small and medium sized cities can be found in dB as

$$\mathbf{a(h_M) = [1.1 \log(f) - 0.7] h_M - 1.56 \log(f) + 0.8}$$

and for large cities depending upon frequency as

$$\mathbf{a(h_M) = 8.29[\log(1.54\, h_M)]^2 - 1.10 \qquad \text{when } f \le 200 \text{ MHz}}$$
$$\mathbf{a(h_M) = 3.2[\log(11.75\, h_M)]^2 - 4.97 \qquad \text{when } f \ge 400 \text{ MHz}}$$

For typical suburban area we subtract a correction factor

$$\mathbf{L_{su} = L_u - 2 [\log(f/28)]^2 - 5.4}$$

For open area we have a different correction factor

$$L_r = L_u - 4.78\,[\log(f)]^2 + 18.33\log(f) - 40.94$$

The open area of the Hata model corresponds to flat deserted area. For path loss of typical rural area a margin of 6-10 dB is often added to the path loss predicted by the open area Hata model.

In Hata model a large city is understood to be heavily built with relatively large buildings averaging more than four floors in height. If the city has lower average buildings height, it is considered to be small or medium.

The range of parameter values, where the Hata model is applicable, is

- f = 150 – 1500 MHz
- h_B = 30 – 200 m
- h_M = 1 – 10 m
- d = 1 - 20 km

For higher carrier frequencies of 1500- 2000 MHz the following modification of Hata model for urban area has been proposed. This is also known as Cost231 Hata model and it is the extension of Hata model for *PCS (1900 MHz)*.

$$L_u = 46.3 + 33.9\log(f) - 13.82\log(h_B) - a(h_M) + [44.9 - 6.55\log(h_B)]\log(d) + C$$

Where the correction factor $a(h_M)$ for small and medium-sized cities

$$a(h_M) = [1.1\log(f) - 0.7)]\,h_M - 1.56\log(f) + 0.8$$

These modified equation have been successfully used for cellular mobile network design at 1800 MHz band. However, it should be noted that (modified) Hata model is only valid for macrocell design. The Hata model is not applicable to microcells with d<1 km.

The Hata model also been extended to distances d= 20…100 km with the following modification

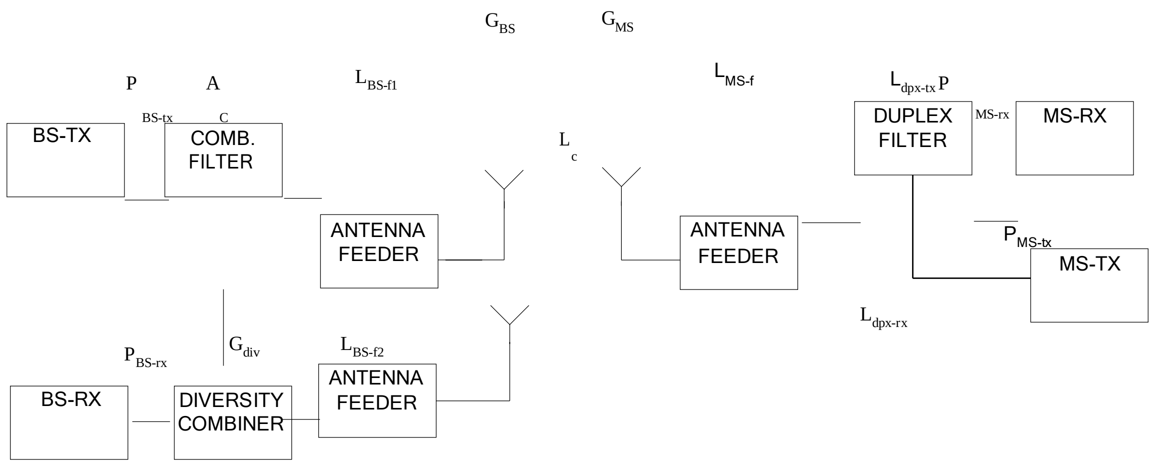
$$\mathbf{L_u = 69.55 + 26.16 \log(f) - 13.82 \log (h_B) - a(h_M) + [\,44.9 - 6.55 \log (h_B)]\,\log(d)]^a}$$

Examples:

1. A transmitter has a power output of 50 W at a carrier frequency of 200MHz. It is connected to an antenna with a gain of 10 dBi. The receiving antenna is 15 Km away and has a gain of 3 dBi. Calculate power delivered to the receiver, assuming free space propagation. Assume also that there are no losses or mismatches in the system.
2. Let us consider a Kathmandu city with different GSM uplink parameters. The MS is transmitting with power 2W .The minimum acceptable received power at BS is -116 dBm. The carrier frequency is 900MHz, the height of base station is 30m and height of mobile station is 1m.Estimate the maximum cell radius and corresponding cell area.

Radio Link Power Budget

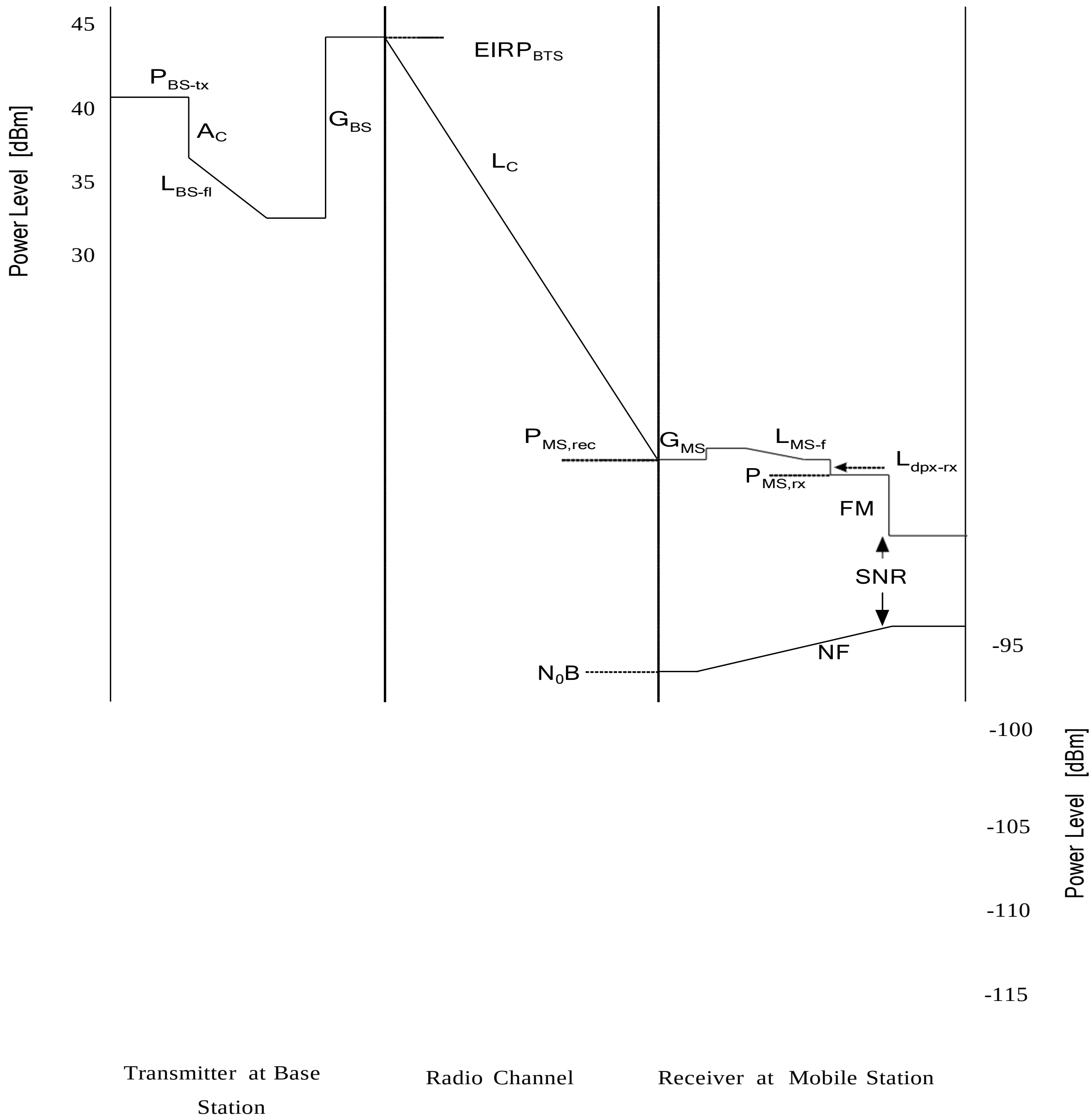
Transmission in uplink and downlink of cellular systems is asymmetric, since the BS transmitter typically uses much higher power than the MS transmitter. However, the transmission quality in uplink and downlink should be equal, especially near the cell edge. Both speech and data services in cellular systems are dimensioned for equal transmission quality in both directions. Transmission quality in uplink and downlink can be determined from the link budget. The terms in the link budget are defined with aid of the block diagram in the adjoining figure.



Where

- P_{BS-tx} is the output power level of the base station
- A_c is the base station transmitter combining filter loss (dB),
- L_{BS-f1} is the Base station antenna feeder loss (dB),
- G_{BS} is the base station antenna gain (dBi, relative to an isotropic radiator),
- L_C is the radio path loss between isotropic antennas (dB),
- G_{MS} is the mobile station gain (dBi),
- L_{MS-f} is the mobile station antenna feeder loss (dB),
- L_{dpx-rx} is the mobile station duplex filter loss in the downlink direction (dB).
- L_{dpx-tx} is the duplex filter loss in the uplink direction (dB),
- G_{div} is the base station diversity gain (dB),
- P_{MS-rx} is the received power level in the mobile station input terminal (dBm),
- P_{BS-rx} is the received power level in the base station input terminal (dBm).

The following Figure shows a pictorial example of a radio link power budget. In addition to the link budget of the transmitted signal, it also illustrates noise at the receiver. The noise has power NoB at the input of the receiver, and it is amplified by the noise figure (NF) of the receiver.



Power Level [dBm]
45
40
35
30
P BS-tx
A C
L BS-fl
G BS
EIRP BTS
L C
P MS,rec
G MS
L MS-f
L dpx-rx
P MS,rx
FM
SNR
NF
N 0 B
-95
-100
-105
-110
-115
Power Level [dBm]
Transmitter at Base Station
Radio Channel
Receiver at Mobile Station

The abbreviation FM stands for fading margin for slow fading. Hence, the signal-to-noise ratio(SNR)in the figure corresponds to the SNR that is encountered, when the receiver is shadowed by an obstruction. The average SNR is better that this value. The SNR during a fade should still be larger than the minimum acceptable E_b / N_o defined for the system. The maximum acceptable path loss can be computed as

$$L_c = EIRP_{max} - P_{rec,\, min}$$

Where

EIRP is Effective Isotropic Radiated Power

P_{rec} is Received power with (hypothetical) isotropic antenna

For downlink:

$EIRP_{BTS,\, max}$ = $P_{BS\text{-}tx}$ maximum BTS power (mean power over burst)

- A_c combiner and filter loss

- $L_{BS\text{-}f1}$ antenna cable loss

+ G_{BS} antenna gain

P_{MS}, rec, min = $P_{MS\text{-}rx}$ MS reference sensitivity (for MS class i)

+ $L_{MS\text{-}f}$ antenna cable loss

- G_{MS} antenna gain

+ $L_{dpx\text{-}rx}$ duplex filter loss

For uplink:

$EIRP_{MS,\,max}$ = $P_{MS\text{-}tx}$ maximum transmission power of MS (mean power over burst for MS class i)

- $L_{dpx\text{-}tx}$ duplex filter loss

- $L_{MS\text{-}f}$ antenna cable loss

+ G_{MS} antenna gain

P_{BTS}, rec, min = $P_{BS\text{-}rx}$ BTS reference sensitivity

+ $L_{BS\text{-}f1}$ antenna cable loss

- G_{BS} antenna gain

- G_{div} diversity gain (if existing)

A balanced power budget is achieved if and only if

$$L_{c,uplink} = L_{c,\,downlink}$$

Which is equivalent to

$$EIRP_{BTS,max} - P_{MS,rec,min} = EIRP_{MS,max} - P_{BTS,rec,min}$$

In reality, balanced power budget is only achieved with an accuracy of ±5 dB, depending primarily on the employed mobile phone class and the initial network design goals. Therefore, uplink and downlink may have different power ranges which may lead to significant performance differences on the cell boundary or at indoor locations.

The acceptable path loss (without interference margin) for any cellular system can be computed separately for uplink and downlink as:

$$L_{c,u} = \left[P_{MS-tx} - L_{dpx-tx} - L_{MS-f} + G_{MS} \right] - \left[P_{BS-rx} + L_{BS-fi} - G_{BS} - G_{div} \right]$$

$$L_{c,d} = \left[P_{BS-tx} - A_c - L_{BS-fi} + G_{BS} \right] - \left[P_{MS-rx} + L_{MS-f} - G_{MS} + L_{dpx-rx} \right]$$

Example: power balance / unbalance in a GSM – system with following parameter values:

- P_{BS-tx} = 10 W (40 dBm)
- P_{MS-tx} = 1 W (30 dBm, handheld telephone)
- P_{MS-tx} = 5 W (37 dBm, car mounted telephone)
- S_{BS} = -104 dBm
- S_{MS} = -102 dBm (handheld telephone)
- S_{MS} = -104 dBm (car mounted telephone)
- L_{dpx-tx} = L_{dpx-rx}
- G_{div} = 7 dB
- A_c = 3 dB

The power unbalance of the handheld telephone:

$$\Delta L = L_{c,d} - L_{c,u} = [P_{BS-tx} - A_c - L_{BS-fi} + G_{BS}]$$

$$- \ [P_{MS-rx} + L_{MS-f} - G_{MS} + L_{dpx-rx}]$$

$$- \ [P_{MS-tx} - L_{dpx-tx} - L_{MS-f} + G_{MS}]$$

$$+ \ [P_{BS-rx} + L_{BS-fi} - G_{BS} - G_{div}]$$

$$= P_{BS-tx} - A_c - P_{MS-rx} - P_{MS-tx} + P_{BS-rx} - G_{div}$$

$$= 40 - 3 + 102 - 30 - 104 - 7 = -2 \text{ dB}$$

The uplink direction is thus 2 dB better. This difference can be neglected in practical network design.

For the car mounted telephone the power unbalance is:

FORMULA

$$\Delta L = P_{BS-tx} - A_c - P_{MS-rx} - P_{MS-tx} + P_{BS-rx} - G_{div}$$

$$= 40 - 3 + 104 - 37 - 7 - 104 = -7 \text{ dB}$$

The uplink direction is now 7 dB better. From the operator's point of view it would be better if the power unbalance were in favor of the downlink. This would guarantee better network control. However the real time power control used in GSM can easily rectify situation and produce almost perfect power balance.

Radio Propagation characterstics :→

Radio wave Propagation :→

The mechanism behind electromagnetic wave propagation are diverse, but can generally attribute to reflection, diffraction and scattering.

Two types of models are there

Large-scale propagation Model → Propagation model that predict the mean signal strength for an arbitary T-R Seperation distance are useful in estimating the radio coverage area of a tx^n and are called LSPM.

Small-scale Propagation model → The model that characterize the rapid fluctuation of the received signal strength over very short travel distance.

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.

The free space power received by a Rx^n antenna which is seperated from a radiating transmitter antenna by a distance di, is given by free space equation

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \quad -(i)$$

where P_t is the transmitted power, $P_r(d)$ is the received power which is a function of T-R Separation. G_t is transmitter antenna gain, G_r is a Rx^n antenna gain, λ is wavelength in mt, d is the T-R Separation distance; L is the system loss factor not related to propagation ($L \geq 1$).

The gain of antenna is related to Effective Apperture

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

Also $\lambda = \frac{c}{f} = \frac{2\pi c}{\omega_c}$

The Friis free space eqn (i) show that received power decays with square of distance b/w T-R. This implies that the Rxed power decays with distance at a rate of 20 dB/decade.

An isotropic radiator is an ideal antenna which radiates power with unit gain uniformly in all directions. The effective radiated power (EIRP) is defined as $\boxed{EIRP = P_t G_t}$

The path loss which represents signal attenuation as a +ve quantity measured in dB and is defined as difference b/w the effective transmitted power and the received power, which may or may not include antenna gain

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

when $G_t = G_r = 1$, then

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

The fraunhofer region, of a transmitting antenna is defined as the region beyond the far-field distance which is related to the largest linear dimension of the Txn antenna aperture.

$$d_f = \frac{2D^2}{\lambda}$$

Also $d_f >> D$

$d_f >> \lambda$.

The received power in free space at a distance greater than do is given by

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

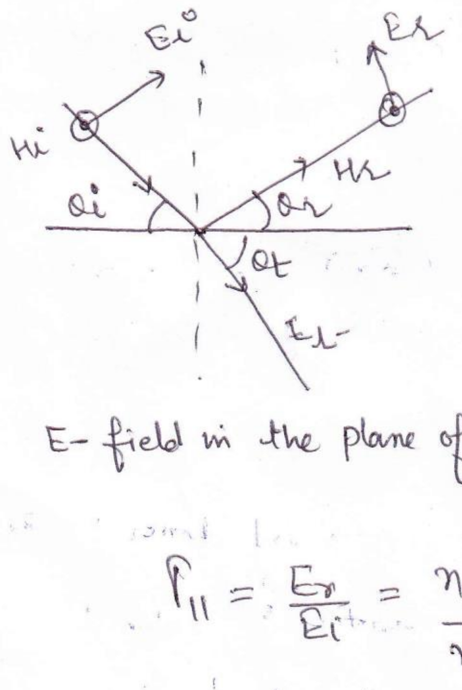
Also $P_r(d)\,dBm = 10\log\left[\frac{P_r(d_0)}{0.001W}\right] + 20\log\left(\frac{d_0}{d}\right)$; $d \geq d_0 \geq d_f$

Where $P_r(d_0)$ is in unit of watts.

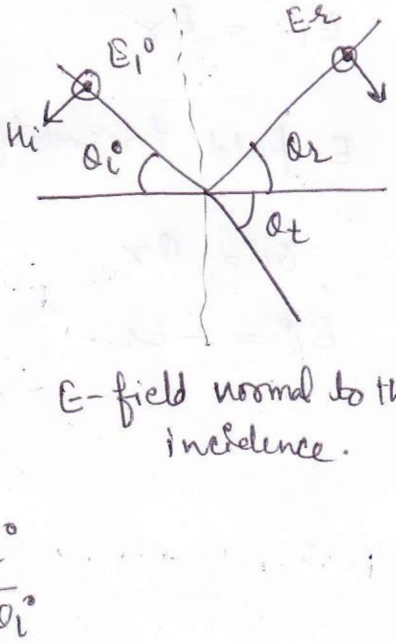
Three Basic Propagation Mechanism :→

(i) Reflection :→ It occurs when a propagating EMW impinges upon an object which has very large dimensions when compared to the wavelength of the propagating wave.

Reflection from Dielectric :→



E-field in the plane of incidence.



E-field normal to the plane of incidence.

$$\Gamma_{\parallel} = \frac{E_r}{E_i} = \frac{\eta_2 \sin\theta_t - \eta_1 \sin\theta_i}{\eta_2 \sin\theta_t + \eta_1 \sin\theta_i}$$

$$\Gamma_{\perp} = \frac{E_r}{E_i} = \frac{\eta_2 \sin\theta_i - \eta_1 \sin\theta_t}{\eta_2 \sin\theta_i + \eta_1 \sin\theta_t}$$

Brewester Angle :→

It is the angle at which no reflection occurs in the medium of origin. It occurs when the incidence angle θ_B is such that the reflection Coff^n Γ_{11} is equal to zero.

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

For the case when the 1st medium is free space and 2nd medium has a relative permittivity ε_r.

$$\sin(\theta_B) = \frac{\sqrt{\varepsilon_r - 1}}{\sqrt{\varepsilon_r^2 - 1}}$$

Reflection from perfect conductor :→

$$\theta_i^\circ = \theta_r$$

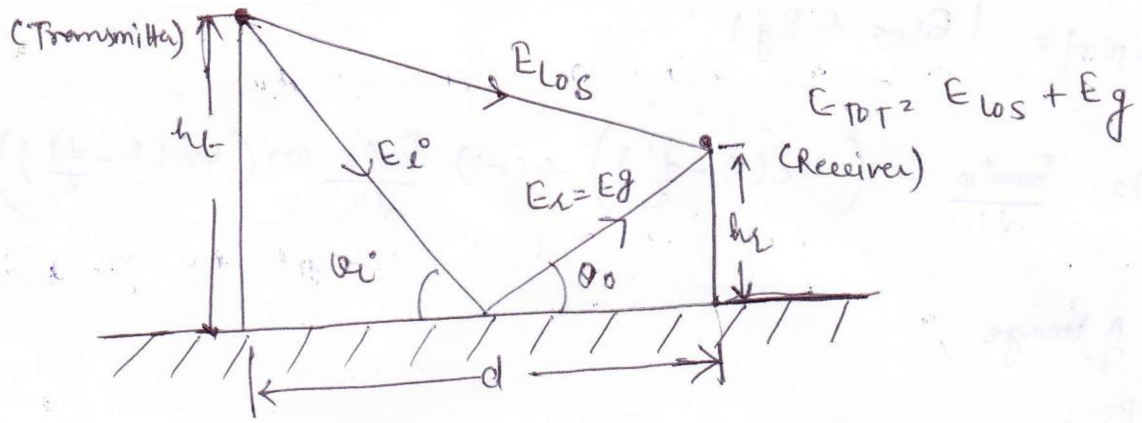
$$E_i^\circ = E_r$$

Case when E-field horizontally polarized.

$$\theta_i^\circ = \theta_r$$

$$E_i^\circ = -E_r.$$

Ground Reflection (Two-Ray) model :→



This is the model based on geometric optics, considered the both the direct path and a ground reflected propagation path b/w transmitter and receiver. This model is good for LSPM.

d = Distance b/w Tx^n and Rx^n

E_{TOT} = Total received field

E_{LOS} = Direct line-of-sight.

E_g = Ground reflected component.

$$E(d,t) = \frac{E_0 d_0}{d} \cos\left[w_c\left(t - \frac{d}{c}\right)\right] \quad (d > d_0)$$

If E_0 is the free space field at reference distance d_0 from Tx^n.

where $E(d,t) = E_0 d_0 / d$ represents envelope of the E-field.

$\|^{ly}$ $$E_{LOS}(d', t) = \frac{E_0 d_0}{d'} \cos\left(w_c\left(t - \frac{d'}{c}\right)\right)$$

$$E_g(d'', t) = \frac{E_0 d_0}{d''} \cos\left[w_c\left(t - \frac{d''}{c}\right)\right)$$

Acc^n to law of reflection

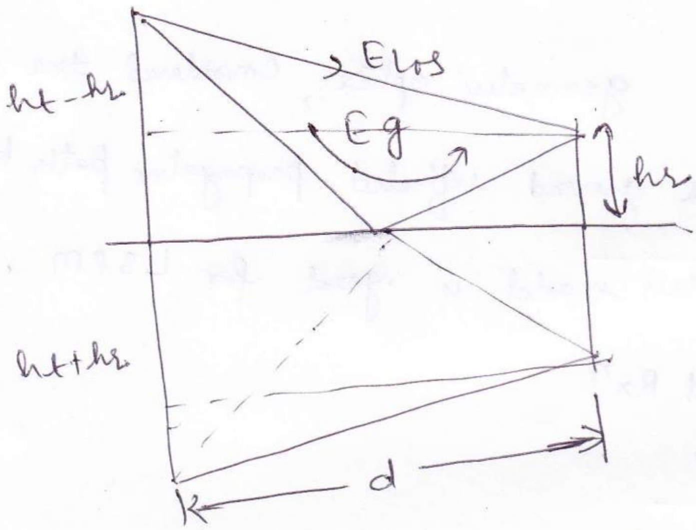
$$\theta_i = \theta_0$$
$$E_g = \Gamma E_i$$
$$E_i = (1+\Gamma) E_t$$

Also we know that

$$|E_{TOT}| = |E_{los} + E_g|$$

$$\therefore E_{TOT}(d,t) = \frac{E_0 d_0}{d'} \cos\left(w_c\left(t - \frac{d'}{c}\right)\right) + (-1)\frac{E_0 d_0}{d''} \cos\left(w_c\left(t - \frac{d''}{c}\right)\right)$$

using Method of image,



The path difference $\Delta = d'' - d'$

$$\Delta = \sqrt{(h_t + h_r)^2 + d^2} - \sqrt{(h_t - h_r)^2 + d^2}$$

$$\Delta = d'' - d' = \frac{2 h_t h_r}{d}.$$

Phase difference $\theta_\Delta = \frac{2\pi\Delta}{d} = \frac{\Delta w}{c}$

Time delay $\tau_d = \frac{\Delta}{c} = \frac{\theta_\Delta}{2\pi f_c}$

Also $\left|\frac{E_0 d_0}{d}\right| = \left|\frac{E_0 d_0}{d'}\right| = \left|\frac{E_0 d_0}{d''}\right|$

of the received E-field is evaluated at some time say at $t = d''/c$

then

$$E_{(TOT)}\left[d, t = \frac{d''}{c}\right] = \frac{E_o d_o}{d'} \cos\left(w_c\left(\frac{d''}{c} - \frac{d'}{c}\right)\right) - \frac{E_o d_o}{d''}\cos 0^\circ$$

$$= \frac{E_o d_o}{d'} \angle \theta_\Delta - \frac{E_o d_o}{d''}$$

$$= \frac{E_o d_o}{d}\left(\angle \theta_\Delta - 1\right]$$

Also

$$|E_{TOT}(d)| = \sqrt{\left(\frac{E_o d_o}{d}\right)^2 (\cos\theta_\Delta - 1)^2 + \left(\frac{E_o d_o}{d}\right)^2 \sin^2\theta_\Delta}$$

or

$$|E_{TOT}(d)| = \frac{E_o d_o}{d}\sqrt{2 - 2\cos\theta_\Delta}$$

$$|E_{TOT}(d)| = \frac{2E_o d_o}{d}\sin\left(\frac{\theta_\Delta}{2}\right)$$

where

$$\frac{\theta_\Delta}{2} = \frac{2\pi h_t h_r}{\lambda d} < 0.3 \text{ radians}$$

$$d > \frac{20\pi h_t h_r}{3\lambda} = \frac{20 h_t h_r}{\lambda}$$

$$E_{TOT}(d) = \frac{2E_o d_o}{d}\,\frac{2\pi h_t h_r}{\lambda d} = \frac{k}{d^2}\ \text{V/m}$$

Now

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

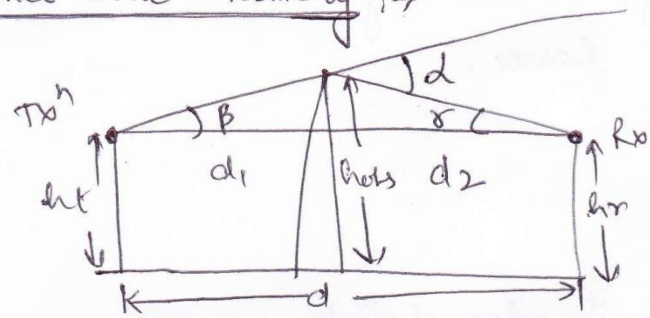
Also

$$PL(dB) = 40\log d - (10\log G_t + 10\log G_r + 20\log h_t + 20\log h_r)$$

Diffraction :→

It occurs when radio path between transmitter and receiver is obstructed by a surface that has sharp irregularities.

Fresnel Zone Geometry :→



In this Model an obstacle is placed between Txn and Rxn

d_1 is the distance between Txn and front end of obstacle and d_2 is the distance between Back end of obstacle and Rxn.

In this case the path difference is

$$\Delta = \frac{h^2}{2}\frac{(d_1+d_2)}{d_1 d_2}$$

The corresponding phase difference is

$$\phi = \frac{2\pi\Delta}{\lambda} = \frac{2\pi}{\lambda}\frac{h^2}{2}\frac{(d_1+d_2)}{d_1 d_2}$$

When $\tan x = x$ then $\alpha = \beta + \gamma$ using trignometric geometry

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

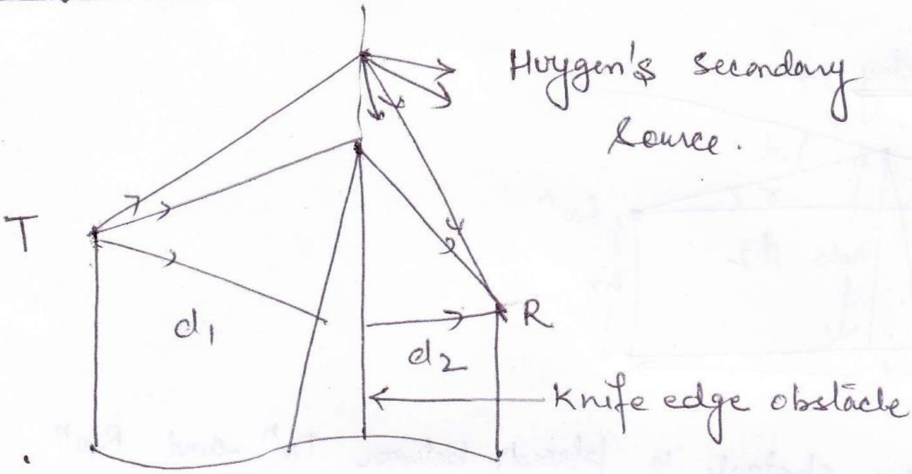
Also Fresnel - kirchoff diffraction parameter v is given by

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

The corresponding phase difference is

$$\phi = \frac{\pi}{2} v^2 .$$

Knife-edge Diffraction Model :→



When shadowing is caused by a single-object such as hill or mountain, the attenuation caused by diffraction can be estimated by treating the obstruction as a diffracting knife edge. This is the simplest of diffraction model, and the diffraction loss in this case can be readily estimated using the classical Fresnel sol^n for the field behind a knife edge.

The electric-field strength E_d, of a knife-edge diffraction wave is given by

$$\frac{E_d}{E_0} = F(v) = \frac{(1+j)}{2} \int_v^{\infty} \exp\left((-j\pi t^2)/2\right) dt .$$

$F(v)$ = Complex Fresnel geometry.

The diffraction gain due to the presence of knife edge is given by

$$G_d(dB) = 20\log\left|F(v)\right|$$

An approximate solⁿ to above eqⁿ is

$$G_d(dB) = 0 \quad , \quad v \leq -1$$

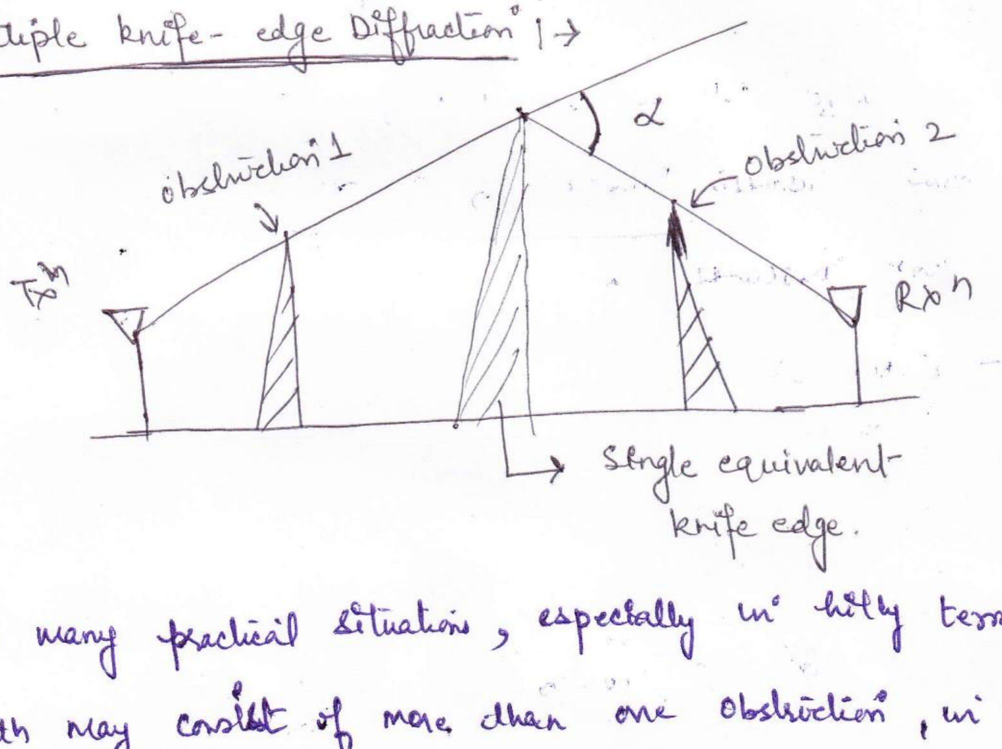
$$G_d(dB) = 20\log(0.5 - 0.62v) \quad , \quad -1 \leq v \leq 0$$

$$G_d(dB) = 20\log(0.5\exp(-0.95v)) \quad , \quad 0 \leq v \leq 1$$

$$G_d(dB) = 20\log\left(0.4 - \sqrt{0.1184 - (0.38 - 0.1v)^2}\right) \quad ; \quad 1 \leq v \leq 2.4$$

$$G_d(dB) = 20\log\left(\frac{0.225}{v}\right)$$

<u>Multiple knife- edge Diffraction</u> :→



In many practical situation, especially in hilly terrain, the propagation path may consist of more than one obstruction, in which case the total diffraction loss due to all of the obstacles must be computed. The series of obstacles can be replaced by a single equivalent obstacle so that the path loss can be obtained using single knife-edge diffraction model.

Scattering :→

It occurs when the medium through which the wave travels consist of objects with dimensions that are small compared to the wavelength and when the number of obstacles per unit volume is large.

Scattering is caterized / characterized into ~~Two~~ models named as

Radar cross section Model :→

In radio channels when large distant object Induce scattering knowledge of the physical location of such objects can be used to accurately predict scattered signal strength. The Radar cross section (RCS) of a scattering object is defined as the Ratio of the power density of the signal scattered in the direction of the receiver to the power density of the Radio wave incident upon the scattering object.

$$P_R(dB) = P_T(dBm) + G_T(dBi) + 20\log(d) + RCS[dBm^2] - 30\log(4\pi) - 20\log d_T - 20\log d_R.$$

Models for Path Loss :→

As discussed earlier path loss is the difference b/w Transmitter power and the received power.

Log- Distance Path loss Model :→

Practically and theoretically it indicates that the average received signal power decreases logarithmically; with distance

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

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

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

n = path loss exponent

d_0 = close-in refrence distance

d = T-R separation.

The bars denote the ensemble average of all possible path loss value for a given value of d.

Log-Normal Shadowing :->

The above model doesnot consider the fact that the surrounding enviornmental clutter may be vastly different at two different locations having the same T-R separation.

$$PL(d)[dB] = \overline{PL}(d) + X_\sigma = \overline{PL}(d_0) + 10n \log\left(\frac{d}{d_0}\right) + X_\sigma$$

and

$$P_r(d)[dBm] = P_t[dBm] - PL(d)[dB]$$

X_σ = Zero-mean Gaussian

σ = Standard deviation.

Outdoor propagation Model :->

(i) Longley - Rice Model :->

(i) This model is applicable to point-to-point communication system in frequency range from 40MHz to 100MHz. Geometric optic technique are used to predict signal strength within the radio horizon.

(ii) The LRM propagation prediction model is also reffered to as the ITS irregular Terrain model.

(iii) This model is also available as a computer program to calculate large-scale median transmission loss relative to free space loss over irregular terrain for frequencies b/w 20MHz to 10GHz.

(iv) This model operates in Two modes point-to-point mode and Area Mode.

(v) One Shortcoming of this Model is that it doesnot provide a way of determing correction due to environmental factor in the immediate vicinity of the Mobile Rxn.

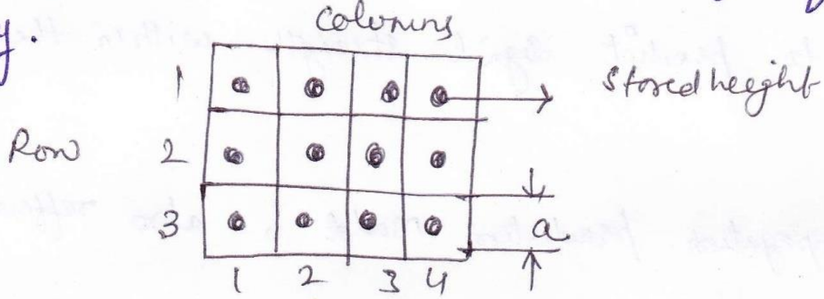
(ii) Durkin's Model :->

This Model is used for predicting field strength Contours over irregular terrain;

The execution of Durkin path loss simulator consist of Two parts.

(i) The 1st part accesses a topographic data base of a proposed service area and reconstruct the ground profile information along the radial joining the transmitter to the Rx?.

The topographical data base can be thought of as a two-dimensional array.



This Durkin's Model is very attractive because it can read in a digital elevation map and perform a site-specific propagation computation on the elevation data. The disadvantage are that it cannot adequately predict propagation effect due to fogile (foliage), buildings other man-made structures and It doesnot account for multipath propagation other than ground reflection, so additional loss factors are often included.

Okumura Model :→

This model is one of the most widely used models for signal predictions in urban area. This model is applicable for frequencies 150 MHz to 1920 MHz and distance 1 km to 100 km. For calculating the signal strength following eq^{n}s are developed.

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

L_{50} = 50th percentile value of PL

L_f = free space propagation loss

A_{mu} = Median Attenuation related to free space

$G(h_{te})$ $G(h_{re})$ = Mobile antenna height Gain factor for Tx^n and Rx^n

G_{Area} = Gain due to the type of environment.

$$G(h_{te}) = 20\log\left(\frac{h_{te}}{200}\right), \quad 1000\,m > h_{te} > 30\,m$$

$$G(h_{re}) = 10\log\left(\frac{h_{re}}{3}\right), \quad h_{re} \leq 3\,m$$

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

Hata Model :→

This model is an empirical formulation of the graphical path loss data provided by Okumura and is valid for 150 MHz to 1500 MHz.

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

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

To obtain path loss, then

$$L_{50}(dB) = L_{50}(\text{urban}) - 2\left[\log (f_c/28)\right]^2 - 5.4$$

Also for Rural Areas

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

PCS Extension to Hata Model :→

It is the extended version of Hata Model.

$$L_{50}(\text{urban}) = 46.3 + 33.9 \log f_c - 13.82 \log h_{te} - a(h_{re}) + (44.9 - 6.55 \log h_{te}) \log d + C_M$$

- $a(h_{re})$ is defined earlier.

$C_M =$ 0 dB for medium sized city and suburban areas

3 dB for Metropolitan cities.

The COST-231 extension of Hata model is restricted to the following ranges

Indoor Propagation Model :→

Partition losses :→

Buildings have a wide variety of partitions and obstacles which form the internal and external structure. office buildings, on the other hand often have large open areas which are constructed by using moveable office partitions so that the space may be reconfigured easily.

Partitions that are formed as part of the building structure are called Hard partition, and partitions that may be moved and which donot span to the ceiling are called soft partition.

Partition losses between floor :→

the losses between floor of a building are determined by the external dimensions and materials of the building as well as the type of construction used to create the floor and the external surroundings.

Log-distance Path loss Model :→

Indoor path loss obeys the distance power law

$$PL(dB) = PL(d_0) + 10n \log\left(\frac{d}{d_0}\right) + X_\sigma$$

where the value of n depends on the surroundings and buildings type and X_σ represent's a normal random variable in dB having standard deviation of σ dB.

Ericsson Multiple Breakpoint Model :→

The Ericsson radio system Model was obtained by Measurements in a multiple floor office building. The model has four breakpoints both an upper and lower bound on the pathloss. This model assume that there is 30 dB attenuation at $d_0 = 1m$, which can be shown to be accurate for $f = 300 MHz$ and unity-gain antennas.

Shadowing and Multipath fading- delay spread :→

Small-scale fading or simply fading, is used to describe the rapid fluctuations of amplitude, phases or multipath delay of a radio signal over a short period of time or travel distance so that large-scale pathloss effect may be ignored.

Small-Scale Multipath propagation :→

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

(i) Rapid change in signal strength over small travel distance or time travel.

(ii) Random frequency modulation due to varying Doppler shift on different multipath signal.

(iii) Time dispersion caused by multipath propagation Delays.

$f = 1500\,MHz$ to $2000\,MHz$

$h_{te} = 30m$ to $200m$

$h_{re} = 1m$ to $10m$

$d = 1km$ to $20km$.

Walfisch and Bertoni Model :→

A model developed by walfisch and Bertoni consider the impact of rooftops and buildings height by using diffraction to predict average s/g strength at street level. The model consider the path loss S, to be a product of three factors

$$S = P_0 Q^2 P_1$$

P_0 represents free space path loss b/w isotropic antenna given by

$$P_0 = \left(\frac{\lambda}{4\pi R}\right)^2$$

Q^2 = gives the reduction in the rooftop due to row of buildings

P_1 = term is based upon diffraction and determine the signal loss from the rooftop to the street.

In dB path loss is given by

$$S(dB) = L_0 + L_{rts} + L_{ms}$$

L_0 = free space loss

L_{rts} = Roof to street loss

Factor Influencing Small-Scale fading :→

(i) Multipath Propagation :→

The presence of reflecting objects and scatters in the channel creates a constantly changing environment that dissipates the signal energy in amplitude, phase and time.

(ii) Speed of Mobile :→

The relative motion between the base station and the mobile result in random frequency modulation due to different Doppler shift on each of the multipath component.

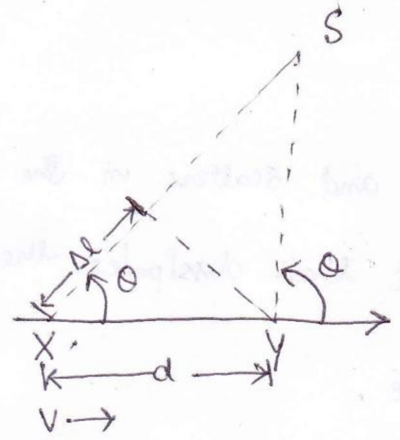
(iii) Speed of Surrounding object :→

If objects in the radio channel are in motion, they induce a time varying Doppler shift on multipath component.

(iv) 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, but the received signal strength will not fade such over a local area.

Doppler shift :→



Consider a mobile moving at a constant velocity v, along a path segement having length d between points X & Y, while it receives signal from a remote source S. The difference in path lengths travelled by the wave from source S to the mobile at point X & Y is $\Delta l = d\cos\theta = v\Delta t \cos\theta$, where Δt is the time required for the mobile to travel from X to Y.

The phase change in the received signal due to the difference in path length is ;

$$\Delta\phi = \frac{2\pi\Delta l}{d} = \frac{2\pi v \Delta t}{d}\cos\theta$$

and hence the apparent change in frequency, or Doppler shift is given by fd,

$$f_d = \frac{1}{2\pi}\frac{\Delta\phi}{\Delta t} = \frac{v}{d}\cdot\cos\theta$$

where
$$\overline{\tau^2} = \frac{\sum_k a_k^2 \tau_k^2}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\tau_k^2}{\sum_k P(\tau_k)}$$

Coherence Bandwidth :→

while the delay spread is a natural phenomenon caused by reflected and scattered propagation path in the radio channel, the coherence Bandwidth B_C, is defined relation derived from the rms delay spread.

Coherence Bandwidth is a statistical measure of the range of frequencies over which the channel can be considered flat which means the channel that passes all the spectral component approximately equal gain & linear phase.

$$\boxed{B_C = \frac{1}{50\sigma_\tau}}$$

If the def^n is relaxed so that the frequency correlation function is above 0.5, then

$$\boxed{B_C = \frac{1}{5\sigma_\tau}}$$

Parameter of Mobile Multipath Channel :→

(i) Time Dispersion parameter (Doppler Spread)

(ii) Coherence Bandwidth

(iii) Coherence Time.

Doppler Spread :→

In order to compare different multipath channels and to develop some general design guidelines for wireless system. The mean excess delay, rms delay spread, excess delay spread, are the multipath channel parameter that can be determined from a power delay profile.

The time dispersive properties of wideband multipath channels are most commonly quantified by their mean excess delay ($\bar{\tau}$) and rms delay spread (σ_τ).

The mean excess delay

$$\bar{\tau} = \frac{\sum_K a_k^2 \tau_k}{\sum_K a_k^2} = \frac{\sum_K P(\tau_k)\tau_k}{\sum_K P(\tau_k)}$$

The rms delay spread

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

Doppler spread and Coherence Time :→

Delay spread and coherence bandwidth are parameters which describe the time dispersive nature of the channel in a local area.

Doppler spread and Coherence time are parameters which describe the time varying nature of the channel in small-scale region.

Doppler spread (B_D) →

It is a measure of the spectral broadening caused by the time rate of change of the mobile radio channel and is defined as the range of frequencies over which the doppler spectrum is "essentially non-zero.

If the baseband signal bandwidth is much greater than B_D, the effect of Doppler spread are negligible at the receiver. This is a slow fading channel.

Coherence Time (T_C) → It is the time domain dual of Doppler spread and is used to characterize the time varying nature of the frequency dispersiveness of the channel in the time domain.

$$T_C = \frac{1}{f_m}$$

If coherence time is defined as the time over which the time correlation function is above 0.5 then the T_C is

$$T_C = \frac{9}{16\pi f_m}$$

f_m = max^n Doppler shift

$$f_m = v/\lambda$$

$$T_C = \sqrt{\frac{9}{16\pi f_m^2}} = \frac{0.423}{f_m}$$

Multipath Fading :→

Types of Small-Scale Fading :→

Depending on the relation between the signal parameters such as bandwidth symbol period etc. and the channel parameter such as rms delay spread and Doppler Spread, different transmitted signals will undergo different type of fading.

Multipath delay spread leads to Time dispersion and frequency selective fading, while Doppler Spread leads to frequency dispersion and time selective fading.

Fading effect Due to Multipath Time Delay Spread :→

Small-Scale Fading
(Based on Multipath time delay spread)

Flat fading

1. BW of Signal < BW of Channel
2. Delay Spread < Symbol period

Frequency selective fading

1. BW of signal > BW of channel
2. Delay spread > Symbol period

Flat fading :→

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

A signal undergoes flat fading if

$$B_S << B_C$$

$$T_S >> \sigma_\tau$$

Frequency selective fading :→

If the channel possesses a constant gain and linear phase response over a bandwidth that is smaller than the bandwidth of transmitted signal, then the channel creates frequency selective fading on the received signal.

A signal undergoes frequency selective fading if

$$B_S > B_C$$

$$T_S < \sigma_\tau$$

Fading effect Due to Doppler spread :→

Small scale fading
(Based on Doppler spread)

Fast fading	Slow fading
1. High Doppler spread	1. low Doppler spread
2. coherence time < Symbol period	2. Coherence time > Symbol period
3. Channel variation faster than baseband signal variation	3. channel variation slower than baseband signal

Fast fading :→

In this, the channel Impulse response changes rapidly within the symbol duration ie the coherence time of the channel is smaller than the symbol period of the transmitted signal. This cause frequency dispersion due to Doppler spreading, which leads to signal distortion.

The signal undergoes fast fading if

$$T_S > T_C$$

$$B_S < B_D$$

Slow Fading :→

In this, the channel impulse response changes at rate much slower than the transmitted baseband signal. In this case, the channel may be assumed to be static over one or several reciprocal bandwidth intervals. This implies that the Doppler spread of the channel is much less than the bandwidth of the baseband signal.

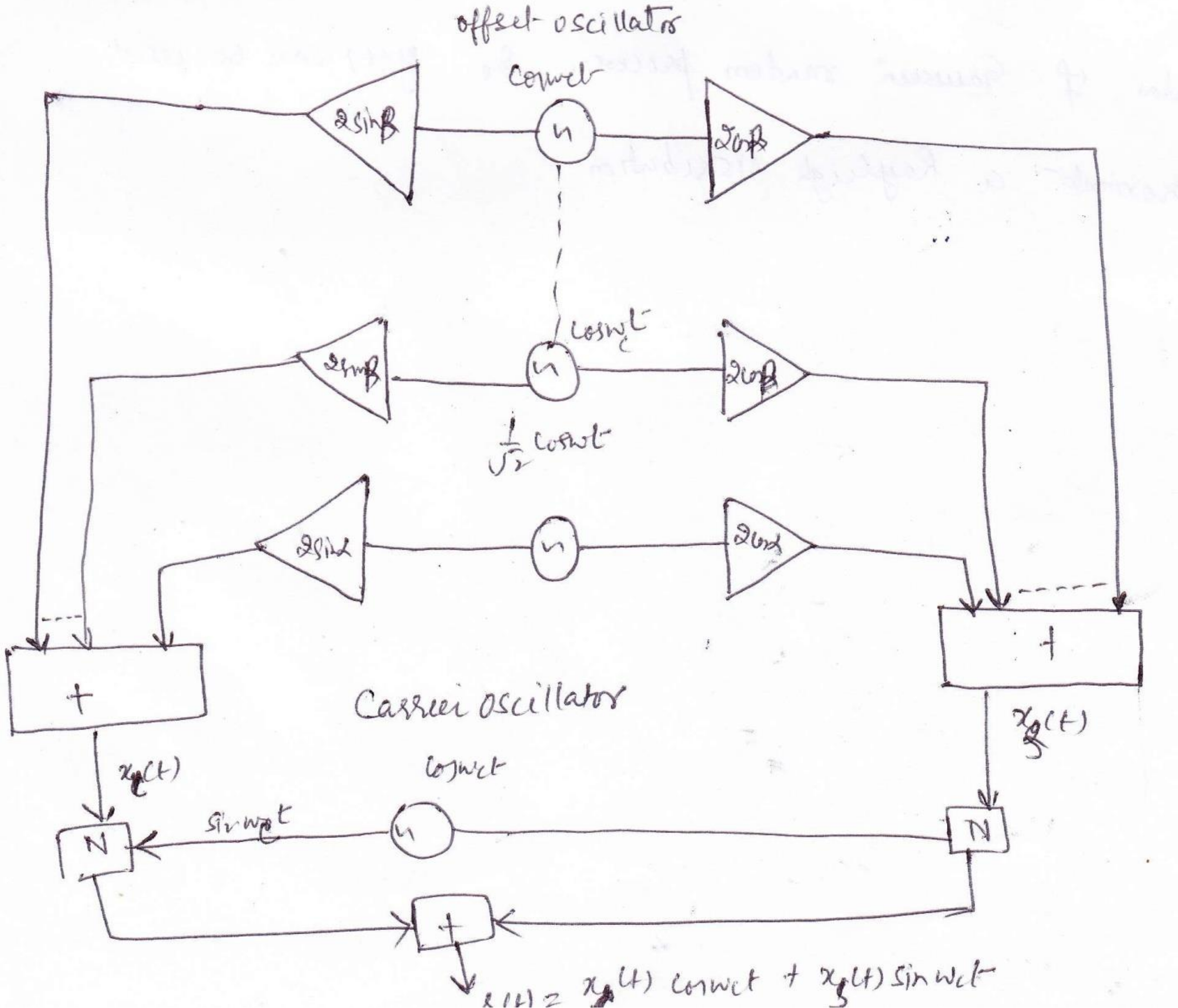
A signal undergoes slow fading if

$$T_S << T_C$$

$$B_S >> B_D$$

Jake's Model :→

Jake's method is a mathematical model suitable to simulated the Rayleigh fading. The computation load is far lower than a model by doing FFT to the Doppler spectrum. It is simplified model of the ring scattering model, which assume the arrival phases of received multipath signals are uniformed distributed from 0 to 2π after passing through different scattering path and angles. In Jake's model, it uses several sinusoids to approximate this effect. The following fig shows the block diagram of Jake's model.



the Simmation of the Sinusoid is

$$y(t) = x_c(t)\cos\omega_c t + x_s(t)\sin(\omega_c t)$$

$$x_c(t) = 2\sum_{n=1}^{N_0} \cos(\beta_n)\cos(\omega_n t) + \sqrt{2}\cos(\alpha)\cos(\omega_m t)$$

$$\omega_m = \frac{2\pi V}{\lambda}$$

$$\omega_n = \omega_m \cos\left(\frac{2\pi n}{N}\right)$$

$$N_0 = \frac{1}{2}\left(\frac{N}{2} - 1\right)$$

It can be observed that Jake's model samples more frequencies near the maximum Doppler shift when $N_0 \leq 8$, $x_c(t)$ and $x_s(t)$ are approximation of Gaussian random process. So $y(t)$ can be used to approximate a Rayleigh distribution.

UNIT III
Mobile Radio Propagation
Introduction

There are two basic ways of transmitting an electro-magnetic (EM) signal, through a guided medium or through an unguided medium. Guided mediums such as coaxial cables and fiber optic cables, are far less hostile toward the information carrying EM signal than the wireless or the unguided medium. It presents challenges and conditions which are unique for this kind of transmissions. A signal, as it travels through the wireless channel, undergoes many kinds of propagation effects such as reflection, diffraction and scattering, due to the presence of buildings, mountains and other such obstructions. Reflection occurs when the EM waves impinge on objects which are much greater than the wavelength of the traveling wave. Diffraction is a phenomena occurring when the wave interacts with a surface having sharp irregularities. Scattering occurs when the medium through the wave is traveling contains objects which are much smaller than the wavelength of the EM wave. These varied phenomena's lead to large scale and small scale propagation losses. Due to the inherent randomness associated with such channels they are best described with the help of statistical models. Models which predict the mean signal strength for arbitrary transmitter receiver distances are termed as large scale propagation models. These are termed so because they predict the average signal strength for large Tx-Rx separations, typically for hundreds of kilometers.

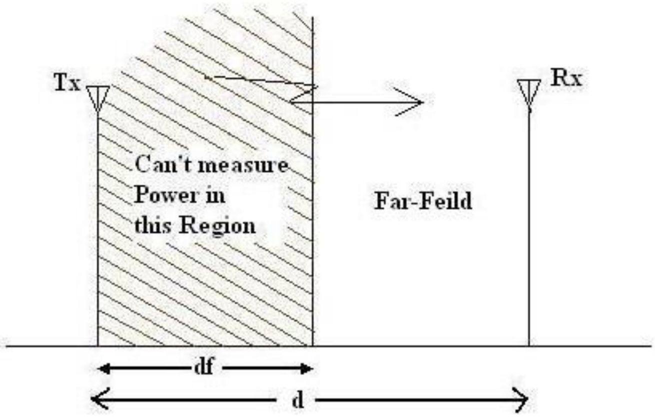


Figure 3.1: Free space propagation model, showing the near and far fields.

Free Space Propagation Model

Although EM signals when traveling through wireless channels experience fading effects due to various effects, but in some cases the transmission is with a direct line of sight such as in satellite communication. Free space model predicts that the received power decays as negative square root of the distance. Friis free space equation is given by

$$P_r(d) = \frac{P_t tt_t tt_r \lambda^2}{(4\pi)^2 d^2 L} \tag{3.1}$$

where P_t is the transmitted power, $P_r(d)$ is the received power, tt_t is the transmitter antenna gain, tt_r is the receiver antenna gain, d is the Tx-Rx separation and L is the system loss factor depended upon line attenuation, filter losses and antenna losses and not related to propagation. The gain of the antenna is related to the effective aperture of the antenna which in turn is dependent upon the physical size of the antenna as given below

$$tt = 4\pi A_e/\lambda^2. \tag{3.2}$$

The path loss, representing the attenuation suffered by the signal as it travels through the wireless channel is given by the difference of the transmitted and re- ceived power in dB and is expressed as:

$$PL(dB) = 10 \log P_t/P_r. \tag{4.3}$$

The fields of an antenna can broadly be classified in two regions, the far field and the near field. It is in the far field that the propagating waves act as plane waves and the power decays inversely with distance. The far field region is also termed as Fraunhofer region and the Friis equation holds in this region. Hence, the Friis
equation is used only beyond the far field distance, d_f, which is dependent upon the largest dimension of the antenna as

$$d_f = 2D^2/\lambda. \tag{3.4}$$

Also we can see that the Friis equation is not defined for d=0. For this reason, we use a close in distance, d_o, as a reference point. The power received, $P_r(d)$, is then given by:

$$P_r(d) = P_r(d_o)(d_o/d)^2. \tag{3.5}$$

Ex. 1: Find the far field distance for a circular antenna with maximum dimension of 1 m and operating frequency of 900 MHz.

Solution: Since the operating frequency f = 900 Mhz, the wavelength

$$\lambda = \frac{3 \times 10^8 m/s}{900 \times 10^6 Hz} m$$

. Thus, with the largest dimension of the antenna, D=1m, the far field distance is

$$d_f = \frac{2D^2}{\lambda} = \frac{2(1)^2}{0.33} = 6m$$

.

Ex. 2: A unit gain antenna with a maximum dimension of 1 m produces 50 W power at 900 MHz. Find (i) the transmit power in dBm and dB, (ii) the received power at a free space distance of 5 m and 100 m.

Solution:

(i) Tx power = 10log(50) = 17 dB = (17+30) dBm = 47 dBm

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

Thus the received power at 5 m can not be calculated using free space distance formula.

At 100 m ,

$$P_R = \frac{P_T tt_T tt_R \lambda^2}{4\pi d^2}$$

$$= \frac{50 \times 1 \times (1/3)^2}{4\pi 100^2}$$

$$= 3.5 \times 10^{-3} mW$$

$$P_R(dBm) = 10logP_r(mW) = -24.5dBm$$

Basic Methods of Propagation

Reflection, diffraction and scattering are the three fundamental phenomena that cause signal propagation in a mobile communication system, apart from LoS com- munication. The most important parameter, predicted by propagation models based on above three phenomena, is the received power. The physics of the above phe- nomena may also be used to describe small scale fading and multipath propagation. The following subsections give an outline of these phenomena.

Reflection

Reflection occurs when an electromagnetic wave falls on an object, which has very large dimensions as compared to the wavelength of the propagating wave. For ex- ample, such objects can be the earth, buildings and walls. When a radio wave falls on another medium having different electrical properties, a part of it is transmitted into it, while some energy is reflected back. Let us see some special cases. If the medium on which the e.m. wave is incident is a dielectric, some energy is reflected back and some energy is transmitted. If the medium is a perfect conductor, all energy is reflected back to the first medium. The amount of energy that is reflected back depends on the polarization of the e.m. wave.

Another particular case of interest arises in parallel polarization, when no re- flection occurs in the medium of origin. This would occur, when the incident angle would be such that the reflection coefficient is equal to zero. This angle is the Brewster's angle. By applying laws of electro-magnetics, it is found to be

$$sin(\theta_B) = \sqrt{\frac{s_1}{s_1 + s_2}} \quad (3.6)$$

Further, considering perfect conductors, the electric field inside the conductor is always zero. Hence all energy is reflected back. Boundary conditions require that

$$\theta_i = \theta_r \quad (3.7)$$

and

$$E_i = E_r \quad (3.8)$$

for vertical polarization, and

$$E_i = -E_r \tag{3.9}$$

for horizontal polarization.

Diffraction

Diffraction is the phenomenon due to which an EM wave can propagate beyond the horizon, around the curved earth's surface and obstructions like tall buildings. As the user moves deeper into the shadowed region, the received field strength decreases. But the diffraction field still exists an it has enough strength to yield a good signal. This phenomenon can be explained by the Huygen's principle, according to which, every point on a wavefront acts as point sources for the production of sec- ondary wavelets, and they combine to produce a new wavefront in the direction of propagation. The propagation of secondary wavelets in the shadowed region results in diffraction. The field in the shadowed region is the vector sum of the electric field components of all the secondary wavelets that are received by the receiver.

Scattering

The actual received power at the receiver is somewhat stronger than claimed by the models of reflection and diffraction. The cause is that the trees, buildings and lamp- posts scatter energy in all directions. This provides extra energy at the receiver. Roughness is tested by a Rayleigh criterion, which defines a critical height h_c of surface protuberances for a given angle of incidence θ_i, given by,

$$h_c = \frac{\lambda}{8\ \ _i}.$$

A surface is smooth if its minimum to maximum protuberance h is less than h_c, and rough if protuberance is greater than h_c. In case of rough surfaces, the surface reflection coefficient needs to be multiplied by a scattering loss factor ρ_S, given by

$$\rho_S = exp(-8(\frac{\pi\sigma_h sin\theta_i}{\lambda})^2) \tag{3.11}$$

where σ_h is the standard deviation of the Gaussian random variable h. The following result is a better approximation to the observed value

$$\rho_S = exp(-8(\frac{\pi\sigma_h sin\theta_i}{\lambda})^2)I_0[-8(\frac{\pi\sigma_h sin\theta_i}{\lambda})^2] \tag{3.12}$$

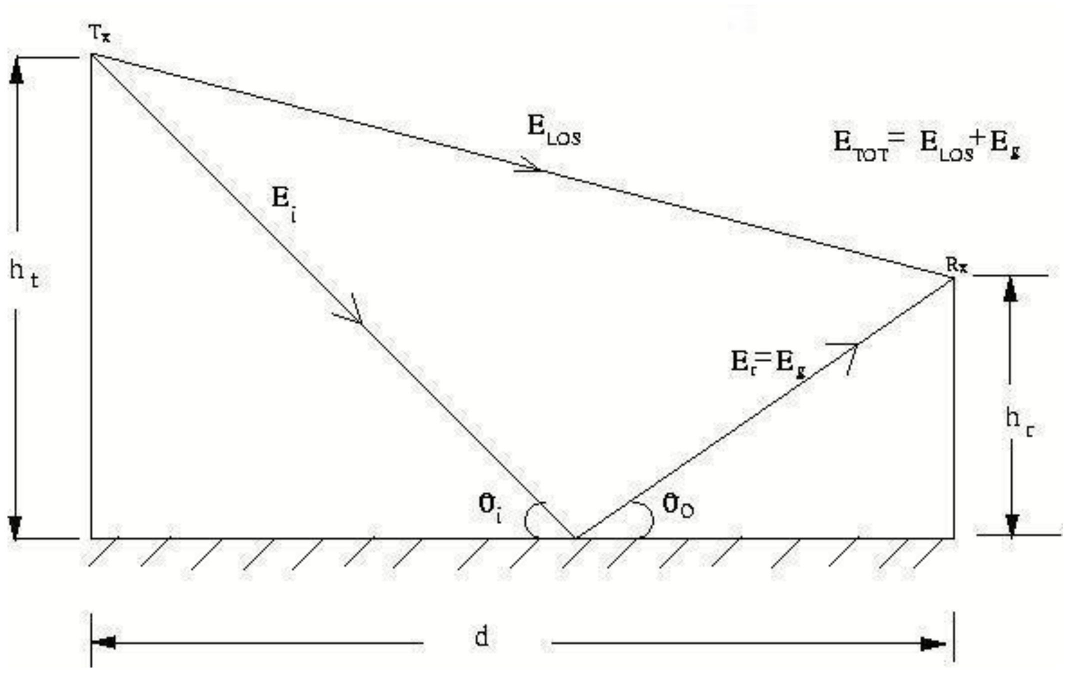


Figure 3.2: Two-ray reflection model.

which agrees very well for large walls made of limestone. The equivalent reflection coefficient is given by,

$$\Gamma_{rough} = \rho s \Gamma. \tag{3.13}$$

Two Ray Reflection Model

Interaction of EM waves with materials having different electrical properties than the material through which the wave is traveling leads to transmitting of energy through the medium and reflection of energy back in the medium of propagation. The amount of energy reflected to the amount of energy incidented is represented by Fresnel reflection coefficient Γ, which depends upon the wave polarization, angle of incidence and frequency of the wave. For example, as the EM waves can not pass through conductors, all the energy is reflected back with angle of incidence equal to the angle of reflection and reflection coefficient $\Gamma = -1$. In general, for parallel and perpendicular polarizations, Γ is given by:

$$\Gamma_{\|} = E_r/E_i = \eta_2 \sin\theta_t - \eta_1 \sin\theta_i/\eta_2 \sin\theta_t + \eta_1 \sin\theta_i \tag{3.14}$$

$$\Gamma_{\perp} = E_r/E_i = \eta_2 \sin\theta_i - \eta_1 \sin\theta_t / \eta_2 \sin\theta_i + \eta_1 \sin\theta_t. \tag{3.15}$$

Seldom in communication systems we encounter channels with only LOS paths and hence the Friis formula is not a very accurate description of the communication link. A two-ray model, which consists of two overlapping waves at the receiver, one direct path and one reflected wave from the ground gives a more accurate description as shown in Figure 4.2. A simple addition of a single reflected wave shows that power varies inversely with the forth power of the distance between the Tx and the Rx. This is deduced via the following treatment. From Figure 4.2, the total transmitted and received electric fields are

$$E_T^{TOT} = E_i + E_{LOS}, \tag{3.16}$$

$$E_R^{TOT} = E_g + E_{LOS}. \tag{3.17}$$

Let E_0 is the free space electric field (in V/m) at a reference distance d_0. Then

$$E(d,t) = \frac{E_0 d_0}{d} cos(\omega_c t - \varphi) \tag{3.18}$$

where

$$\varphi = \omega_c \frac{d}{c} \tag{3.19}$$

and $d > d_0$. The envelop of the electric field at d meters from the transmitter at any time t is therefore

$$|E(d, t)| = \frac{E_0 d_0}{d}. \tag{3.20}$$

This means the envelop is constant with respect to time.

Two propagating waves arrive at the receiver, one LOS wave which travels a distance of d^j and another ground reflected wave, that travels d^{jj}. Mathematically, it can be expressed as:

$$E(d^j,t) = \frac{E_0 d_0}{d^j} cos(\omega_c t - \varphi^j) \tag{3.21}$$

where

$$\varphi^j = \omega_c \frac{d^j}{c} \tag{3.22}$$

and

$$E(d^{jj},t) = \frac{E_0 d_0}{d^{jj}} cos(\omega_c t - \varphi^{jj}) \tag{3.23}$$

where

$$\varphi^{jj} = \omega_c \frac{d^{jj}}{c}. \tag{3.24}$$

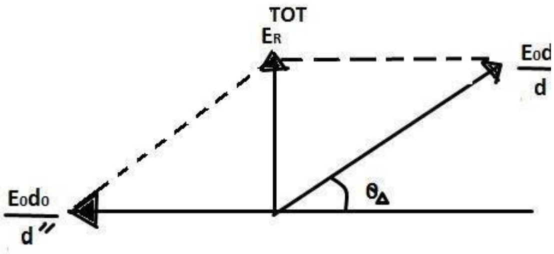


Figure 3.3: Phasor diagram of electric fields.

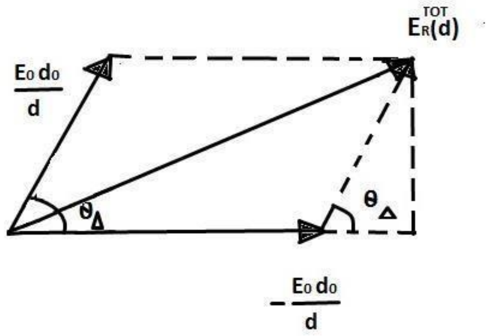


Figure 3.4: Equivalent phasor diagram of Figure 4.3.

According to the law of reflection in a dielectric, $\theta_i = \theta_0$ and $E_g = \Gamma E_i$ which means the total electric field,

$$E_t = E_i + E_g = E_i(1 + \Gamma). \tag{3.25}$$

For small values of θ_i, reflected wave is equal in magnitude and 180^o out of phase with respect to incident wave. Assuming perfect horizontal electric field polarization, i.e.,

$$\Gamma_\perp = -1 \Longrightarrow E_t = (1-1)E_i = 0, \tag{3.26}$$

the resultant electric field is the vector sum of E_{LOS} and E_g. This implies that,

$$E_R^{TOT} = |E_{LOS} + E_g|. \tag{3.27}$$

It can be therefore written that

$$E_R^{TOT}(d,t) = \frac{E_0 d_0}{d'}\cos\left(\omega_c\left(t - \frac{d'}{c}\right)\right) + (-1)\frac{E_0 d_0}{d''}\cos\left(\omega_c\left(t - \frac{d''}{c}\right)\right) \tag{3.28}$$

In such cases, the path difference is

$$\Delta = d'' - d' = \sqrt{(h_t + h_r)^2 + d^2} - \sqrt{(h_t - h_r)^2 + d^2}. \tag{3.29}$$

However, when T-R separation distance is very large compared to $(h_t + h_r)$, then

$$\Delta \approx \frac{2h_t h_r}{d} \tag{3.30}$$

Ex 3: Prove the above two equations, i.e., equation (4.29) and (4.30).

Once the path difference is known, the phase difference is

$$\theta_\Delta = \frac{2\pi\Delta}{\lambda} = \frac{\Delta\omega_c}{\lambda} \tag{3.31}$$

and the time difference,

$$\tau_d = \frac{\Delta}{c} = \frac{\theta_\Delta}{2\pi f_c}. \tag{3.32}$$

When d is very large, then Δ becomes very small and therefore E_{LOS} and E_g are virtually identical with only phase difference,i.e.,

$$\left|\frac{E_0 d_0}{d}\right| \approx \left|\frac{E_0 d_0}{d'}\right| \approx \left|\frac{E_0 d_0}{d''}\right|. \tag{3.33}$$

Say, we want to evaluate the received E-field at any $t = \frac{d''}{c}$. Then,

$$E_R^{TOT}\left(d, t = \frac{d''}{c}\right) = \frac{E_0 d_0}{d'} cos\left(\omega_c \frac{d''}{c} - \omega_c \frac{d'}{c}\right) - \frac{E_0 d_0}{d''} cos\left(\omega_c \frac{d''}{c} - \omega_c \frac{d''}{c}\right) \tag{3.34}$$

$$= \frac{E_0 d_0}{d'} cos\left(\frac{\Delta \omega_c}{c}\right) - \frac{E_0 d_0}{d''} cos(0^o) \tag{3.35}$$

$$= \frac{E_0 d_0}{d'} \angle \theta_\Delta - \frac{E_0 d_0}{d''} \tag{3.36}$$

$$\approx \frac{E_0 d_0}{d} (\angle \theta_\Delta - 1). \tag{3.37}$$

Using phasor diagram concept for vector addition as shown in Figures 3.3 and 3.4, we get

$$|E_R^{TOT}(d)| = \sqrt{\left(\frac{E_0 d_0}{d} + \frac{E_0 d_0}{d} cos(\theta_\Delta)\right)^2 + \left(\frac{E_0 d_0}{d} sin(\theta_\Delta)\right)^2} \tag{3.38}$$

$$= \frac{E_0 d_0}{d} \sqrt{(cos(\theta_\Delta) - 1)^2 + sin^2(\theta_\Delta)} \tag{3.39}$$

$$= \frac{E_0 d_0}{d} \sqrt{2 - 2cos\theta_\Delta} \tag{3.40}$$

$$= 2 \frac{E_0 d_0}{d} sin\left(\frac{\theta_\Delta}{2}\right). \tag{3.41}$$

For $\frac{\theta_\Delta}{2} < 0.5rad$, $sin(\frac{\theta_\Delta}{2}) \approx \frac{\theta_\Delta}{2}$. Using equation (4.31) and further equation (4.30), we can then approximate that

$$sin\left(\frac{\theta_\Delta}{2}\right) \approx \frac{\pi}{\lambda} \Delta = \frac{2\pi h_t h_r}{\lambda d} < 0.5rad. \tag{3.42}$$

This raises the wonderful concept of ‘cross-over distance’ d_c, defined as

$$d > d_c = \frac{20\pi h_t h_r}{5\lambda} = \frac{4\pi h_t h_r}{\lambda}. \tag{4.43}$$

The corresponding approximate received electric field is

$$E_R^{TOT}(d) \approx 2 \frac{E_0 d_0}{d} \frac{2\pi h_t h_r}{\lambda d} = k \frac{h_t h_r}{d^2}. \tag{4.44}$$

Therefore, using equation (4.43) in (4.1), we get the received power as

$$P_r = \frac{P_t tt_t tt_r h_t^2 h_r^2}{L d^4}. \tag{4.45}$$

The cross-over distance shows an approximation of the distance after which the received power decays with its fourth order. The basic difference between equation (4.1) and (4.45) is that when $d < d_c$, equation (4.1) is sufficient to calculate the path loss since the two-ray model does not give a good result for a short distance due to the oscillation caused by the constructive and destructive combination of the two rays, but whenever we distance crosses the 'cross-over distance', the power falls off rapidly as well as two-ray model approximation gives better result than Friis equation.

Observations on Equation (4.45): The important observations from this equation are:

1. This equation gives fair results when the T-R separation distance crosses the cross-over distance.

1. In that case, the power decays as the fourth power of distance

$$P_r(d) = \frac{K}{d^4}, \tag{4.46}$$

with K being a constant.

2. Path loss is independent of frequency (wavelength).
3. Received power is also proportional to h_t^2 and h_r^2, meaning, if height of any of the antennas is increased, received power increases.

Diffraction

Diffraction is the phenomena that explains the digression of a wave from a straight line path, under the influence of an obstacle, so as to propagate behind the obstacle. It is an inherent feature of a wave be it longitudinal or transverse. For e.g the sound can be heard in a room, where the source of the sound is another room without having any line of sight. The similar phenomena occurs for light also but the diffracted light intensity is not noticeable. This is because the obstacle or slit need to be of the order of the wavelength of the wave to have a significant effect. Thus radiation from a point source radiating in all directions can be received at any

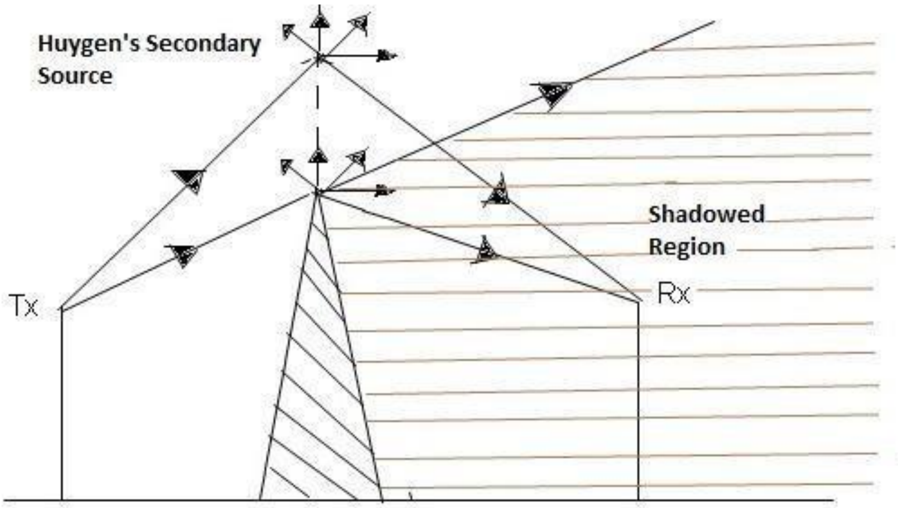


Figure 3.5: Huygen's secondary wavelets.

point, even behind an obstacle (unless it is not completely enveloped by it), as shown in Figure 4.5. Though the intensity received gets smaller as receiver is moved into the shadowed region. Diffraction is explained by Huygens-Fresnel principle which states that all points on a wavefront can be considered as the point source for secondary wavelets which form the secondary wavefront in the direction of the prorogation. Normally, in absence of an obstacle, the sum of all wave sources is zero at a point not in the direct path of the wave and thus the wave travels in the straight line. But in the case of an obstacle, the effect of wave source behind the obstacle cannot be felt and the sources around the obstacle contribute to the secondary wavelets in the shadowed region, leading to bending of wave. In mobile communication, this has a great advantage since, by diffraction (and scattering, reflection), the receiver is able to receive the signal even when not in line of sight of the transmitter. This we show in the subsection given below.

Knife-Edge Diffraction Geometry

As shown in Figure 4.6, consider that there's an impenetrable obstruction of hight h at a distance of d_1 from the transmitter and d_2 from the receiver. The path difference between direct path and the diffracted path is

$$\delta = \sqrt{d_1^2 + h^2} + \sqrt{d_2^2 + h^2} - (d_1 + d_2) \qquad (4.47)$$

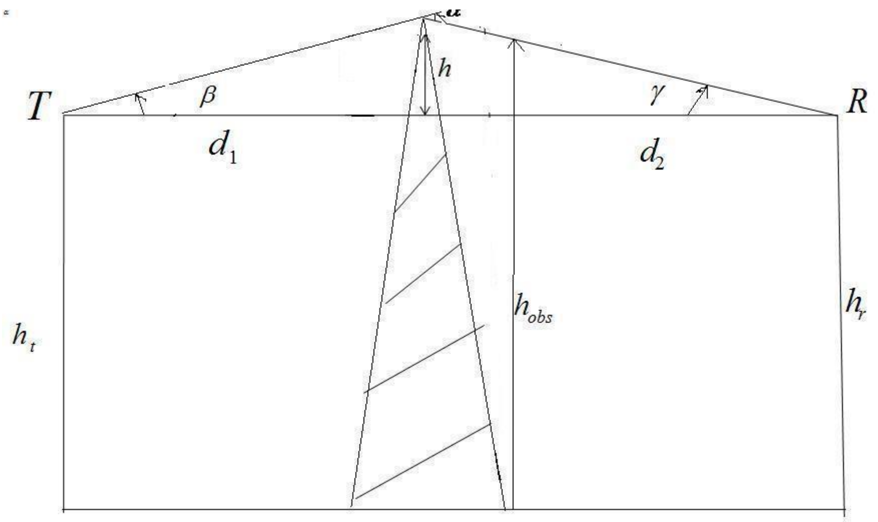


Figure 3.6: Diffraction through a sharp edge. which can

be further simplified as

$$\begin{aligned}\delta &= d_1(1 + h^2/2d^2) + d_2(1 + h^2/2d^2) - (d_1 + d_2) \\ &= h^2/(2d_1) + h^2/(2d_2) = h^2(d_1 + d_2)/(2d_1d_2). \end{aligned} \tag{4.48}$$

Thus the phase difference equals

$$\varphi = 2\pi\delta/\lambda = 2\pi h^2(d_1 + d_2)/\lambda 2(d_1d_2). \tag{4.49}$$

With the following considerations that

$$\alpha = \beta + \gamma \tag{4.50}$$

and

$$\alpha \approx \tan\alpha \tag{4.51}$$

we can write,

$$\alpha\tan\alpha = \tan\beta + \tan\gamma = h/d_1 + h/d_2 = h(d_1 + d_2)/d_1d_2. \tag{4.52}$$

In order to normalize this, we usually use a Fresnel-Kirchoff diffraction parameter v, expressed as

$$v = h\sqrt{2(d_1 + d_2)/(\lambda d_1d_2)} = \alpha\sqrt{(2d_1d_2)/(\lambda(d_1 + d_2))} \tag{4.53}$$

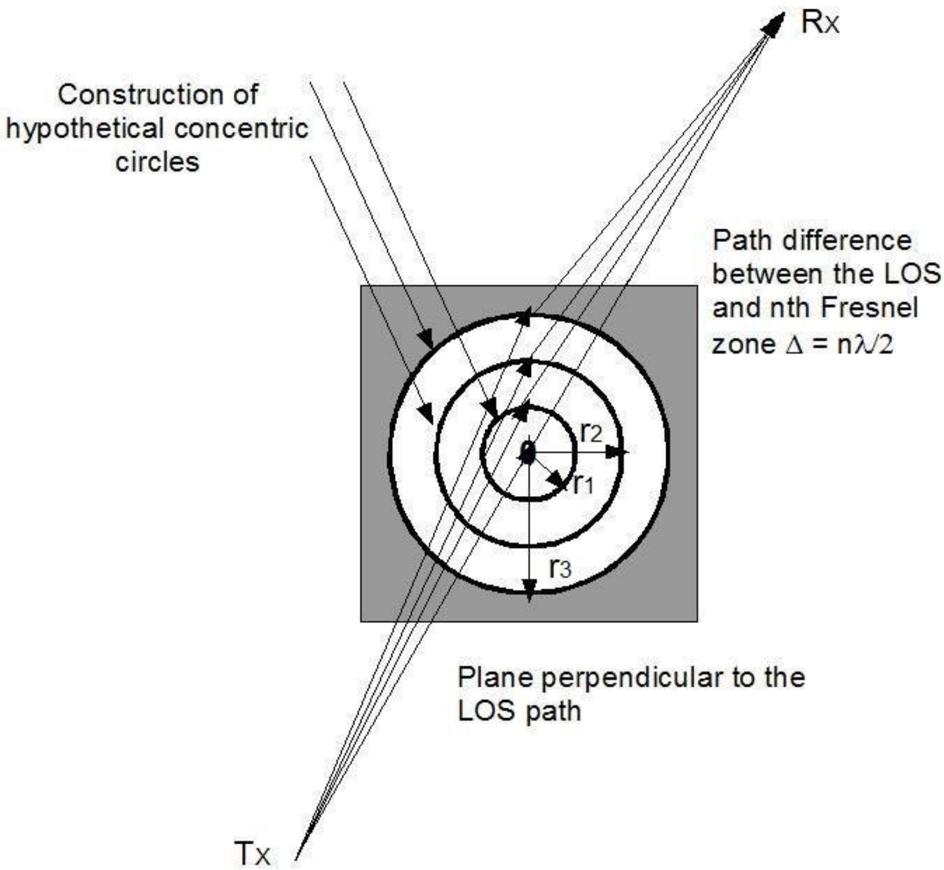


Figure 4.7: Fresnel zones.

and therefore the phase difference becomes

$$\varphi = \pi v^2/2. \tag{4.54}$$

From this, we can observe that: (i) phase difference is a function of the height of the obstruction, and also, (ii) phase difference is a function of the position of the obstruction from transmitter and receiver.

Fresnel Zones: the Concept of Diffraction Loss

As mentioned before, the more is the object in the shadowed region greater is the diffraction loss of the signal. The effect of diffraction loss is explained by Fresnel zones as a function of the path difference. The successive Fresnel zones are limited by the circular periphery through which the path difference of the secondary waves is $n\lambda/2$ greater than total length of the LOS path, as shown in Figure 4.7. Thus successive Fresnel zones have phase difference of π which means they alternatively provide constructive and destructive interference to the received the signal. The radius of the each Fresnel zone is maximum at middle of transmitter and receiver (i.e. when $d_1 = d_2$) and decreases as moved to either side. It is seen that the loci of a Fresnel zone varied over d1 and d2 forms an ellipsoid with the transmitter and receiver at its focii. Now, if there's no obstruction, then all Fresnel zones result in only the direct LOS prorogation and no diffraction effects are observed. But if an obstruction is present, depending on its geometry, it obstructs contribution from some of the secondary wavelets, resulting in diffraction and also the loss of energy, which is the vector sum of energy from unobstructed sources. please note that height of the obstruction can be positive zero and negative also. The diffraction losses are minimum as long as obstruction doesn't block volume of the 1st Fresnel zone. As a rule of thumb, diffraction effects are negligible beyond 55% of 1st Fresnel zone.

Ex 4: Calculate the first Fresnel zone obstruction height maximum for f = 800 MHz.

Solution:

$$\lambda = \frac{c}{f} = \frac{3 \times 10^8}{8 \times 10^2 \times 10^6} = \frac{3}{8}\, m$$

$$\bar{H} = \sqrt{\frac{\lambda\, d_1 d_2}{(d_1 + d_2)}}$$

$$H_1 = \sqrt{\frac{3}{8} \times \frac{250 \times 250}{500}} = 6.89\ m$$

Thus $H_1 = 10 + 6.89 = 16.89m$

(b)

$$H_2 = \sqrt{\frac{\frac{3}{8} \times 100 \times 400}{500}} = 10\sqrt{(0.3)} = 5.48\ m$$

Thus

$$H_2 = 10 + 5.6 = 15.48m$$

. To have good power strength, obstacle should be within the 60% of the first fresnel zone.

Ex 5: Given f=900 MHz, $d_1 = d_2 = 1$ km, h = 25m, where symbols have usual meaning. Compute the diffraction loss. Also find out in which Fresnel zone the tip of the obstruction lies.

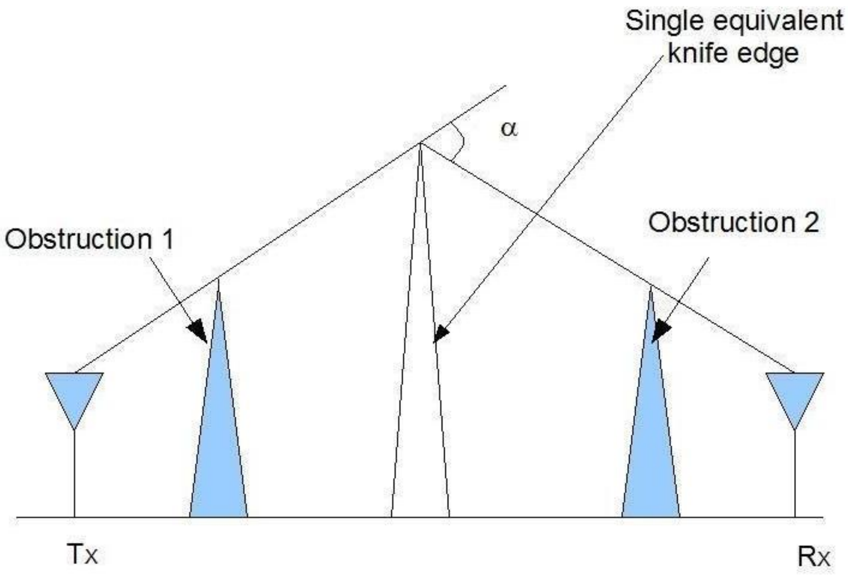


Figure 3.8: Knife-edge Diffraction Model

Given,

$$tt_d(dB) = 20\log(0.5 - 0.62v) \qquad -1 < v <= 0$$

$$tt_d(dB) = 20\log(0.225/v) \qquad v > 2.24$$

Solution:

$$v = h\sqrt{\frac{2(d_1 + d_2)}{\lambda d_1 d_2}} = 25\sqrt{\frac{2 \times 2000}{\frac{1}{3}10}} = 2.74$$

$$G_d(dB) = 20\log(\frac{225}{v}) = -21.7dB$$

Since loss = $-tt_d$ (dB) = 21.7 dB

$$n = \frac{(2.74)^2}{2} = 3\,5\cdot$$

Thus n=4.

Knife-edge diffraction model

Knife-edge diffraction model is one of the simplest diffraction model to estimate the diffraction loss. It considers the object like hill or mountain as a knife edge sharp object. The electric field strength, E_d of a knife-edge diffracted wave is given by

$$E_d/E_o = F(v) = (1+j)/2 \int_v^{\infty} (exp((-j\pi t^2)/2)dt. \quad (3.55)$$

The diffraction gain due to presence of knife edge can be given as

$$tt_d(db) = 20\log|F(v)| \quad (3.56)$$

$$tt_d(db) = 0 v <= -1 \quad (3.57)$$

$$tt_d(db) = 20\log(0.5 - 0.62) \qquad -1 <= v <= 0 \quad (3.58)$$

$$tt_d(db) = 20\log(0.5 exp(-0.95v)) \qquad 0 <= v <= 1 \quad (3.59)$$

$$tt_d(db) = 20\log(0.4 - sqrt(0.1184 - (0.38 - 0.1v^2))) \qquad 1 <= v <= 2.4 \quad (3.60)$$

$$tt_d(db) = 20\log(0.225/v) \qquad v > 2.4 \quad (3.61)$$

When there are more than one obstruction, then the equivalent model can be found by one knife-edge diffraction model as shown in Figure 4.8.

Link Budget Analysis

Log-distance Path Loss Model

According to this model the received power at distance d is given by,

$$PL(d)\left(\frac{d}{d_0}\right)^n \Longrightarrow PL(dB) = PL(d_0) + 10n\log\left(\frac{d}{d_0}\right) \quad (3.62)$$

The value of n varies with propagation environments. The value of n is 2 for free space. The value of n varies from 4 to 6 for obstruction of building, and 3 to 5 for urban scenarios. The important factor is to select the correct reference distance d_0. For large cell area it is 1 Km, while for micro-cell system it varies from 10m-1m.

Limitations:

Surrounding environmental clutter may be different for two locations having the same transmitter to receiver separation. Moreover it does not account for the shadowing effects.

Log Normal Shadowing

The equation for the log normal shadowing is given by,

$$P\,L(dB) = \overline{P\,L}(dB) + X_\sigma = P\,L(\overline{d_0}) + 10n\log\left(\frac{d}{d_0}\right) + X_\sigma \tag{3.63}$$

where X_σ is a zero mean Gaussian distributed random variable in dB with standard deviation σ also in dB. In practice n and σ values are computed from measured data.

Average received power

The 'Q' function is given by,

$$Q(z) = 0.5\left(1 - erf\left(\frac{z}{\sqrt{2}}\right)\right) \tag{3.64}$$

and

$$Q(z) = 1 - Q(-z) \tag{3.65}$$

So the probability that the received signal level (in dB) will exceed a certain value γ is

$$P\,(P_d > \gamma) = Q\left(\frac{\gamma - \overline{P_r}}{\sigma}\right). \tag{3.66}$$

Outdoor Propagation Models

There are many empirical outdoor propagation models such as Longley-Rice model, Durkin's model, Okumura model, Hata model etc. Longley-Rice model is the most commonly used model within a frequency band of 40 MHz to 100 GHz over different terrains. Certain modifications over the rudimentary model like an extra urban factor (UF) due to urban clutter near the reciever is also included in this model. Below, we discuss some of the outdoor models, followed by a few indoor models too.

Okumura Model

The Okumura model is used for Urban Areas is a Radio propagation model that is used for signal prediction.The frequency coverage of this model is in the range of 200 MHz to 1900 MHz and distances of 1 Km to 100 Km.It can be applicable for base station effective antenna heights (h_t) ranging from 30 m to 1000 m.

Okumura used extensive measurements of base station-to-mobile signal attenua- tion throughout Tokyo to develop a set of curves giving median attenuation relative to free space (A_{mu}) of signal propogation in irregular terrain. The empirical path- loss formula of Okumura at distance d parameterized by the carrier frequency f_c is given by

$$P_L(d)dB = L(f_c, d) + A_{mu}(f_c, d) - tt(h_t) - tt(h_r) - tt_{AREA} \tag{3.67}$$

where $L(f_c, d)$ is free space path loss at distance d and carrier frequency f_c, $A_{mu}(f_c, d)$ is the medianattenuation inadditiontofree-space pathlossacrossallenvironments,$tt(h_t)$ is the base station antenna height gain factor,$tt(h_r)$ is the mobile antenna height gain factor,tt_{AREA} is the gain due to type of environment. The values of $A_{mu}(f_c, d)$ and tt_{AREA} are obtained from Okumura's empirical plots. Okumura derived empirical formulas for $tt(h_t)$ and $tt(h_r)$ as follows:

$$tt(h_t) = 20 \log_{10}(h_t/200), \qquad 30\,m < h_t < 1000m \tag{3.68}$$

$$tt(h_r) = 10 \log_{10}(h_r/3), \qquad h_r \leq 3m \tag{3.69}$$

$$tt(h_r) = 20 \log_{10}(h_r/3), \qquad 3\,m < h_r < 10m \tag{3.70}$$

Correlation factors related to terrain are also developed in order to improve the models accuracy. Okumura's model has a 10-14 dB empirical standard deviation between the path loss predicted by the model and the path loss associated with one of the measurements used to develop the model.

Hata Model

The Hata model is an empirical formulation of the graphical path-loss data provided by the Okumura and is valid over roughly the same range of frequencies, 150-1500 MHz. This empirical formula simplifies the calculation of path loss because it is closed form formula and it is not based on empirical curves for the different param- eters. The standard formula for empirical path loss in urban areas under the Hata model is

$$P_{L,urban}(d)dB = 69.55+26.16 \log_{10}(f_c)-13.82 \log_{10}(h_t)-a(h_r)+(44.9-6.55 \log_{10}(h_t)) \log_{10}(d) \tag{3.71}$$

The parameters in this model are same as in the Okumura model,and $a(h_r)$ is a correction factor for the mobile antenna height based on the size of coverage area.For small to medium sized cities this factor is given by

$$a(h_r) = (1.11 \log_{10}(f_c) - 0.7)h_r - (1.56 \log_{10}(f_c) - 0.8)dB$$ and for

larger cities at a frequencies $f_c > 300$ MHz by

$$a(h_r) = 3.2(\log_{10}(11.75h_r))^2 - 4.97dB$$

else it is

$$a(h_r) = 8.29(\log_{10}(1.54h_r))^2 - 1.1dB$$

Corrections to the urban model are made for the suburban, and is given by

$$P_{L,suburban}(d)dB = P_{L,urban}(d)dB - 2(\log_{10}(f_c/28))^2 - 5.4 \quad (3.72)$$

Unlike the Okumura model,the Hata model does not provide for any specific path-correlation factors. The Hata model well approximates the Okumura model for distances d > 1 Km. Hence it is a good model for first generation cellular systems, but it does not model propogation well in current cellular systems with smaller cell sizes and higher frequencies. Indoor environments are also not captured by the Hata model.

Indoor Propagation Models

The indoor radio channel differs from the traditional mobile radio channel in ways- the distances covered are much smaller ,and the variability of the environment is much greater for smaller range of Tx-Rx separation distances.Features such as lay-out of the building,the construction materials,and the building type strongly in- fluence the propagation within the building.Indoor radio propagation is dominated by the same mechanisms as outdoor: reflection, diffraction and scattering with vari- able conditions. In general,indoor channels may be classified as either line-of-sight or obstructed.

Partition Losses Inside a Floor (Intra-floor)

The internal and external structure of a building formed by partitions and obstacles vary widely.Partitions that are formed as a part of building structure are called hard partitions , and partitions that may be moved and which do not span to the ceiling are called soft partitions. Partitions vary widely in their physical and electrical characteristics,making it difficult to apply general models to specific indoor installations.

Partition Losses Between Floors (Inter-floor)

The losses between floors of a building are determined by the external dimensions and materials of the building,as well as the type of construction used to create the floors and the external surroundings. Even the number of windows in a building and the presence of tinting can impact the loss between floors.

Log-distance Path Loss Model

It has been observed that indoor path loss obeys the distance power law given by

$$PL(dB) = PL(d_0) + 10n \log_{10}(d/d_0) + X_\sigma \tag{3.73}$$

where n depends on the building and surrounding type, and X_σ represents a normal random variable in dB having standard deviation of σ dB.

Multipath Wave Propagation and Fading

Multipath Propagation

In wireless telecommunications, multipath is the propagation phenomenon that re- sults in radio signals reaching the receiving antenna by two or more paths. Causes of multipath include atmospheric ducting, ionospheric reflection and refraction, and reflection from water bodies and terrestrial objects such as mountains and buildings. The effects of multipath include constructive and destructive interference, and phase shifting of the signal. In digital radio communications (such as GSM) multipath can cause errors and affect the quality of communications. We discuss all the related issues in this chapter.

Multipath & Small-Scale Fading

Multipath signals are received in a terrestrial environment, i.e., where different forms of propagation are present and the signals arrive at the receiver from transmitter via a variety of paths. Therefore there would be multipath interference, causing multi- path fading. Adding the effect of movement of either Tx or Rx or the surrounding clutter to it, the received overall signal amplitude or phase changes over a small amount of time. Mainly this causes the fading.

Fading

The term **fading**, or, small-scale fading, means rapid fluctuations of the amplitudes, phases, or multipath delays of a radio signal over a short period or short travel distance. This might be so severe that large scale radio propagation loss effects might be ignored.

Multipath Fading Effects

In principle, the following are the main multipath effects:

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

Factors Influencing Fading

The following physical factors influence small-scale fading in the radio propagation channel:

(1) **Multipath propagation –** Multipath is the propagation phenomenon that re- sults in radio signals reaching the receiving antenna by two or more paths. The effects of multipath include constructive and destructive interference, and phase shifting of the signal.

(2) **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 multipath components.

(3) **Speed of surrounding objects –** If objects in the radio channel are in mo- tion, 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 fading.

(4) **Transmission Bandwidth of the signal –** If the transmitted radio signal bandwidth is greater than the "bandwidth" of the multipath channel (quanti- fied by *coherence bandwidth*), the received signal will be distorted.

Types of Small-Scale Fading

The type of fading experienced by the signal through a mobile channel depends on the relation between the signal parameters (bandwidth, symbol period) and the channel parameters (rms delay spread and Doppler spread). Hence we have four different types of fading. There are two types of fading due to the time dispersive nature of the channel.

Fading Effects due to Multipath Time Delay Spread

Flat Fading

Such types of fading occurs when the bandwidth of the transmitted signal is less than the coherence bandwidth of the channel. Equivalently if the symbol period of the signal is more than the rms delay spread of the channel, then the fading is flat fading.

So we can say that flat fading occurs when

$$B_S \quad B_C$$

where B_S is the signal bandwidth and B_C is the coherence bandwidth. Also

$$T_S \times \sigma_\tau$$

where T_S is the symbol period and σ_τ is the rms delay spread. And in such a case, mobile channel has a constant gain and linear phase response over its bandwidth.

Frequency Selective Fading

Frequency selective fading occurs when the signal bandwidth is more than the co-herence bandwidth of the mobile radio channel or equivalently the symbols duration of the signal is less than the rms delay spread.

$$B_S \times B_C$$

and

$$T_S \quad \sigma_\tau$$

At the receiver, we obtain multiple copies of the transmitted signal, all attenuated and delayed in time. The channel introduces inter symbol interference. A rule of thumb for a channel to have flat fading is if

$$\frac{\sigma_\tau}{T_S} \leq 0.1$$

Fading Effects due to Doppler Spread

Fast Fading

In a fast fading channel, the channel impulse response changes rapidly within the symbol duration of the signal. Due to Doppler spreading, signal undergoes frequency dispersion leading to distortion. Therefore a signal undergoes fast fading if

$$T_S \times T_C$$

where T_C is the coherence time and

$$B_S \times B_D$$

where B_D is the Doppler spread. Transmission involving very low data rates suffer from fast fading.

Slow Fading

In such a channel, the rate of the change of the channel impulse response is much less than the transmitted signal. We can consider a slow faded channel a channel in which channel is almost constant over atleast one symbol duration. Hence

$$T_S \quad T_C$$

and

$$B_S \times B_D$$

We observe that the velocity of the user plays an important role in deciding whether the signal experiences fast or slow fading.

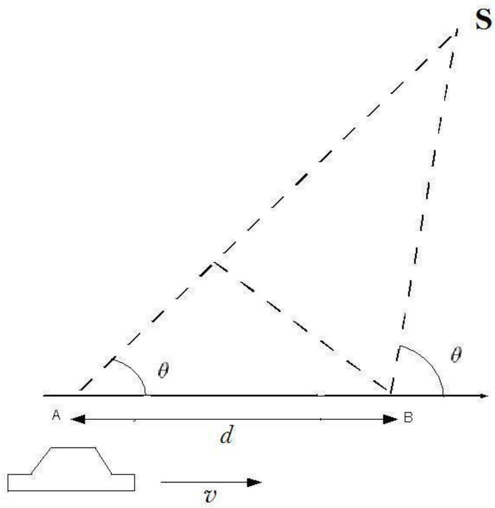


Figure Illustration of Doppler effect.

Doppler Shift

The Doppler effect (or Doppler shift) is the change in frequency of a wave for an observer moving relative to the source of the wave. In classical physics (waves in a medium), the relationship between the observed frequency f and the emitted frequency f_o is given by:

$$f = \left(\frac{v \pm v_r}{v \pm v_s}\right) f_0$$

where v is the velocity of waves in the medium, v_s is the velocity of the source relative to the medium and v_r is the velocity of the receiver relative to the medium.

In mobile communication, the above equation can be slightly changed according to our convenience since the source (BS) is fixed and located at a remote elevated level from ground. The expected Doppler shift of the EM wave then comes out to be $\pm \frac{v_r}{} f_o$ or $\lambda \pm \frac{v_r}{}$. As the BS is located at an elevated place, a cos φ factor would also be multiplied with this. The exact scenario, as given in Figure 5.1, is illustrated below.

Consider a mobile moving at a constant velocity v, along a path segment length d between points A and B, while it receives signals from a remote BS source S. The difference in path lengths traveled by the wave from source S to the mobile at points A and B is $\Delta l = d \cos\theta = v\Delta t \cos\theta$, where Δt is the time required for the mobile to travel from A to B, and θ is assumed to be the same at points A and B 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

$$\Delta\phi = \frac{2\pi\Delta l}{\lambda} = \frac{2\pi v\Delta t}{\lambda}\cos\theta \tag{5.11}$$

and hence the apparent change in frequency, or Doppler shift (f_d) is

$$f_d = \frac{1}{2\pi}\cdot\frac{\Delta\phi}{\Delta t} = \frac{v}{\lambda}\cdot\cos\theta.$$

Example 1

An aircraft is heading towards a control tower with 500 kmph, at an elevation of 20°. Communication between aircraft and control tower occurs at 900 MHz. Find out the expected Doppler shift.

Solution As given here,

$$v = 500\ kmph$$

the horizontal component of the velocity is

$$v^{j} = v\cos\theta = 500 \times \cos 20^\circ = 130\ m/s$$

Hence, it can be written that

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

$$f_d = \frac{130}{1/3} = 390\ Hz$$

If the plane banks suddenly and heads for other direction, the Doppler shift change will be 390 Hz to −390 Hz.

Impulse Response Model of a Multipath Channel

Mobile radio channel may be modeled as a linear filter with time varying impulse response in continuous time. To show this, consider time variation due to receiver motion and time varying impulse response $h(d, t)$ and $x(t)$, the transmitted signal.

The received signal $y(d, t)$ at any position d would be

$$y(d,t) = x(t) * h(d,t) = \int_{-\infty}^{\infty} x(\tau)\, h(d, t-\tau)\, d\tau$$

For a causal system: $h(d, t) = 0$, for $t < 0$ and for a stable system $\int_{-\infty}^{\infty} |h(d,t)|\, dt < \infty$

Applying causality condition in the above equation, $h(d, t-\tau) = 0$ for $t-\tau < 0$ $\Rightarrow \tau > t$, i.e., the integral limits are changed to

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

Since the receiver moves along the ground at a constant velocity v, the position of the receiver is $d = vt$, i.e.,

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

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

$$y(t) = \int_{-\infty}^{t} x(\tau)\, h(vt, t-\tau)\, d\tau = x(t) * h(vt, t) = x(t) * h(d, t)$$

It is useful to discretize the multipath delay axis τ of the impulse response into equal time delay segments called *excess delay bins*, each bin having a time delay width equal to $(\tau_{i+1} - \tau_i) = \Delta\tau$ and $\tau_i = i\Delta\tau$ for $i \in \{0, 1, 2, ..N-1\}$, where N represents the total number of possible equally-spaced multipath components, including the first arriving component. The useful frequency span of the model is $2/\Delta\tau$. The model may be used to analyze transmitted RF signals having bandwidth less than $2/\Delta\tau$.
If there are N multipaths, maximum excess delay is given by $N\Delta\tau$.

$$\{y(t) = x(t) * h(t, \tau_i) | i = 0, 1, ...N-1\}$$

Bandpass channel impulse response model is

$$x(t) \rightarrow h(t, \tau) = Re\{h_b(t, \tau)e^{j\omega_c t}\} \rightarrow y(t) = Re\{r(t)e^{j\omega_c t}\}$$

Baseband equivalent channel impulse response model is given by

$$c(t) \rightarrow \frac{1}{2}\bar{h}_b(t, \tau) \rightarrow r(t) = c(t) * \frac{1}{2}\bar{h}_b(t, \tau)$$

Average power is

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

The baseband impulse response of a multipath channel can be expressed as

$$h_b(t,\tau) = \sum_{i=0}^{N-1} a_i(t,\tau)\exp[j(2\pi f_c\tau_i(t) + \phi_i(t,\tau))]\delta(\tau - \tau_i(t))$$

where $a_i(t, \tau)$ and $\tau_i(t)$ are the real amplitudes and excess delays, respectively, of the ith multipath component at time t. The phase term $2\pi f_c\tau_i(t) + \phi_i(t, \tau)$ in the above equation represents the phase shift due to free space propagation of the ith multipath component, plus any additional phase shifts which are encountered in the channel.

If the channel impulse response is wide sense stationary over a small-scale time or distance interval, then

$$h_b(\tau) = \sum_{i=0}^{N-1} a_i \exp[j\theta_i]\delta(\tau - \tau_i)$$

For measuring $h_b(\tau)$, we use a probing pulse to approximate $\delta(t)$ i.e.,

$$p(t) \approx \delta(t - \tau)$$

Power delay profile is taken by spatial average of $|h_b(t, \tau)|^2$ over a local area. The received power delay profile in a local area is given by

$$p(\tau) \approx k\overline{|h_b(t;\tau)|^2}.$$

Relation Between Bandwidth and Received Power

In actual wireless communications, impulse response of a multipath channel is measured using channel sounding techniques. Let us consider two extreme channel sounding cases.

Consider a pulsed, transmitted RF signal

$$x(t) = Re\{p(t)e^{j2\pi f_c t}\}$$

where $p(t) = \sqrt{\frac{4\tau_{max}}{T_{bb}}}$ for $0 \le t \le T_{bb}$ and 0 elsewhere. The low pass channel output is

$$r(t) = \frac{1}{2}\sum_{i=0}^{N-1} a_i \exp[j\theta_i]p(t - \tau_i)$$

$$= \sum_{i=0}^{N-1} a_i \exp[j\theta_i].\sqrt{\frac{\tau_{max}}{T_{bb}}}\, rect(t - \frac{T_b}{2} - \tau_i.)$$

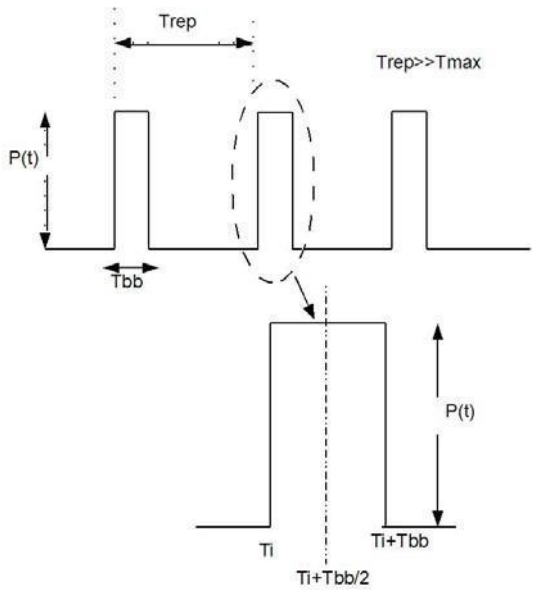


Figure: A generic transmitted pulsed RF signal.

The received power at any time t_0 is

$$|r(t_0)|^2 = \frac{1}{\tau_{\max}} \int_0^{\tau_{\max}} r(t)r^*(t)dt$$

$$= \frac{1}{\tau_{\max}} \int_0^{\tau_{\max}} \frac{1}{4}\left\{\sum_{k=0}^{N-1} a_k^2(t_0)p^2(t-\tau_k)\right\} dt$$

$$= \frac{1}{\tau_{\max}} \sum_{k=0}^{N-1} a_k^2(t_0) \int_0^{\tau_{\max}} \left\{\sqrt{\frac{\tau_{\max}}{T_{bb}}}\, rect\left(t - \frac{T_b}{2} - \tau_i\right)\right\}^2 dt$$

$$= \sum_{k=0}^{N-1} a_k^2(t_0).$$

Interpretation: If the transmitted signal is able to resolve the multipaths, then average small-scale receiver power is simply sum of average powers received from each multipath components.

$$E_{a,\theta}[P_{WB}] = E_{a,\theta}\left[\sum_{i=0}^{N-1} |a_i \exp(j\theta_i)|^2\right] \approx \sum_{i=0}^{N-1} \overline{a_i^2}$$

Now instead of a pulse, consider a CW signal, transmitted into the same channel and for simplicity, let the envelope be $c(t) = 2$. Then

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

and the instantaneous power is

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

Over local areas, a_i varies little but θ_i varies greatly resulting in large fluctuations.

$$\begin{aligned} E_{a,\theta}[P_{CW}] &= E_{a,\theta}\left[\left|\sum_{i=0}^{N-1} a_i \exp(j\theta_i)\right|^2\right] \\ &\approx \sum_{i=0}^{N-1} \overline{a_i^2} + 2\sum_{i=0}^{N-1}\sum_{i,j\neq i}^{N} r_{ij}\,\overline{\cos(\theta_i - \theta_j)} \end{aligned}$$

where $r_{ij} = E_a[a_i a_j]$.

If, $r_{ij} = \overline{\cos(\theta_i - \theta_j)} = 0$, then $E_{a,\theta}[P_{CW}] = E_{a,\theta}[P_{WB}]$. This occurs if multipath components are uncorrelated or if multipath phases are i.i.d over $[0, 2\pi]$.

Bottomline:

1. If the signal bandwidth is greater than multipath channel bandwidth then fading effects are negligible

2. If the signal bandwidth is less than the multipath channel bandwidth, large fading occurs due to phase shift of unresolved paths.

Linear Time Varying Channels (LTV)

The time variant transfer function(TF) of an LTV channel is FT of $h(t, \tau)$ w.r.t. τ .

$$H(f, t) = FT[h(\tau, t)] = \int_{-\infty}^{\infty} h(\tau, t)e^{-j2\pi f\tau}\, d\tau$$

$$h(\tau, t) = FT^{-1}[H(f, t)] = \int_{-\infty}^{\infty} H(f, t)e^{j2\pi f\tau}\, df$$

The received signal

$$r(t) = \int_{-\infty}^{\infty} R(f, t)e^{j2\pi ft}\, df$$

where $R(f, t) = H(f, t)X(f)$.

For flat fading channel, $h(\tau, t) = Z(t)\delta(\tau - \tau_i)$ where $Z(t) = \sum \alpha_n(t)e^{-j2\pi f_c\tau_n(t)}$. In this case, the received signal is

$$r(t) = \int_{-\infty}^{\infty} h(\tau, t)x(t-\tau)\, d\tau = Z(t)x(t-\tau_i)$$

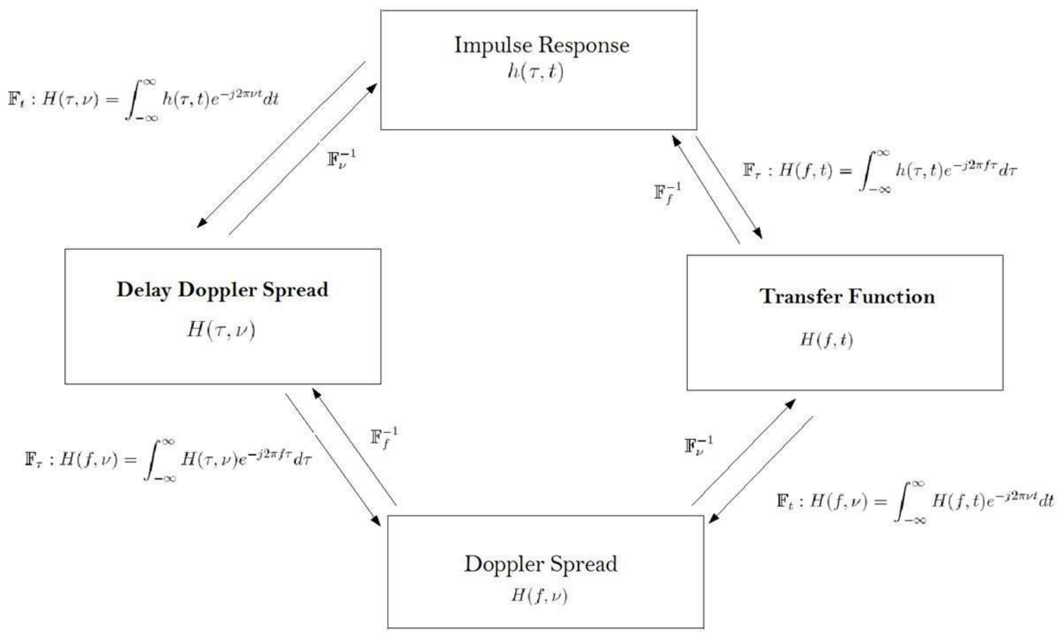


Figure : Relationship among different channel functions.

where the channel becomes multiplicative.

Doppler spread functions:

$$H(f, \nu) = FT\,[H(f, t)] = \int_{-\infty}^{\infty} H(f, t)e^{-j2\pi\nu t}dt$$

and

$$H(f, t) = FT^{-1}[H(f, \nu)] = \int_{-\infty}^{\infty} H(f, \nu)e^{j2\pi\nu t}d\nu$$

Delay Doppler spread:

$$H(\tau, \nu) = FT\,[h(\tau, t)] = \int_{-\infty}^{\infty} h(\tau, t)e^{-j2\pi\nu t}dt$$

Small-Scale Multipath Measurements

Direct RF Pulse System

A wideband pulsed bistatic radar usually transmits a repetitive pulse of width T_{bb} s, and uses a receiver with a wide bandpass filter ($BW = \frac{2}{T_{bb}}$ Hz). The signal is then amplified, envelope detected, and displayed and stored on a high speed oscilloscope. Immediate measurements of the square of the channel impulse response convolved with the probing pulse can be taken. If the oscilloscope is set on averaging mode, then this system provides a local average power delay profile.

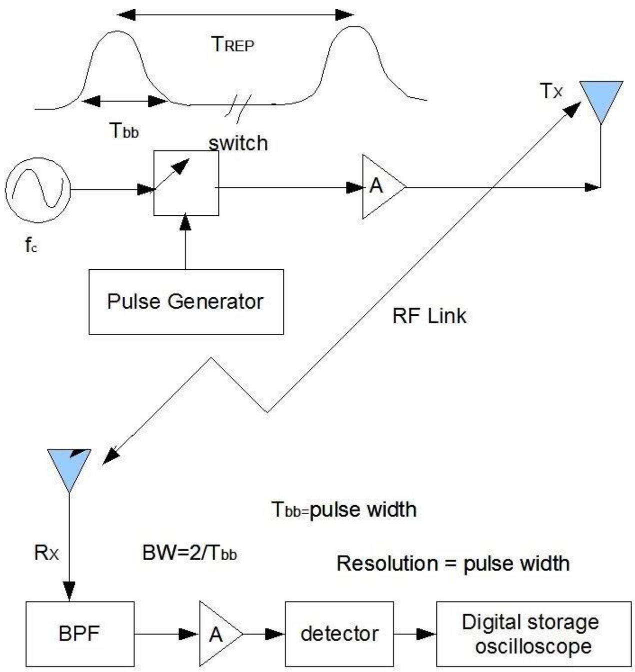


Figure : Direct RF pulsed channel IR measurement.

This system is subject to interference noise. If the first arriving signal is blocked or fades, severe fading occurs, and it is possible the system may not trigger properly.

Frequency Domain Channel Sounding

In this case we measure the channel in the frequency domain and then convert it into time domain impulse response by taking its inverse discrete Fourier transform (IDFT). A vector network analyzer controls a swept frequency synthesizer. An S- parameter test set is used to monitor the frequency response of the channel. The sweeper scans a particular frequency band, centered on the carrier, by stepping through discrete frequencies. The number and spacing of the frequency step impacts the time resolution of the impulse response measurement. For each frequency step, the S-parameter test set transmits a known signal level at port 1 and monitors the received signal at port 2. These signals allow the analyzer to measure the complex response, $S_{21}(\omega)$, of the channel over the measured frequency range. The $S_{21}(\omega)$ measure is the measure of the signal flow from transmitter antenna to receiver

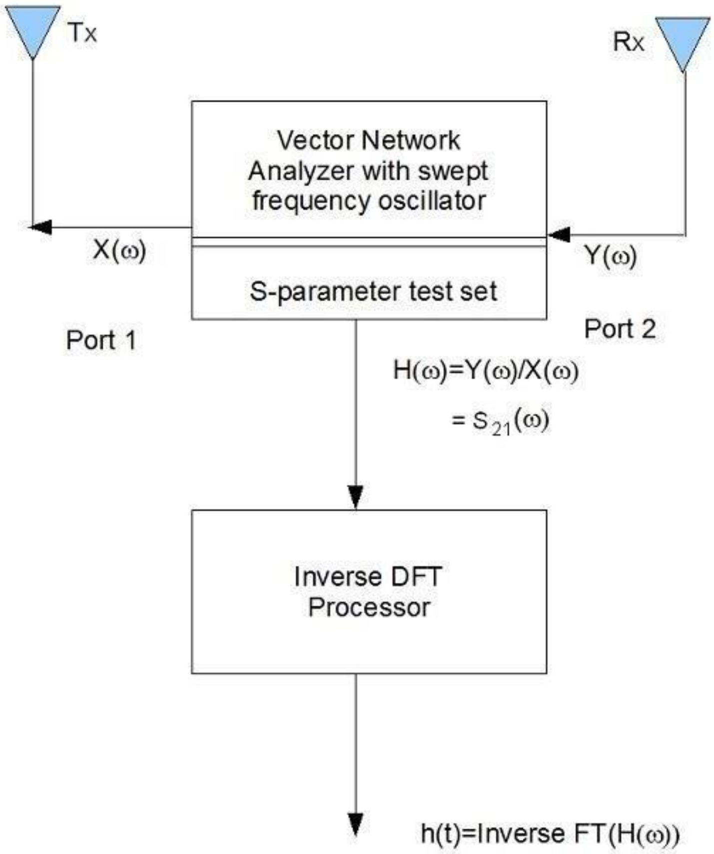


Figure : Frequency domain channel IR measurement.

antenna (i.e., the channel).

This system is suitable only for indoor channel measurements. This system is also non real-time. Hence, it is not suitable for time-varying channels unless the sweep times are fast enough.

Multipath Channel Parameters

To compare the different multipath channels and to quantify them, we define some parameters. They all can be determined from the power delay profile. These pa- rameters can be broadly divided in to two types.

Time Dispersion Parameters

These parameters include the mean excess delay,rms delay spread and excess delay spread. The mean excess delay is the first moment of the power delay profile and is defined as

$$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)}$$

where a_k is the amplitude, τ_k is the excess delay and $P(\tau_k)$ is the power of the individual multipath signals.

The mean square excess delay spread is defined as

$$\overline{\tau^2} = \frac{\sum_k P(\tau_k)\tau_k^2}{\sum_k P(\tau_k)}$$

Since the rms delay spread is the square root of the second central moment of the power delay profile, it can be written as

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

As a rule of thumb, for a channel to be flat fading the following condition must be satisfied

$$\frac{\sigma_\tau}{T_S} \leq 0.1$$

where T_S is the symbol duration. For this case, no equalizer is required at the receiver.

Example 2

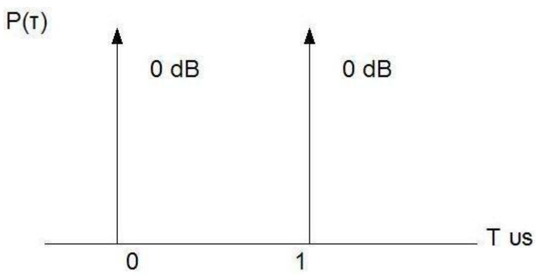
1. Sketch the power delay profile and compute RMS delay spread for the follow- ing: $P(\tau) = \sum_{n=0} \delta(\tau - n \times 10^{-6})$ (in watts)
2. If BPSK modulation is used, what is the maximum bit rate that can be sent through the channel without needing an equalizer?

Solution

1. $P(0) = 1$ watt, $P(1) = 1$ watt

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

$$\overline{\tau^2} = 0.5\mu s^2 \quad \sigma_\tau = 0.5\mu s$$

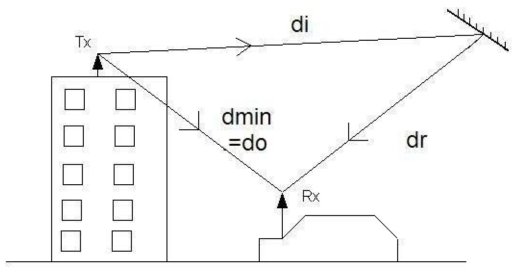


2. For flat fading channel, we need $\frac{\sigma_\tau}{T_s} 0.1 \quad \Rightarrow R_s = \frac{1}{T_s} = 0.2 \times 10^4 = 200\ kbps$

   For BPSK we need $R_b = R_s = 200\ kbps$

Example 3 A simple delay spread bound: **Feher's upper bound**

Consider a simple worst-case delay spread scenario as shown in figure below.



Here $d_{min} = d_0$ and $d_{max} = d_i + d_r$

Transmitted power = P_T, Minimum received power = $P_{R_{\min}} = P_{T\,hreshold}$

$$\frac{P_{R_{\min}}}{P_T} = tt_T\, tt_R \left(\frac{\lambda}{4\pi d_{\max}}\right)^2$$

Put $tt_T = tt_R = 1$ i.e., considering omni-directional unity gain antennas

$$d_{\max} = \left(\frac{\lambda}{4\pi}\right)\left(\frac{P_T}{P_{R_{\min}}}\right)^{\frac{1}{2}}$$

$$\tau_{\max} = \frac{d_{\max}}{c} = \left(\frac{\lambda}{4\pi c}\right)\left(\frac{P_T}{P_{R_{\min}}}\right)^{\frac{1}{2}}$$

$$\tau_{\max} = \left(\frac{1}{4\pi f}\right)\left(\frac{P_T}{P_{R_{\min}}}\right)^{\frac{1}{2}}$$

Frequency Dispersion Parameters

To characterize the channel in the frequency domain, we have the following param- eters.

(1) Coherence bandwidth: it is a statistical measure of the range of frequencies over which the channel can be considered to pass all the frequency components with almost equal gain and linear phase. When this condition is satisfied then we say the channel to be flat.

Practically, coherence bandwidth is the minimum separation over which the two frequency components are affected differently. If the coherence bandwidth is con- sidered to be the bandwidth over which the frequency correlation function is above 0.9, then it is approximated as

$$B_C \approx \frac{1}{50\sigma_\tau}.$$

However, if the coherence bandwidth is considered to be the bandwidth over which the frequency correlation function is above 0.5, then it is defined as

$$B_C \approx \frac{1}{5\sigma_\tau}.$$

The coherence bandwidth describes the time dispersive nature of the channel in the local area. A more convenient parameter to study the time variation of the channel is the coherence time. This variation may be due to the relative motion between the mobile and the base station or the motion of the objects in the channel.

(2) Coherence time: this is a statistical measure of the time duration over which the channel impulse response is almost invariant. When channel behaves like this, it is said to be slow faded. Essentially it is the minimum time duration over which two received signals are affected differently. For an example, if the coherence time is considered to be the bandwidth over which the time correlation is above 0.5, then it can be approximated as

$$T_C \approx \frac{9}{16\pi f_m}$$

where f_m is the maximum doppler spread given be $f_m = \frac{v}{\lambda}$.

Another parameter is the Doppler spread (B_D) which is the range of frequencies over which the received Doppler spectrum is non zero.

Statistical models for multipath propagation

Many multipath models have been proposed to explain the observed statistical na- ture of a practical mobile channel. Both the first order and second order statistics

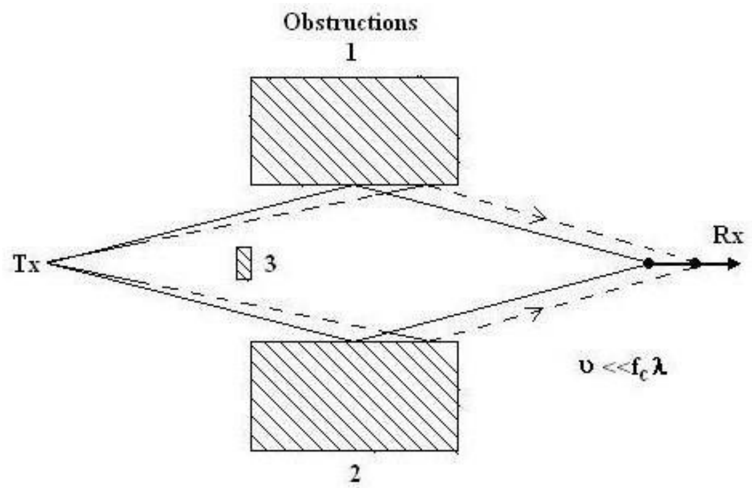


Figure: Two ray NLoS multipath, resulting in Rayleigh fading.

have been examined in order to find out the effective way to model and combat the channel effects. The most popular of these models are Rayleigh model, which de- scribes the NLoS propagation. The Rayleigh model is used to model the statistical time varying nature of the received envelope of a flat fading envelope. Below, we discuss about the main first order and second order statistical models.

NLoS Propagation: Rayleigh Fading Model

Let there be two multipath signals S_1 and S_2 received at two different time instants due to the presence of obstacles as shown in Figure 5.6. Now there can either be constructive or destructive interference between the two signals.

Let E_n be the electric field and Θ_n be the relative phase of the various multipath signals.So we have

$$\widetilde{E} = \sum_{n=1}^{N} E_n e^{j\theta_n}$$

Now if N→ ∞(i.e. are sufficiently large number of multipaths) and all the E_n are IID distributed, then by Central Limit Theorem we have,

$$\lim_{N\to\infty} \tilde{E} = \lim_{N\to\infty} \sum_{n=1}^{N} E_n e^{j\theta_n}$$

$$= Z_r + jZ_i = Re^{j\varphi}$$

where Z_r and Z_i are Gaussian Random variables. For the above case

$$R = \sqrt{Z_r^2 + Z_i^2}$$

and

$$\varphi = \tan^{-1}\frac{Z_i}{Z_r}$$

For all practical purposes we assume that the relative phase Θ_n is uniformaly distributed.

$$E[e^{j\theta_n}] = \frac{1}{2\pi}\int^{2\pi} e^{j\theta}d\theta = 0$$

It can be seen that E_n and Θ_n are independent. So,

$$E[\widetilde{E}] = E[\sum E_n e^{j\theta_n}] = 0$$

$$E[|\widetilde{E}|^2] = E[\sum E_n e^{j\theta_n} \sum E_n^* e^{-j\theta_n}] = E[\sum_m \sum_n E_n E_m e^{j(\theta_n - \theta_m)}] = \sum_{n=1}^{N} E_n^2 = P_0$$

where P_0 is the total power obtained. To find the Cumulative Distribution Func- tion(CDF) of R, we proceed as follows.

$$F_R(r) = P_r(R \leq r) = \int\int_A f_{Z_i,Z_r}(z_i, z_r)dz_i dz_r$$

where A is determined by the values taken by the dummy variable r. Let Z_i and Z_r be zero mean Gaussian RVs. Hence the CDF can be written as

$$F_R(r) = \int\int_A \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(Z_r^2+Z_i^2)}{2\sigma^2}} dZ_i dZ_r$$

Let $Z_r = p\cos(\Theta)$ and $Z_i = p\sin(\Theta)$ So we have

$$F_R(r) = \int_0^{2\pi}\int_0^{2\pi} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-p^2}{2\sigma^2}} p\,dp\,d\theta$$

$$= 1 - e^{\frac{-r^2}{2\sigma^2}}$$

Above equation is valid for all $r \geq 0$. The pdf can be written as

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}$$

and is shown in Figure with different σ values. This equation too is valid for all $r \geq 0$. Above distribution is known as Rayleigh distribution and it has been derived

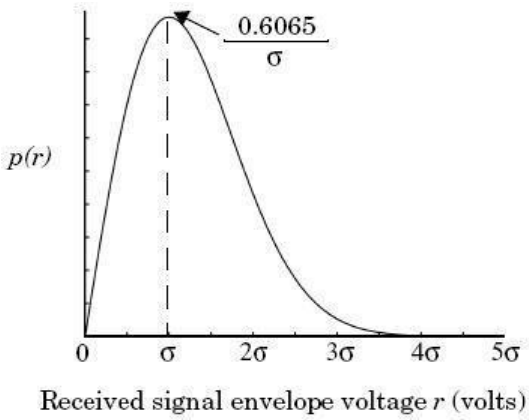


Figure 5.7: Rayleigh probability density function.

for slow fading. However, if f_D 1 Hz, we call it as Quasi-stationary Rayleigh fading. We observe the following:

$$E[R] = \sqrt{\frac{\pi}{2}}\sigma$$

$$E[R^2] = 2\sigma^2$$

$$\mathrm{var}[R] = (2 - \frac{\pi}{2})\sigma^2$$

$$median[R] = 1.77\sigma.$$

LoS Propagation: Rician Fading Model

Rician Fading is the addition to all the normal multipaths a direct LOS path.

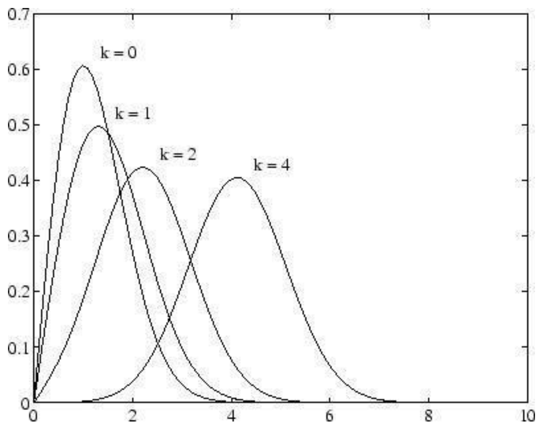


Figure : Ricean probability density function.

$$f_R(r) = \frac{r}{\sigma^2} e^{\frac{-(r^2+A^2)}{2\sigma^2}} I_0\left(\frac{A_r}{\sigma^2}\right)$$

for all $A \geq 0$ and $r \geq 0$. Here A is the peak amplitude of the dominant signal and $I_0(.)$ is the modified Bessel function of the first kind and zeroth order.

A factor K is defined as

$$K_{dB} = 10 \log \frac{A^2}{2\sigma^2}$$

As A $\rightarrow$ 0 then $K_{dB} \rightarrow \infty$.

Generalized Model: Nakagami Distribution

A generalization of the Rayleigh and Rician fading is the Nakagami distribution.

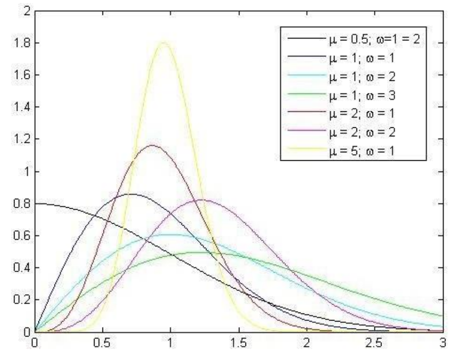


Figure: Nakagami probability density function.

Its pdf is given as,

$$f_R(r) = \frac{2r^{m-1}}{\Gamma(m)} \left(\frac{m^m}{\Omega^m}\right) e^{\frac{-mr^2}{\Omega}}$$

where,

$\Gamma(m)$ is the gamma function

Ω is the average signal power and

m is the fading factor.It is always greater than or equal to 0.5.

When m=1, Nakagami model is the Rayleigh model. When

$$m = \frac{(M+1)^2}{2M+1}$$

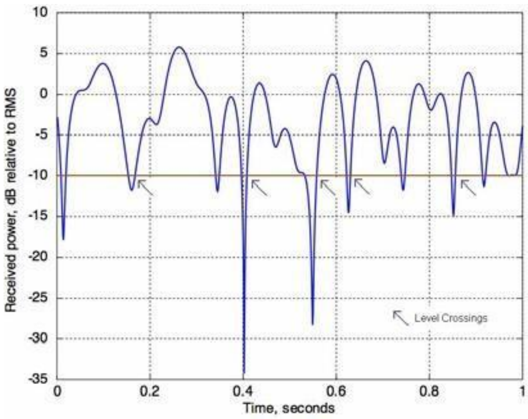


Figure : Schematic representation of level crossing with a Rayleigh fading enve- lope at 10 Hz Doppler spread.

where

$$M = \frac{A}{2\sigma}$$

Nakagami fading is the Rician fading.

As m $\rightarrow\infty$ Nakagami fading is the impulse channel and no fading occurs.

Second Order Statistics

To design better error control codes, we have two important second order param- eters of fading model, namely the **level crossing rate** (LCR) and **average fade duration** (AFD). These parameters can be utilized to assess the speed of the user by measuring them through the reverse channel. The LCR is the expected rate at which the Rayleigh fading envelope normalized to the local rms amplitude crosses a specific level 'R' in a positive going direction.

$$N_R = \int_0^{\infty} \dot{r}\, p(R, \dot{r})\, d\dot{r} = \sqrt{2\pi} f_D \rho e^{-\rho^2}$$

where $\dot{r}$ is the time derivative of r(t), f_D is the maximum Doppler shift and ρ is the value of the specified level R, normalized to the local rms amplitude of the fading envelope.

The other important parameter, AFD, is the average period time for which the receiver power is below a specified level R.

$$\bar{\tau} = \frac{1}{N_r} P_r(r \le R)$$

As

$$P_r(r \le R) = \int_0^R p(r)dr = 1 - e^{-\rho^2},$$

therefore,

$$\bar{\tau} = \frac{1 - e^{-\rho^2}}{\sqrt{2\pi} f_D \rho e^{-\rho^2}} = \frac{e^{\rho^2} - 1}{\sqrt{2\pi} f_D \rho}.$$

Apart from LCR, another parameter is fading rate, which is defined as the number of times the signal envelope crosses the middle value (r_m) in a positive going direction per unit time. The average rate is expressed as

$$N(r_m) = \frac{2v}{\lambda}.$$

Another statistical parameter, sometimes used in the mobile communication, is called as depth of fading. It is defined as the ratio between the minimum value and the mean square value of the faded signal. Usually, an average value of 10% as depth of fading gives a marginal fading scenario.

Simulation of Rayleigh Fading Models

Clarke’s Model: without Doppler Effect

In it, two independent Gaussian low pass noise sources are used to produce in-phase and quadrature fading branches. This is the basic model and is useful for slow fading channel. Also the Doppler effect is not accounted for.

Clarke and Gans’ Model: with Doppler Effect

In this model, the output of the Clarke’s model is passed through Doppler filter in the RF or through two initial baseband Doppler filters for baseband processing as shown in Figure 5.11. Here, the obtained Rayleigh output is flat faded signal but not frequency selective.

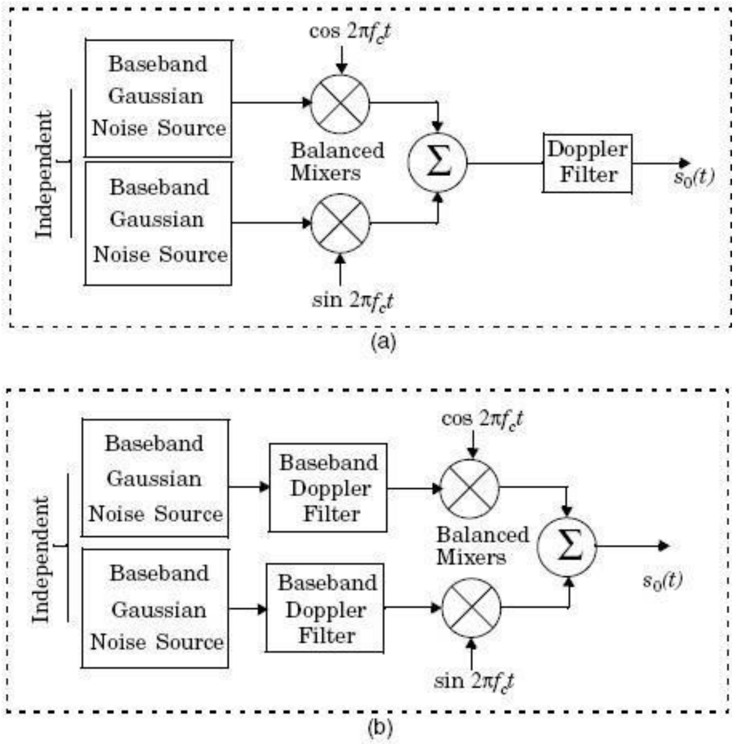


Figure: Clarke and Gan's model for Rayleigh fading generation using quadra- ture amplitude modulation with (a) RF Doppler filter, and, (b) baseband Doppler filter.

Rayleigh Simulator with Wide Range of Channel Conditions

To get a frequency selective output we have the following simulator through which both the frequency selective and flat faded Rayleigh signal may be obtained. This is achieved through varying the parameters a_i and τ_i, as given in Figure

Two-Ray Rayleigh Faded Model

The above model is, however, very complex and difficult to implement. So, we have the two ray Rayleigh fading model which can be easily implemented in software as shown in Figure 5.13.

$$h_b(t) = \alpha_1 e^{j\varphi_1}\delta(t) + \alpha_2 e^{j\varphi_2}\delta(t-\tau)$$

where α_1 and α_2 are independent Rayleigh distributed and φ_1 and φ_2 are indepen- dent and uniformaly distributed over 0 to 2π. By varying τ it is possible to create a wide range of frequency selective fading effects.

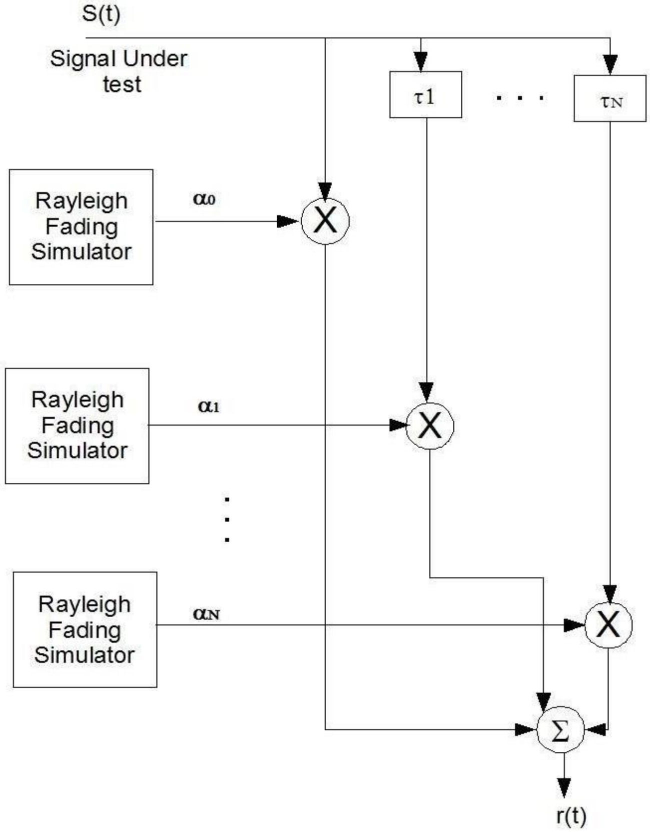


Figure : Rayleigh fading model to get both the flat and frequency selective channel conditions.

Saleh and Valenzuela Indoor Statistical Model

This method involved averaging the square law detected pulse response while sweep- ing the frequency of the transmitted pulse. The model assumes that the multipath components arrive in clusters. The amplitudes of the received components are in- dependent Rayleigh random variables with variances that decay exponentially with cluster delay as well as excess delay within a cluster. The clusters and multipath components within a cluster form Poisson arrival processes with different rates.

SIRCIM/SMRCIM Indoor/Outdoor Statistical Models

SIRCIM (Simulation of Indoor Radio Channel Impulse-response Model) generates realistic samples of small-scale indoor channel impulse response measurements. Sub-

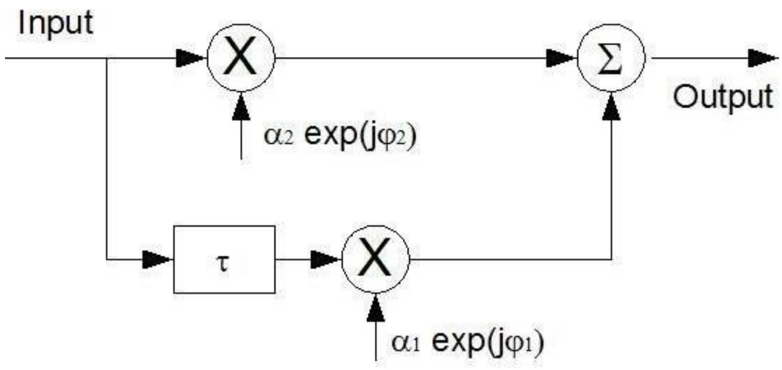


Figure 5.13: Two-ray Rayleigh fading model.

sequent work by Huang produced SMRCIM (Simulation of Mobile Radio Channel Impulse-response Model), a similar program that generates small-scale urban cellu- lar and micro-cellular channel impulse responses.

Conclusion

In this chapter, the main channel impairment, i.e., fading, has been introduced which becomes so severe sometimes that even the large scale path loss becomes insignificant in comparison to it. Some statistical propagation models have been presented based on the fading characteristics. Mainly the frequency selective fading, fast fading and deep fading can be considered the major obstruction from the channel severity view point.

UNIT IV

Equalization and Diversity

Introduction

Apart from the better transmitter and receiver technology, mobile communications require signal processing techniques that improve the link performance. Equaliza- tion, Diversity and channel coding are channel impairment improvement techniques. Equalization compensates for Inter Symbol Interference (ISI) created by multipath within time dispersive channels. An equalizer within a receiver compensates for the average range of expected channel amplitude and delay characteristics. In other words, an equalizer is a filter at the mobile receiver whose impulse response is inverse of the channel impulse response. As such equalizers find their use in frequency selec- tive fading channels. Diversity is another technique used to compensate fast fading and is usually implemented using two or more receiving antennas. It is usually em- ployed to reduce the depths and duration of the fades experienced by a receiver in a flat fading channel. Channel coding improves mobile communication link perfor- mance by adding redundant data bits in the transmitted message.At the baseband portion of the transmitter, a channel coder maps a digital message sequence in to another specific code sequence containing greater number of bits than original con- tained in the message. Channel Coding is used to correct deep fading or spectral null. We discuss all three of these techniques in this chapter. A general framework of the fading effects and their mitigation techniques is shown in Figure 4.1.

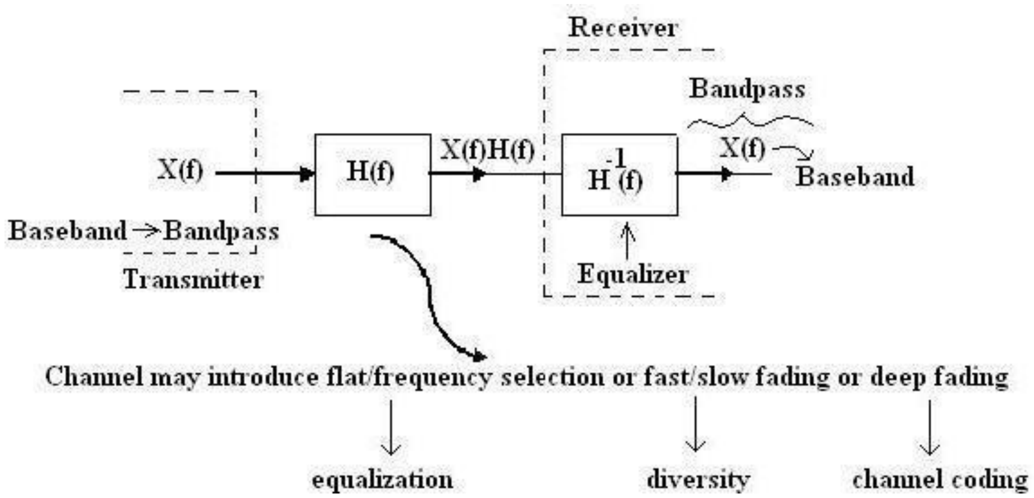


Figure 4.1: A general framework of fading effects and their mitigation techniques.

Equalization

ISI has been identified as one of the major obstacles to high speed data transmission over mobile radio channels. If the modulation bandwidth exceeds the coherence bandwidth of the radio channel (i.e., frequency selective fading), modulation pulses are spread in time, causing ISI. An equalizer at the front end of a receiver compen- sates for the average range of expected channel amplitude and delay characteristics. As the mobile fading channels are random and time varying, equalizers must track the time-varying characteristics of the mobile channel and therefore should be time- varying or adaptive. An adaptive equalizer has two phases of operation: training and tracking. These are as follows.

Training Mode:

- Initially a known, fixed length training sequence is sent by the transmitter so that the receiver equalizer may average to a proper setting.
- Training sequence is typically a pseudo-random binary signal or a fixed, of prescribed bit pattern.
- The training sequence is designed to permit an equalizer at the receiver to acquire the proper filter coefficient in the worst possible channel condition. An adaptive filter at the receiver thus uses a recursive algorithm to evaluate the channel and estimate filter coefficients to compensate for the channel.

Tracking Mode:

- When the training sequence is finished the filter coefficients are near optimal.
- Immediately following the training sequence, user data is sent.
- When the data of the users are received, the adaptive algorithms of the equal- izer tracks the changing channel.
- As a result, the adaptive equalizer continuously changes the filter characteris- tics over time.

A Mathematical Framework

The signal received by the equalizer is given by

$$x(t) = d(t) * h(t) + n_b(t) \tag{4.1}$$

where $d(t)$ is the transmitted signal, $h(t)$ is the combined impulse response of the transmitter,channel and the RF/IF section of the receiver and $n_b(t)$ denotes the baseband noise.

If the impulse response of the equalizer is $h_{eq}(t)$, the output of the equalizer is

$$\hat{y}(t) = d(t) * h(t) * h_{eq}(t) + n_b(t) * h_{eq}(t) = d(t) * g(t) + n_b(t) * h_{eq}(t) . \quad (4.2)$$

However, the desired output of the equalizer is $d(t)$ which is the original source data. Assuming $n_b(t)$=0, we can write $y(t) = d(t)$, which in turn stems the following equation:

$$g(t) = h(t) * h_{eq}(t) = \delta(t) \quad (4.3)$$

The main goal of any equalization process is to satisfy this equation optimally. In frequency domain it can be written as

$$H_{eq}(f)\, H(f) = 1 \quad (4.4)$$

which indicates that an equalizer is actually an inverse filter of the channel. If the channel is frequency selective, the equalizer enhances the frequency components with small amplitudes and attenuates the strong frequencies in the received frequency spectrum in order to provide a flat, composite received frequency response and linear phase response. For a time varying channel, the equalizer is designed to track the channel variations so that the above equation is approximately satisfied.

Zero Forcing Equalization

In a zero forcing equalizer, the equalizer coefficients c_n are chosen to force the samples of the combined channel and equalizer impulse response to zero. When each of the delay elements provide a time delay equal to the symbol duration T, the frequency response $H_{eq}(f)$ of the equalizer is periodic with a period equal to the symbol rate 1/T. The combined response of the channel with the equalizer must satisfy Nyquist's criterion

$$H_{ch}(f)\, H_{eq}(f) = 1, |f| < {}^{1}/_{2T} \quad (4.5)$$

where $H_{ch}(f)$ is the folded frequency response of the channel. Thus, an infinite length zero-forcing ISI equalizer is simply an inverse filter which inverts the folded frequency response of the channel.

Disadvantage: Since $H_{eq}(f)$ is inverse of $H_{ch}(f)$ so inverse filter may excessively amplify the noise at frequencies where the folded channel spectrum has high atten- uation, so it is rarely used for wireless link except for static channels with high SNR such as local wired telephone. The usual equalizer model follows a time varying or adaptive structure which is given next.

A Generic Adaptive Equalizer

The basic structure of an adaptive filter is shown in Figure 4.2. This filter is called the transversal filter, and in this case has N delay elements, N+1 taps and N+1 tunable complex multipliers, called weights. These weights are updated continuously by an adaptive algorithm. In the figure the subscript k represents discrete time index. The adaptive algorithm is controlled by the error signal e_k. The error signal is derived by comparing the output of the equalizer, with some signal d_k which is replica of transmitted signal. The adaptive algorithm uses e_k to minimize the cost function and uses the equalizer weights in such a manner that it minimizes the cost function iteratively. Let us denote the received sequence vector at the receiver and

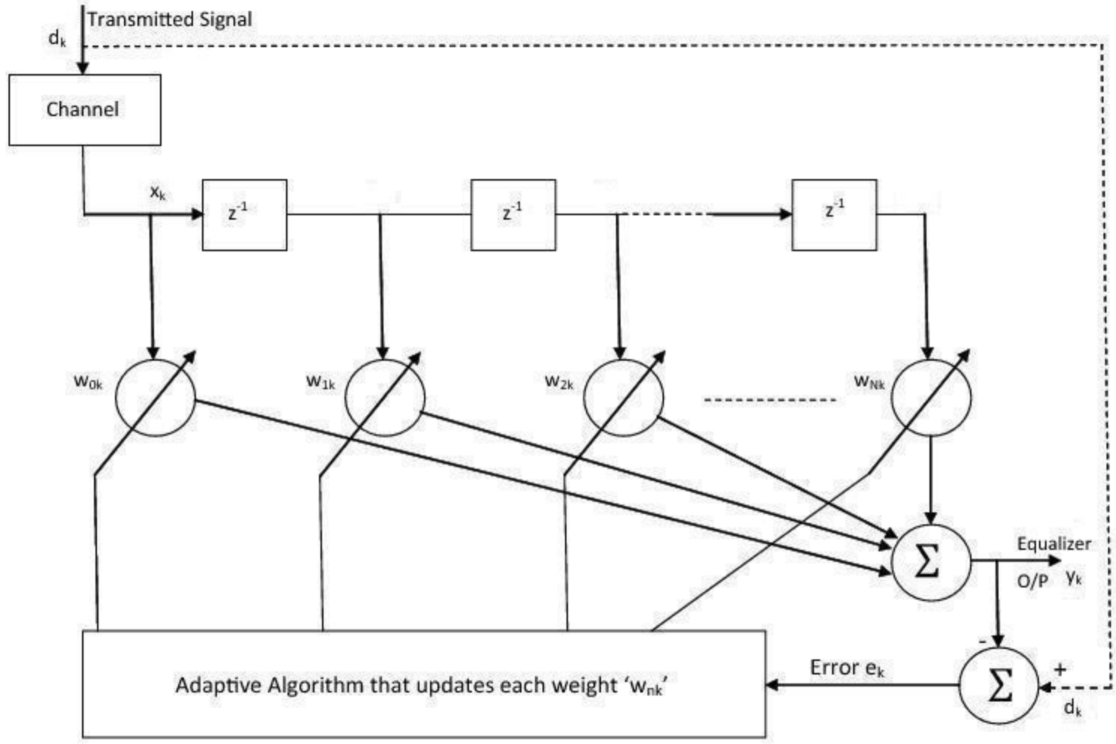


Figure 4.2: A generic adaptive equalizer.

the input to the equalizer as

$$\mathbf{x}_k = [x_k, x_{k-1}, \ldots.., x_{k-N}]^T, \tag{4.6}$$

and the tap coefficient vector as

$$\mathbf{w}_k = [w_k^0, w_k^1, \ldots.., w_k^N]^T. \tag{4.7}$$

Now, the output sequence of the equalizer y_k is the inner product of $\mathbf{x}_k$ and $\mathbf{w}_k$, i.e.,

$$y_k = \langle \mathbf{x}_k, \mathbf{w}_k \rangle = \mathbf{x}_k^T \mathbf{w}_k = \mathbf{w}_k^T \mathbf{x}_k. \tag{4.8}$$

The error signal is defined as

$$e_k = d_k - y_k = d_k - \mathbf{x}_k^T \mathbf{w}_k. \tag{4.9}$$

Assuming d_k and $\mathbf{x}_k$ to be jointly stationary, the Mean Square Error (MSE) is given as

$$\begin{aligned} MSE &= E[e_k^2] = E[(d_k - y_k)^2] \\ &= E[(d_k - \mathbf{x}_k^T \mathbf{w}_k)^2] \\ &= E[d_k^2] + \mathbf{w}_k^T E[\mathbf{x}_k \mathbf{x}_k^T]\mathbf{w}_k - 2E[d_k \mathbf{x}_k^T]\mathbf{w}_k \end{aligned} \tag{4.10}$$

where $\mathbf{w}_k$ is assumed to be an array of optimum values and therefore it has been taken out of the $E()$ operator. The MSE then can be expressed as

$$MSE = \xi = \sigma_d^2 + \mathbf{w}_k^T \mathbf{R} \mathbf{w}_k - 2\mathbf{p}^T \mathbf{w}_k \tag{4.11}$$

where the signal variance $\sigma_d^2 = E[d_k^2]$ and the cross correlation vector $\mathbf{p}$ between the desired response and the input signal is defined as

$$p = E\left[d_k \mathbf{x}_k\right] = E\left[\begin{matrix} d_k x_k & d_k x_{k-1} & d_k x_{k-2} & \cdots & d_k x_{k-N} \end{matrix}\right]^T. \tag{4.12}$$

The input correlation matrix $\mathbf{R}$ is defined as an $(N+1) \times (N+1)$ square matrix, where

$$R = E\left[\mathbf{x}_k \mathbf{x}_k^T\right] = E\begin{bmatrix} x_k^2 & x_k x_{k-1} & x_k x_{k-2} & \cdots & x_k x_{k-N} \\ x_{k-1} x_k & x_{k-1}^2 & x_{k-1} x_{k-2} & \cdots & x_{k-1} x_{k-N} \\ x_{k-2} x_k & x_{k-2} x_{k-1} & x_{k-2}^2 & \cdots & x_{k-2} x_{k-N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{k-N} x_k & x_{k-N} x_{k-1} & x_{k-N} x_{k-2} & \cdots & x_{k-N}^2 \end{bmatrix}. \tag{4.13}$$

Clearly, MSE is a function of $\mathbf{w}_k$. On equating $\frac{\partial \xi}{\partial w_k}$

to 0, we get the condition for
minimum MSE (MMSE) which is known as Wiener solution:

$$\mathbf{w}_k = \mathbf{R}^{-1}\mathbf{p}. \tag{4.14}$$

Hence, MMSE is given by the equation

$$MMSE = \xi_{min} = \sigma_d^2 - \mathbf{p}^T\mathbf{w}_k. \tag{4.15}$$

Choice of Algorithms for Adaptive Equalization

Since an adaptive equalizer compensates for an unknown and time varying channel, it requires a specific algorithm to update the equalizer coefficients and track the channel variations. Factors which determine algorithm's performance are:

Rate of convergence: Number of iterations required for an algorithm, in re- sponse to a stationary inputs, to converge close enough to optimal solution. A fast rate of convergence allows the algorithm to adapt rapidly to a stationary environ- ment of unknown statistics.

Misadjustment: Provides a quantitative measure of the amount by which the final value of mean square error, averaged over an ensemble of adaptive filters, deviates from an optimal mean square error.

Computational complexity: Number of operations required to make one com- plete iteration of the algorithm.

Numerical properties: Inaccuracies like round-off noise and representation errors in the computer, which influence the stability of the algorithm.

Three classic equalizer algorithms are primitive for most of today's wireless stan- dards. These include the Zero Forcing Algorithm (ZF), the Least Mean Square Algo- rithm (LMS), and the Recursive Least Square Algorithm (RLS). Below, we discuss a few of the adaptive algorithms.

Least Mean Square (LMS) Algorithm

LMS algorithm is the simplest algorithm based on minimization of the MSE between the desired equalizer output and the actual equalizer output, as discussed earlier. Here the system error, the MSE and the optimal Wiener solution remain the same as given the adaptive equalization framework.

In practice, the minimization of the MSE is carried out recursively, and may be performed by use of the stochastic gradient algorithm. It is the simplest equalization algorithm and requires only 2N+1 operations per iteration. The filter weights are updated by the update equation. Letting the variable n denote the sequence of iteration, LMS is computed iteratively by

$$w_k(n+1) = w_k(n) + \mu e_k(n)\, x(n-k) \tag{4.16}$$

where the subscript k denotes the kth delay stage in the equalizer and μ is the step size which controls the convergence rate and stability of the algorithm.

The LMS equalizer maximizes the signal to distortion ratio at its output within the constraints of the equalizer filter length. If an input signal has a time dispersion characteristics that is greater than the propagation delay through the equalizer, then the equalizer will be unable to reduce distortion. The convergence rate of the LMS algorithm is slow due to the fact that there is only one parameter, the step size, that controls the adaptation rate. To prevent the adaptation from becoming unstable, the value of μ is chosen from

$$0 < \mu < \frac{2}{\sum_{i=1}^{N} \lambda_i} \qquad (4.17)$$

where λ_i is the i-th eigenvalue of the covariance matrix R.

Normalized LMS (NLMS) Algorithm

In the LMS algorithm, the correction that is applied to $w_k(n)$ is proportional to the input sample $x(n-k)$. Therefore when $x(n-k)$ is large, the LMS algorithm experiences gradient noise amplification. With the normalization of the LMS step size by $\|\mathbf{x}(n)\|^2$ in the NLMS algorithm, this problem is eliminated. Only when $x(n-k)$ becomes close to zero, the denominator term $\|\mathbf{x}(n)\|^2$ in the NLMS equation becomes very small and the correction factor may diverge. So, a small positive number ε is added to the denominator term of the correction factor. Here, the step size is time varying and is expressed as

$$\mu(n) = \frac{\beta}{\|\mathbf{x}(n)\|^2 + \varepsilon}. \tag{4.18}$$

Therefore, the NLMS algorithm update equation takes the form of

$$w_k(n+1) = w_k(n) + \frac{\beta}{\|\mathbf{x}(n)\|^2 + \varepsilon} e_k(n)\, x(n-k) . \tag{4.19}$$

Diversity

Diversity is a method used to develop information from several signals transmitted over independent fading paths. It exploits the random nature of radio propagation by finding independent signal paths for communication. It is a very simple concept where if one path undergoes a deep fade, another independent path may have a strong signal. As there is more than one path to select from, both the instantaneous and average SNRs at the receiver may be improved. Usually diversity decisions are made by receiver. Unlike equalization, diversity requires no training overhead as a training sequence is not required by transmitter. Note that if the distance between two receivers is a multiple of $\lambda/2$, there might occur a destructive interference be- tween the two signals. Hence receivers in diversity technique are used in such a way that the signal received by one is independent of the other. Diversity can be of various forms, starting from space diversity to time diversity. We take up the types one by one in the sequel.

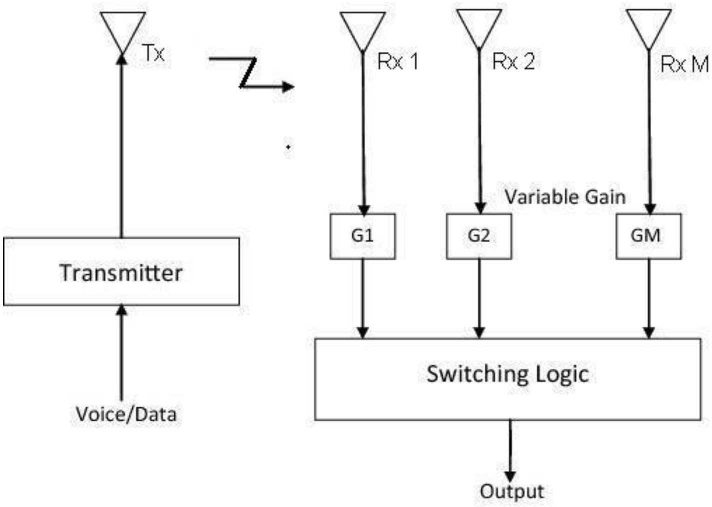


Figure 4.3: Receiver selection diversity, with M receivers.

Different Types of Diversity

Space Diversity

A method of transmission or reception, or both, in which the effects of fading are minimized by the simultaneous use of two or more physically separated antennas, ideally separated by one half or more wavelengths. Signals received from spatially separated antennas have uncorrelated envelopes. Space diversity reception methods can be classified into four categories: selection, feedback or scanning, maximal ratio combining and equal gain combining.

(a) Selection Diversity:

The basic principle of this type of diversity is selecting the best signal among all the signals received from different branches at the receiving end. Selection Diversity is the simplest diversity technique. Figure 7.3 shows a block diagram of this method where 'M' demodulators are used to provide M diversity branches whose gains are adjusted to provide the same average SNR for each branch. The receiver branches having the highest instantaneous SNR is connected to the demodulator.

Let M independent Rayleigh fading channels are available at a receiver. Each channel is called a diversity branch and let each branch has the same average SNR. The signal to noise ratio is defined as

$$SNR = \Gamma = \frac{E_b}{N_0}\alpha^2 \qquad (4.20)$$

where E_b is the average carrier energy, N_0 is the noise PSD, α is a random variable used to represent amplitude values of the fading channel.

The instantaneous SNR(γ_i) is usually defined as γ_i = instantaneous signal power per branch/mean noise power per branch. For Rayleigh fading channels, α has a Rayleigh distribution and so α^2 and consequently γ_i have a chi-square distribution

with two degrees of freedom. The probability density function for such a channel is

$$p(\gamma_i) = \frac{1}{\Gamma} e^{\frac{-\gamma_i}{\Gamma}} \tag{7.21}$$

The probability that any single branch has an instantaneous SNR less than some defined threshold γ is

$$\Pr[\gamma_i \le \gamma] = \int_0^{\gamma} p(\gamma_i)\, d\gamma_i = \int_0^{\gamma} \frac{1}{\Gamma} e^{\frac{-\gamma_i}{\Gamma}}\, d\gamma_i = 1 - e^{\frac{-\gamma}{\Gamma}} = P(\Gamma). \tag{7.22}$$

Similarly, the probability that all M independent diversity branches receive signals which are simultaneously less than some specific SNR threshold γ is

$$\Pr[\gamma_1, \gamma_2, \ldots, \gamma_M \le \gamma] = \left(1 - e^{\frac{-\gamma}{\Gamma}}\right)^M = P_M(\gamma) \tag{7.23}$$

where $P_M(\gamma)$ is the probability of all branches failing to achieve an instantaneous SNR $= \gamma$. Quite clearly, $P_M(\Gamma) < P(\Gamma)$. If a single branch achieves SNR $> \gamma$, then the probability that SNR $> \gamma$ for one or more branches is given by

$$\Pr[\gamma_i > \gamma] = 1 - P_M(\gamma) = 1 - \left(1 - e^{\frac{-\gamma}{\Gamma}}\right)^M \tag{7.24}$$

which is more than the required SNR for a single branch receiver. This expression shows the advantage when a selection diversity is used.

To determine of average signal to noise ratio, we first find out the pdf of γ as

$$p_M(\gamma) = \frac{d}{d\gamma} P_M(\gamma) = \frac{M}{\Gamma}\left(1 - e^{-\gamma/\Gamma}\right)^{M-1} e^{-\gamma/\Gamma}. \tag{7.25}$$

The average SNR, $\bar{\gamma}$, can be then expressed as

$$\bar{\gamma} = \int_0^{\infty} \gamma\, p_M(\gamma)\, d\gamma = \Gamma \int_0^{\infty} Mx\left(1 - e^{-x}\right)^{M-1} e^{-x} dx \tag{7.26}$$

where $x = \gamma/\Gamma$ and Γ is the average SNR for a single branch, when no diversity is used.

This equation shows an average improvement in the link margin without requir- ing extra transmitter power or complex circuitry, and it is easy to implement as it needed a monitoring station and an antenna switch at the receiver. It is not an optimal diversity technique as it doesn't use all the possible branches simultaneously.

(b) Feedback or Scanning Diversity:

Scanning all the signals in a fixed sequence until the one with SNR more than a predetermined threshold is identified. Feedback or scanning diversity is very similar to selection diversity except that instead of always using the best of N signals, the N signals are scanned in a fixed sequence until one is found to be above a predetermined threshold. This signal is then received until it falls below threshold and the scanning process is again initiated. The resulting fading statistics are somewhat inferior, but the advantage is that it is very simple to implement(only one receiver is required).

(c) Maximal Ratio Combining:

Signals from all of the m branches are weighted according to their individual signal voltage to noise power ratios and then summed. Individual signals must be cophased before being summed, which generally requires an individual receiver and phasing circuit for each antenna element. Produces an output SNR equal to the sum of all individual SNR. Advantage of producing an output with an acceptable SNR even when none of the individual signals are themselves acceptable. Modern DSP techniques and digital receivers are now making this optimal form, as it gives the best statistical reduction of fading of any known linear diversity combiner. In terms of voltage signal,

$$r_m = \sum_{i=1}^{m} tt_i r_i \tag{7.27}$$

where tt_i is the gain and r_i is the voltage signal from each branch.

(d) Equal Gain Combining:

In some cases it is not convenient to provide for the variable weighting capability required for true maximal ratio combining. In such cases, the branch weights are all set unity, but the signals from each branch are co-phased to provide equal gain combining diversity. It allows the receiver to exploit signals that are simultaneously received on each branch. Performance of this method is marginally inferior to max- imal ratio combining and superior to Selection diversity. Assuming all the tt_i to be

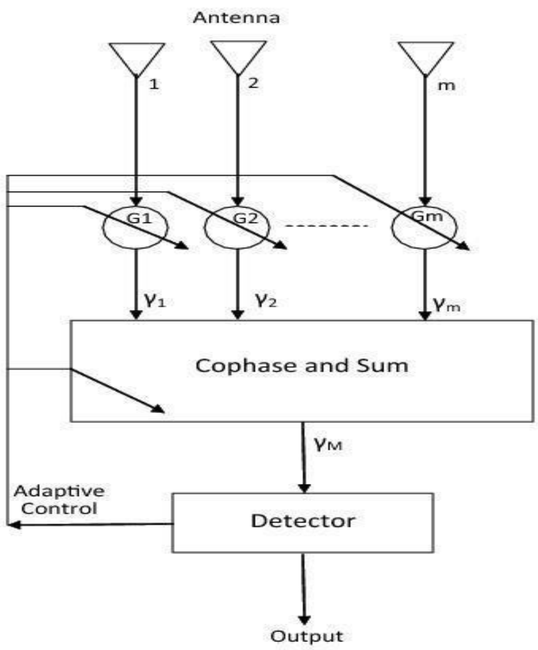


Figure 4.4: Maximal ratio combining technique.

unity, here,

$$r_m = \sum_{i=1}^{m} r_i. \tag{7.28}$$

Polarization Diversity

Polarization Diversity relies on the decorrelation of the two receive ports to achieve diversity gain. The two receiver ports must remain cross-polarized. Polarization Diversity at a base station does not require antenna spacing. Polarization diversity combines pairs of antennas with orthogonal polarizations (i.e. horizontal/vertical, ± slant 45^o, Left-hand/Right-hand CP etc). Reflected signals can undergo polarization changes depending on the channel. Pairing two complementary polarizations, this scheme can immunize a system from polarization mismatches that would otherwise cause signal fade. Polarization diversity has prove valuable at radio and mobile com- munication base stations since it is less susceptible to the near random orientations of transmitting antennas.

Frequency Diversity

In Frequency Diversity, the same information signal is transmitted and received simultaneously on two or more independent fading carrier frequencies. Rationale behind this technique is that frequencies separated by more than the coherence bandwidth of the channel will be uncorrelated and will thus not experience the same fades. The probability of simultaneous fading will be the product of the individual fading probabilities. This method is employed in microwave LoS links which carry several channels in a frequency division multiplex mode (FDM). Main disadvantage is that it requires spare bandwidth also as many receivers as there are channels used for the frequency diversity.

Time Diversity

In time diversity, the signal representing the same information are sent over the same channel at different times. Time diversity repeatedly transmits information at time spacings that exceeds the coherence time of the channel. Multiple repetition of the signal will be received with independent fading conditions, thereby providing for diversity. A modern implementation of time diversity involves the use of RAKE receiver for spread spectrum CDMA, where the multipath channel provides redun- dancy in the transmitted message. Disadvantage is that it requires spare bandwidth also as many receivers as there are channels used for the frequency diversity. Two important types of time diversity application is discussed below.

Application 1: RAKE Receiver

In CDMA spread spectrum systems, CDMA spreading codes are designed to provide very low correlation between successive chips, propagation delay spread in the radio channel provides multiple version of the transmitted signal at the receiver. Delaying multipath components by more than a chip duration, will appear like uncorrelated noise at a CDMA receiver. CDMA receiver may combine the time delayed versions of the original signal to improve the signal to noise ratio at the receiver. RAKE

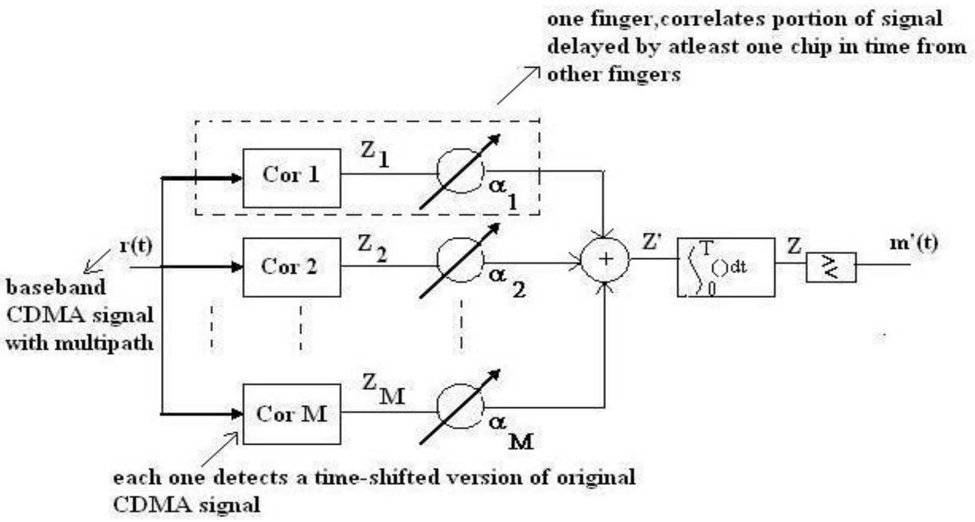


Figure 4.5: RAKE receiver.

receiver collect the time shifted versions of the original signal by providing a sep- arate correlation receiver for M strongest multipath components. Outputs of each correlator are weighted to provide a better estimate of the transmitted signal than provided by a single component. Demodulation and bit decisions are based on the weighted output of the correlators. Schematic of a RAKE receiver is shown in Figure 4.5.

Application 2: Interleaver

In the encoded data bits, some source bits are more important than others, and must be protected from errors. Many speech coder produce several important bits in succession. Interleaver spread these bit out in time so that if there is a deep fade or noise burst, the important bits from a block of source data are not c orrupted at the same time. Spreading source bits over time, it becomes possible to make use of error control coding. Interleaver can be of two forms, a block structure or a convolutional structure.

A block interleaver formats the encoded data into a rectangular array of m rows and n columns, and interleaves nm bits at a time. Each row contains a word of source data having n bits. an interleaver of degree m consists of m rows. source bits are placed into the interleaver by sequentially increasing the row number for each successive bit, and forming the columns. The interleaved source data is then read out row-wise and transmitted over the channel. This has the effect of separating the original source bits by m bit periods. At the receiver, de-interleaver stores the received data by sequentially increasing the row number of each successive bit, and then clocks out the data row-wise, one word at a time. Convolutional interleavers are ideally suited for use with convolutional codes.

Channel Coding

In channel coding, redundant data bits are added in the transmitted message so that if an instantaneous fade occurs in the channel, the data may still be recov- ered at the receiver without the request of retransmission. A channel coder maps the transmitted message into another specific code sequence containing more bits. Coded message is then modulated for transmission in the wireless channel. Channel Coding is used by the receiver to detect or correct errors introduced by the channel. Codes that used to detect errors, are error detection codes. Error correction codes can detect and correct errors.

Shannon's Channel Capacity Theorem

In 1948, Shannon showed that by proper encoding of the information, errors induced by a noise channel can be reduced to any desired level without sacrificing the rate of information transfer. Shannon's channel capacity formula is applicable to the AWGN channel and is given by:

$$C = B\log_2\left(1 + \frac{S}{N}\right) = B\log_2\left(1 + \frac{P}{N_0 B}\right) = B\log_2\left(1 + \frac{E_b R_b}{N_0 B}\right) \qquad (4.29)$$

where C is the channel capacity (bit/s), B is the channel bandwidth (Hz), P is the received signal power (W), N_0 is the single sided noise power density (W/Hz), E_b is the average bit energy and R_b is transmission bit rate. and the ratio C/B is denoted as bandwidth efficiency. Introduction of redundant bits increases the transmission bit rate and hence it increases the bandwidth require- ment, which reduces the bandwidth efficiency of the link in high SNR conditions, but provides excellent BER performance at low SNR values. This leads to the following two inferences.

Corollary 1 : While dealing within maximum channel capacity, introduction of re- dundant bits increase the transmitter rate and hence bandwidth requirement also increases, while decreasing the bandwidth efficiency, but it also decreases the BER. *Corollary 2* : If data redundancy is not introduced in a wideband noisy environment, error free performance in not possible (for example, CDMA communication in 3G mobile phones).

A channel coder operates on digital message (or source) data by encoding the source information into a code sequence for transmission through the channel. The error correction and detection codes are classified into three groups based on their struc- ture.

1. Block Code
2. Convolution Code
3. Concatenated Code.

Block Codes

Block codes are *forward error correction* (FEC) codes that enable a limited number of errors to be detected and corrected without retransmission. Block codes can be used to improve the performance of a communications system when other means of improvement (such as increasing transmitter power or using a more sophisticated demodulator) are impractical.

In block codes, parity bits are added to blocks of message bits to make codewords or code blocks. In a block encoder, k information bits are encoded into n code bits. A total of $n-k$ redundant bits are added to the k information bits for the purpose of detecting and correcting errors. The block code is referred to as an (n, k) code, and the rate of the code is defined as $R_c = k/n$ and is equal to the rate of information divided by the raw channel rate.

Parameters in Block Code

(a) Code Rate (R_c): As defined above, $R_c = k/n$.

(b) Code Distance (d): Distance between two codewords is the number of elements in which two codewords C_i and C_j differs denoted by d (C_i, C_j). If the code used is binary, the distance is known as 'Hamming distance'. For example d(10110, 11011) is 3. If the code 'C' consists of the set of codewords, then the minimum

distance of the code is given by $d_{\min} = min \{d(C_i, C_j)\}$.

(c) Code Weight (w): Weight of a codeword is given by the number of nonzero elements in the codeword. For a binary code, the weight is basically the number of 1s in the codeword. For example weight of a code 101101 is 4.

Ex 1: The block code C = 00000, 10100, 11110, 11001 can be used to represent two bit binary numbers as:

- 00 – 00000
- 01 – 10100
- 10 – 11110
- 11 – 11001

Here number of codewords is 4, k = 2, and n = 5. To encode a bit stream 1001010011

- First step is to break the sequence in groups of two bits, i.e., 10 01 01 00 11
- Next step is to replace each block by its corresponding codeword, i.e., 11110 10100 10100 00000 11001

Quite clearly, here, $d_{\min} = min\{d(C_i,C_j)\}= 2$.

Properties of Block Codes

(a) Linearity: Suppose C_i and C_j are two code words in an (n, k) block code. Let α_1 and α_2 be any two elements selected from the alphabet. Then the code is said to be linear if and only if $\alpha_1 C_1 + \alpha_2 C_2$ is also a code word. A linear code must contain the all-zero code word.

(b) Systematic: A systematic code is one in which the parity bits are appended to the end of the information bits. For an (n, k) code, the first k bits are identical

to the information bits, and the remaining $n - k$ bits of each code word are linear combinations of the k information bits.

(c) Cyclic: Cyclic codes are a subset of the class of linear codes which satisfy the following cyclic shift property: If $C = [C_{n-1}, C_{n-2}, ..., C_0]$ is a code word of a cyclic code, then $[C_{n-2}, C_{n-3}, ..., C_0, C_{n-1}]$, obtained by a cyclic shift of the elements of C, is also a code word. That is, all cyclic shifts of C are code words.

In this context, it is important to know about **Finite Field or Galois Field**. Let F be a finite set of elements on which two binary operations – addition (+) and multiplication (.) are defined. The set F together with the two binary operations is called a *field* if the following conditions are satisfied:

1. F is a commutative group under addition.
2. The set of nonzero elements in F is a commutative group under multiplication.
3. Multiplication is distributive over addition; that is, for any three elements a, b, and c in F, $a(b + c) = ab + ac$
4. Identity elements 0 and 1 must exist in F satisfying $a + 0 = a$ and $a.1 = a$.
5. For any a in F, there exists an additive inverse $(-a)$ such that $a + (-a) = 0$.
6. For any a in F, there exists an multiplicative inverse a^{-1} such that $a.a^{-1} = 1$. Depending upon the number of elements in it, a field is called either a finite or an infinite field. The examples of infinite field include Q (set of all rational numbers), R (set of all real numbers), C (set of all complex numbers) etc. A field with a finite number of elements (say q) is called a 'Galois Field' and is denoted by GF(q). A finite field entity p(x), called a polynomial, is introduced to map all symbols (with several bits) to the element of the finite field. A polynomial is a mathematical expression

$$p(x) = p_0 + p_1x + ... + p_mx_m \quad (7.31)$$

where the symbol x is called the indeterminate and the coefficients $p_0, p_1, ..., p_m$ are the elements of GF(q). The coefficient p_m is called the leading coefficient. If p_m is not equal to zero, then m is called the degree of the polynomial, denoted as deg p(x). A polynomial is called monic if its leading coefficient is unity. The division algorithm states that for every pair of polynomials a(x) and b(x) in F(x), there exists a unique pair of polynomials q(x), the quotient, and r(x), the remainder, such that $a(x) = q(x)b(x) + r(x)$, where deg r(x)¡deg b(x). A polynomial p(x) in F(x) is said to be reducible if p(x)=a(x)b(x), otherwise it is called irreducible. A monic irreducible polynomial of degree at least one is called a prime polynomial.

An irreducible polynomial $p(x)$ of degree ‘m’ is said to be primitive if the smallest integer ‘n’ for which $p(x)$ divides x^n+1 is $n = 2^m-1$. A typical primitive polynomial is given by $p(x) = x^m + x + 1$.

A specific type of code which obeys both the cyclic property as well as poly- nomial operation is cyclic codes. Cyclic codes are a subset of the class of linear codes which satisfy the cyclic property. These codes possess a considerable amount of structure which can be exploited. A cyclic code can be generated by using a generator polynomial g(p) of degree (n-k). The generator polynomial of an (n,k) cyclic code is a factor of $p^n + 1$ and has the form

$$g(p) = p^{n-k} + g_{n-k-1}p^{n-k-1} + \cdots + g_1 p + 1. \qquad (7.32)$$

A message polynomial x(p) can also be defined as

$$x(p) = x_{k-1}p^{k-1} + \cdots + x_1 p + x_0 \qquad (7.33)$$

where $(x_{k-1},..., x_0)$ represents the k information bits. The resultant codeword c(p) can be written as

$$c(p) = x(p)\, g(p) \qquad (7.34)$$

where c(p) is a polynomial of degree less than n. We would see an application of such codes in Reed-Solomon codes.

Examples of Block Codes

(a) Single Parity Check Code: In single parity check codes (example: ASCII code), an overall single parity check bit is appended to ’k’ information bits. Let the infor- mation bit word be: $(b_1, b_2, ..., b_k)$, then parity check bit: $p = b_1 + b_2 + ... + b_k$ modulo 2 is appended at the (k+1)th position, making the overall codeword: $C = (b_1, b_2, ..., b_k, p)$. The parity bit may follow an even parity or an odd parity pattern. All error patterns that change an odd number of bits are detectable, and all even numbered error patterns are not detectable. However, such codes can only detect the error, it cannot correct the error.

Ex. 2: Consider a (8,7) ASCII code with information codeword (0, 1, 0, 1, 1, 0, 0) and encoded with overall even parity pattern. Thus the overall codeword is (0, 1, 0, 1, 1, 0, 0, 1) where the last bit is the parity bit. If there is a single error in bit 3: (0, 1, **1**, 1, 1, 0, 0, 1), then it can be easily checked by the receiver that now there are odd number of 1's in the codeword and hence there is an error. On the other hand, if there are two errors, say, errors in bit 3 and 5: (0, 1, **1**, 1, **0**, 0, 0, 1), then error

will not be detected.

After decoding a received codeword, let p_c be the probability that the decoder gives correct codeword C, p_e is the probability that the decoder gives incorrect codeword $C^j \pm= C$, and p_f is the probability that the decoder fails to give a codeword. In this case, we can write $p_c + p_e + p_f = 1$.

If in an n-bit codeword, there are j errors and p is the bit error probability, then the probability of obtaining j errors in this codeword is $P_j = {}^nC_j p^j (1 - p)^{n-j}$.

Using this formula, for any $(n, n - 1)$ single parity check block code, we get

- $p_c = P_0$,
- $p_e = P_2 + P_4 + ... + P_{n^j}$ ($n^j = n$ if n is even, otherwise $n^j = n - 1$), ($n^j = n - 1$ if n is even, otherwise $n^j = n$).
- $p_f = P_1 + P_3 + ... + P_{n^j}$

As an example, for a (5,4) single parity check block code, $p_c = P_0$, $p_e = P_2 + P_4$, and $p_f = P_1 + P_3 + P_5$.

(b) Product Codes: Product codes are a class of linear block codes which provide error detection capability using product of two block codes. Consider that nine information bits (1, 0, 1, 0, 0, 1, 1, 1, 0) are to be transmitted. These 9 bits can be divided into groups of three information bits and (4,3) single parity check codeword can be formed with even parity. After forming three codewords, those can be ap- pended with a vertical parity bit which will form the fourth codeword. Thus the following codewords are transmitted:

$C1 = [$

$C2 = [$

$C3 = [$

$C4 = [$.

Now if an error occurs in the second bit of the second codeword, the received code- words at the receiver would then be

$$C1 = [1 \quad 0 \quad 1 \quad 0]$$

$C2 = [$ $\leftarrow$

$C3 = [$

$C4 = [$

$\uparrow$

and these would indicate the corresponding row and column position of the erroneous bit with vertical and horizontal parity check. Thus the bit can be corrected. Here we get a horizontal (4, 3) codeword and a vertical (4, 3) codeword and concatenating them we get a (16, 9) product code. In general, a product code can be formed as (n_1, k_1) & $(n_2, k_2) \rightarrow (n_1 n_2, k_1 k_2)$.

(c) Repetition Codes: In a (n,1) repetition code each information bit is repeated n times (n should be odd) and transmitted. At the receiver, the majority decoding principle is used to obtain the information bit. Accordingly, if in a group of n received bit, 1 occurs a higher number of times than 0, the information bit is decoded as 1. Such majority scheme works properly only if the noise affects less than n/2 number of bits.

Ex 3: Consider a (3,1) binary repetition code.

- For input bit 0, the codeword is (0 0 0) and for input bit 1, the codeword is (1 1 1).
- If the received codeword is (0 0 0), i.e. no error, it is decoded as 0.
- Similarly, if the received codeword is (1 1 1), i.e. no error, it is decoded as 1.
- If the received codeword is (0 0 1) or (0 1 0) or (1 0 0), then error is detected and it is decoded as 0 with majority decoding principle.
- If the received codeword is (0 1 1) or (1 1 0) or (1 0 1), once again error is detected and it is decoded as 1 with majority decoding principle.

For such a (3,1) repetition code, $p_c = P_0 + P_1$, $p_e = P_2 + P_3$, and $p_f = 0$.

(d) Hamming Codes: A binary Hamming code has the property that

$$(n, k) = (2^m - 1, 2^m - 1 - m) \qquad (7.35)$$

where k is the number of information bits used to form a n bit codeword, and m is any positive integer. The number of parity symbols are $n - k = m$. Thus, a codeword is represented by $C = [i_1, ...i_n, p_1, ..., p_{n-k}]$. This is quite a useful code in communication which is illustrated via the following example.

Ex 4: Consider a (7, 4) Hamming code. With three parity bits we can correct exactly 1 error. The parity bits may follow such a modulo 2 arithmetic:

$p_1 = i_1 + i_2 + i_3$, $p_2 = i_2 + i_3 + i_4$, $p_3 = i_1 + i_3 + i_4$,

which is same as,

$$+ i_1 + i_2 + i_3 \quad =$$

$$+ i_2 + i_3 + i_4 \quad =$$

$$+ i_1 + i_3 + i_4 \quad =$$

The transmitted codeword is then $C = [i_1, i_2, ..., i_4, p_1, p_2, p_3]$.

Syndrome Decoding: For this Hamming code, let the received codeword be $V = [v_1, v_2, ..., v_4, v_5, v_6, v_7]$.

We define a syndrome vector S as

$$=_1 S_2 S_3]$$

$$= + v_2 + v_3 + v_5$$

$$= + v_3 + v_4 + v_6$$

$$= + v_2 + v4 + v7$$

It is obvious that in case of no error, the syndrome vector is equal to zero. Corre- sponding to this syndrome vector, there is an error vector e which can be obtained from a syndrome table and finally the required codeword is taken as $C = V + e$. In a nutshell, to obtain the required codeword, we perform the following steps:

1. Calculate S from decoder input V.
2. From syndrome table, obtain e corresponding to S.
3. The required codeword is then $C = V + e$.

A few cases are given below to illustrate the syndrome decoding.

1. Let C = [0 1 1 1 0 1 0] and V = [0 1 1 1 0 1 0]. This implies S = [0 0 0], and it corresponds to e = [0 0 0 0 0 0 0]. Thus, C = V + e = [0 1 1 1 0 1 0].

2. Let C = [1 1 0 0 0 1 0] and V = [1 1 0 1 0 1 0]. This means S = [0 1 1], from which we get e = [0 0 0 1 0 0 0] which means a single bit error is there in the received bit v_4. This will be corrected by performing the operation C = V + e.

3. Another interesting case is, let C = [0 1 0 1 1 0 0] and V = [0 0 1 1 1 0 1] (two errors at second and third bits). This makes S = [0 0 0] and as a result, e = [0 0 0 0 0 0 0]. However, $C \pm= V$, and $C = V + e$ implies the double error cannot becorrected. Therefore a (7,4) Hamming code can correct only single bit error.

(e) Golay Codes: Golay codes are linear binary (23,12) codes with a minimum distance of seven and a error correction capability of three bits. This is a special, one of a kind code in that this is the only nontrivial example of a perfect code. Every codeword lies within distance three of any codeword, thus making maximum likelihood decoding possible.

(f) BCH Codes: BCH code is one of the most powerful known class of linear cyclic block codes, known for their multiple error correcting ability, and the ease of encoding and decoding. It's block length is $n = 2^m - 1$ for $m \geq 3$ and number of errors that they can correct is bounded by $t < (2^m - 1)/2$. Binary BCH codes can be generalized to create classes of non binary codes which use m bits per code symbol.

(g) Reed Solomon (RS) Codes: Reed-Solomon code is an important subset of the BCH codes with a wide range of applications in digital communication and data storage. Typical application areas are storage devices (CD, DVD etc.), wireless communications, digital TV, high speed modems. It's coding system is based on groups of bits, such as bytes, rather than individual 0 and 1. This feature makes it particularly good at dealing with burst of errors: six consecutive bit errors. Block length of these codes is $n = 2^m - 1$, and can be extended to 2^m or $2^m + 1$. Number of parity symbols that must be used to correct e errors is $n - k = 2e$. Minimum distance $d_{min} = 2e + 1$, and it achieves the largest possible d_{min} of any linear code.

For US-CDPD, the RS code is used with $m = 6$. So each of the 64 field elements is represented by a 6 bit symbol. For this case, we get the primitive polynomial as $p(x) = x^6 + x + 1$. Equating $p(x)$ to 0 implies $x^6 = x + 1$.

The 6 bit representation of the finite field elements is given in Table 7.1. The table elements continue up to α^{62}. However, to follow linearity property there should be

Table 7.1: Finite field elements for US-CDPD

	α	α	α			
1	0	0	0			
α^1	0	0	0			
α^2	0	0	1			
.	.	.	.			
.	.	.	.			
$\alpha^6 = \alpha + 1$	0	0	0			
.	.	.	.			
.	.	.	.			

a zero codeword, hence α^{63} is assigned zero.

The encoding part of the RS polynomial is done as follows: Information polynomial: $d(x) = C_{n-1}x^{n-1} + C_{n-2}x^{n-2} + ... + C_{2t}x^{2t}$, Parity polynomial: $p(x) = C_{2t-1}x^{2t-1} + ... + C_0$, Codeword polynomial: $c(x) = d(x) + p(x)$.

Since generating an information polynomial is difficult, so a generating polynomial is used instead. Information polynomial is then the multiple of generating polynomial. This process is given below. Since this kind of codes are cyclic codes, we take a generating polynomial $g(x)$ such that $d(x) = g(x)q(x) + r(x)$ where $q(x)$ is the quotient polynomial and $r(x)$ is the remainder polynomial. The codeword polynomial would then be given as: $c(x) = g(x)q(x) + r(x) = p(x)$. If we assign a parity polynomial $p(x) = r(x)$, then the codeword polynomial $c(x) = g(x)p(x)$ and the entire process becomes easier.

On the decoder side one has to find a specific $r(x) = p(x)$ or vice-versa, but due to its complexity, it is mainly done using syndrome calculation. The details of such a syndrome calculation can be found in [1].

Convolutional Codes

A continuous sequence of information bits is mapped into a continuous sequence of encoder output bits. A convolutional code is generated by passing the information sequence through a finite state shift register. Shift register contains 'N' k-bit stages

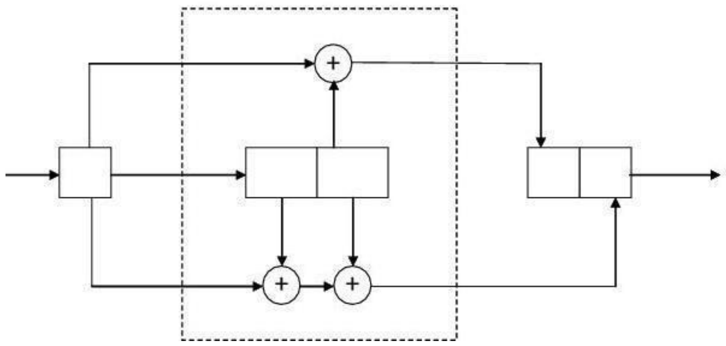


Figure 7.6: A convolutional encoder with n=2 and k=1.

and m linear algebraic function generators based on the generator polynomials. Input data is shifted into and along the shift register, k-bits at a time. Number of output bits for each k-bit user input data sequence is n bits, so the code rate $R_c = k/n$. The shift register of the encoder is initialized to all-zero-state before

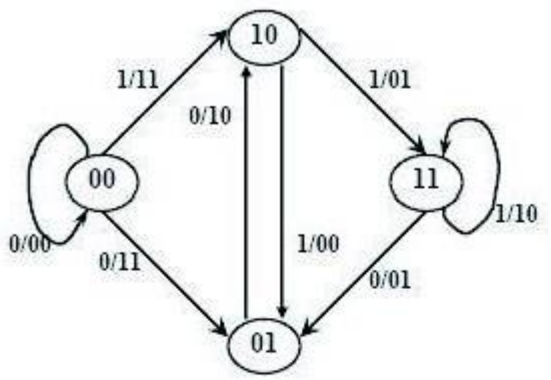


Figure 7.7: State diagram representation of a convolutional encoder.

encoding operation starts. It is easy to verify that encoded sequence is 00 11 10 00 01 . . . for an input message sequence of 01011 .. .. Convolution codes may be represented in various ways as given below.

State Diagram:

Since the output of the encoder is determined by the input and the current state of the encoder, a state diagram can be used to represent the encoding process. The state diagram is simply a graph of the possible states of the encoder and the possible transitions from one state to another. The path information between the states, denoted as b/c_1c_2, represents input information bit 'b' and the corresponding

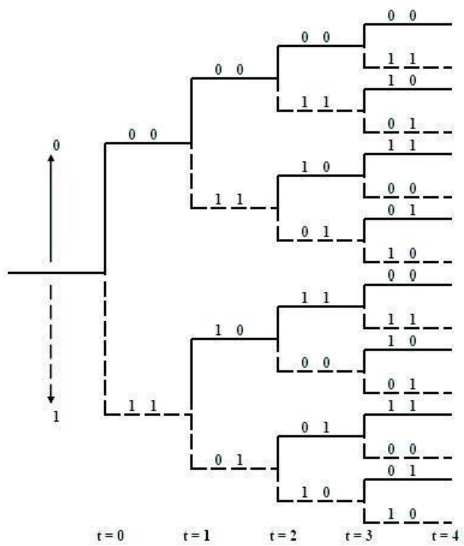


Figure 7.8: Tree diagram representation of a convolutional encoder.

output bits (c_1c_2). Again, it is not difficult to verify from the state diagram that an input information sequence b = (1011) generates an encoded sequence c = (11, 10, 00, 01).

Tree Diagram:

The tree diagram shows the structure of the encoder in the form of a tree with the branches representing the various states and the outputs of the coder. The encoded bits are labeled on the branches of the tree. Given an input sequence, the encoded sequence can be directly read from the tree. As an example, an input sequence (1011) results in the encoded sequence (11, 10, 00, 01).

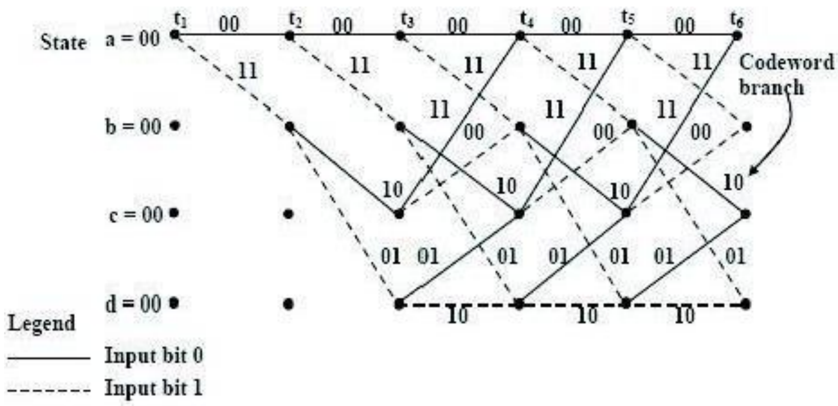


Figure 7.9: Trellis diagram of a convolutional encoder.

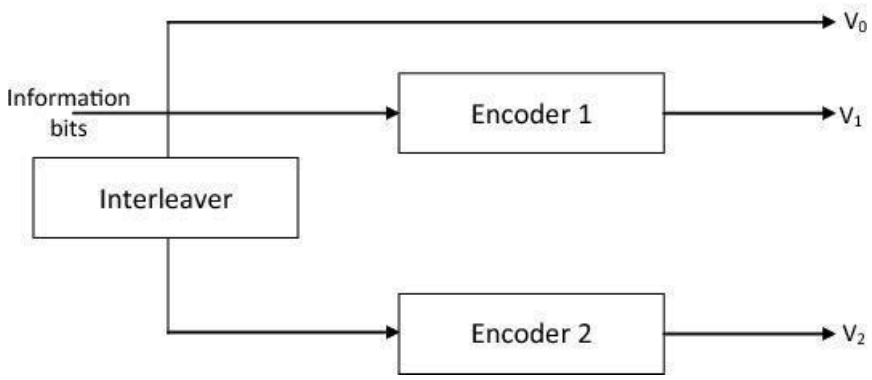


Figure 7.10: Block diagram of a turbo encoder.

Trellis Diagram:

Tree reveals that the structure repeats itself once the number of stages is greater than the constraint length. It is observed that all branches emanating from two nodes having the same state are identical in the sense that they generate identical output sequences. This means that the two nodes having the same label can be merged. By doing this throughout the tree diagram, we obtain another diagram called a Trellis Diagram which is more compact representation.

Concatenated Codes

Concatenated codes are basically concatenation of block and convolutional codes. It can be of two types: serial and parallel codes. Below, we discuss a popular parallel concatenated code, namely, turbo code.

Turbo Codes: A turbo encoder is built using two identical convolutional codes of special type with parallel concatenation. An individual encoder is termed a com- ponent encoder. An interleaver separates the two component encoders. The inter- leaver is a device that permutes the data sequence in some predetermined manner. Only one of the systematic outputs from the two component encoders is used to form a codeword, as the systematic output from the other component encoder is only a permuted version of the chosen systematic output. Figure 7.10 shows the block di- agram of a turbo encoder using two identical encoders. The first encoder outputs the systematic V_0 and recursive convolutional V_1 sequences while the second en- coder discards its systematic sequence and only outputs the recursive convolutional V_2 sequence. Depending on the number of input bits to a component encoder it may be binary or m-binary encoder. Encoders are also categorized as systematic or non-systematic. If the component encoders are not identical then it is called an asymmetric turbo code.

Conclusion

Although a lot of advanced powerful techniques for mitigating the fading effects such as space diversity in MIMO systems, space-time block coding scheme, MIMO equalization, BLAST architectures etc. have taken place in modern wireless com- munication, nevertheless, the discussed topics in this chapter are the basic building blocks for all such techniques and that stems the necessity for all these discussions. The effectiveness of the discussed topics would be more clear in the next chapter in the context of different multiple access techniques.

References

1. T. S. Rappaport, *Wireless Communications: Principles and Practice,* 2nd ed. Singapore: Pearson Education, Inc., 2002.
2. J. R. Treichler, C. R. Johnson (Jr.) and M. G. Larimore, *Theory and Design of Adaptive Filters.* New Delhi: PHI, 2002.
3. S. Gravano, *Introduction to Error Control Codes.* NY: Oxford University Press, 2001.

UNIT -V
WIRELESS NETWORKS

Specific Instructional Objectives

On completion, the student will be able to:

- Explain the need for wireless LAN
- Identify the limitations and challenges of wireless LAN
- Understand different aspects of IEEE 802.11 WLAN
  - Transmission media
  - Topology
  - Medium Access Control
  - Security

Introduction

In the last two decades the wired version of LAN has gained wide popularity and large- scale deployment. The IEEE 802.3 standard has been revised and extended every few years. High-speed versions with transmission rate as high as 1000 Mbps are currently available. Until recently wireless version of LANs were not popular because of the following reasons:

- **High cost**: Previously the equipments cost more.
- **Low data rate**: Initially, the data rate supported by the WLAN is too less, so it supports only a few applications.
- **Occupational safety concerns**
- **Licensing requirements**

In the last couple of years the situation has changed significantly. Cheaper, smaller and powerful notebook computers and other mobile computing equipment have proliferated in homes and offices. These devices share various resources such as printers, files and Broadband Internet connections. This has opened up the need for wireless LAN. Wireless LANs also offer a number of other advantages compared to their wired counterpart.

Before going into the technical details of Wireless LAN let us first look at various reasons which have led to the development of WLANs. Some of the advantages are mentioned below:

- **Availability of low-cost portable equipments:** Due to the technology enhancements, the equipment cost that are required for WLAN set-up have reduced a lot.
- **Mobility:** An increasing number of LAN users are becoming mobile. These mobile users require that they are connected to the network regardless of where they are because they want simultaneous access to the network. This makes the use of cables, or wired LANs, impractical if not impossible. Wireless LAN can provide users mobility, which is likely to increase productivity, user convenience and various service opportunities.
- **Installation speed and simplicity:** Wireless LANs are very easy to install. There is no requirement for wiring every workstation and every room. This ease of installation makes wireless LANs inherently flexible. If a workstation must be

moved, it can be done easily and without additional wiring, cable drops or reconfiguration of the network.

- **Installation flexibility:** If a company moves to a new location, the wireless system is much easier to move than ripping up all of the cables that a wired system would have snaked throughout the building. This also provides portability. Wireless technology allows network to go anywhere wire cannot reach.
- **Reduced cost of ownership:** While the initial cost of wireless LAN can be higher than the cost of wired LAN hardware, it is envisaged that the overall installation expenses and life cycle costs can be significantly lower. Long-term cost-benefits are greater in dynamic environment requiring frequent moves and changes.
- **Scalability:** Wireless LAN can be configured in a variety of topologies to meet the users need and can be easily scaled to cover a large area with thousands of users roaming within it.

However, wireless LAN technology needs to overcome a number of inherent limitations and challenges. Some of the limitations and challenges are mentioned below:

- Lower reliability due to susceptibility of radio transmission to noise and interference.
- Fluctuation of the strength of the received signal through multiple paths causing fading.
- Vulnerable to eavesdropping leading to security problem.
- Limited data rate because of the use of spread spectrum transmission techniques enforced to ISM band users.

In this lesson we shall introduce the wireless LAN technology based on IEEE 802.11 standard. Its predecessor the IEEE 802.3, commonly referred to as the Ethernet, is the most widely deployed member of the family. IEEE 802.11 is commonly referred to as wireless Ethernet because of its close similarity with the IEEE 802.3. Like IEEE 802.3, it also defines only two bottom levels of ISO's open system Interconnection (OSI) model as shown in Fig. 5.7.1. As it shares the upper layers with other LAN standards, it is relatively easy to bridge the IEEE 802.11 wireless LANs to other IEEE 802.11 wired LANs to form an extended interconnected wired and wireless LAN network. Although initially wireless LANs were perceived to be as a substitute to wired LANs, now it is recognized as an indispensable adjunct to wired LANs.

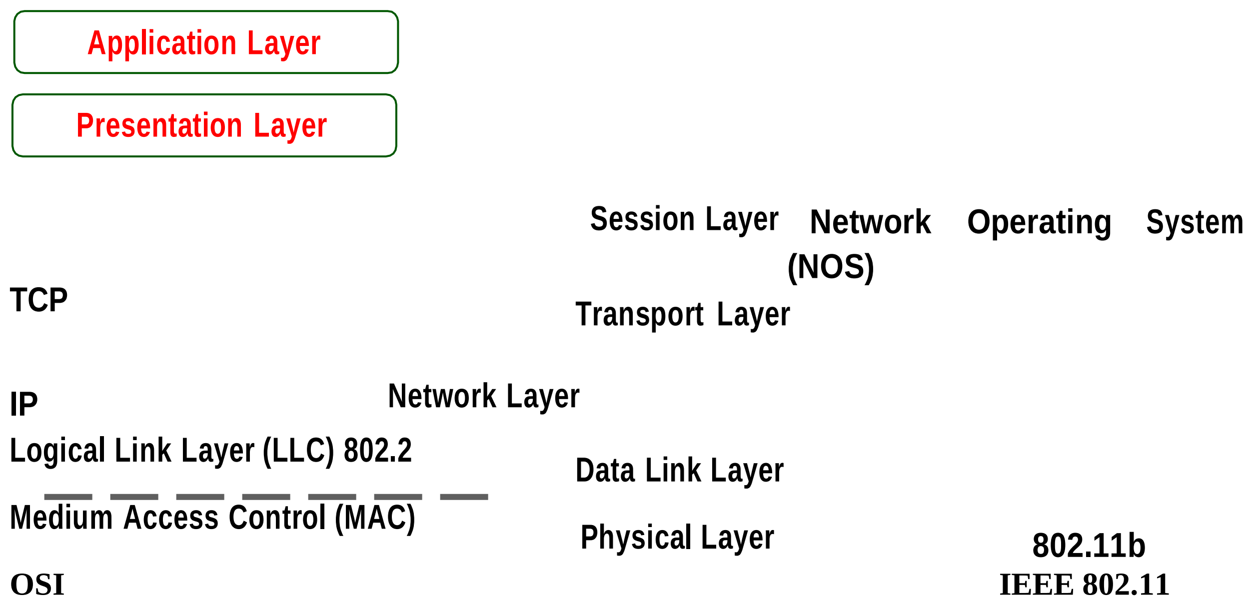


Figure 5.1 OSI Reference Model and IEEE 802.11

The IEEE 802.11 standard basically defines the physical and data link layer. In the later sections we shall look at detailed implementations.

Transmission Media

There are three media that can be used for transmission over wireless LANs. Infrared, radio frequency and microwave. In 1985 the United States released the industrial, scientific, and medical (ISM) frequency bands. These bands are 902 - 928MHz, 2.4 - 2.4853 GHz, and 5.725 - 5.85 GHz and do not require licensing by the Federal Communications Commission (FCC). This prompted most of the wireless LAN products to operate within ISM bands. The FCC did put restrictions on the ISM bands however. In the U.S. radio frequency (RF) systems must implement spread spectrum technology. RF systems must confine the emitted spectrum to a band. RF is also limited to one watt of power. Microwave systems are considered very low power systems and must operate at 500 milliwatts or less.

Infrared

Infrared systems (IR systems) are simple in design and therefore inexpensive. They use the same signal frequencies used on fiber optic links. IR systems detect only the amplitude of the signal and so interference is greatly reduced. These systems are not bandwidth limited and thus can achieve transmission speeds greater than the other systems. Infrared transmission operates in the light spectrum and does not require a license from the FCC to operate. There are two conventional ways to set up an IR LAN.

The infrared transmissions can be **aimed**. This gives a good range of a couple of kilometer and can be used outdoors. It also offers the highest bandwidth and throughput.

The other way is to transmit **omni-directionally** and bounce the signals off of everything in every direction. This reduces coverage to 30 - 60 feet, but it is area coverage. IR technology was initially very popular because it delivered high data rates and relatively cheap price.

The drawbacks to IR systems are that the transmission spectrum is shared with the sun and other things such as fluorescent lights. If there is enough interference from other sources it can render the LAN useless. IR systems require an unobstructed line of sight (LOS). IR signals cannot penetrate opaque objects. This means that walls, dividers, curtains, or even fog can obstruct the signal. InfraLAN is an example of wireless LANs using infrared technology.

Microwave

Microwave (MW) systems operate at less than 500 milliwatts of power in compliance with FCC regulations. MW systems are by far the fewest on the market. They use narrow-band transmission with single frequency modulation and are set up mostly in the 5.8GHz band. The big advantage to MW systems is higher throughput achieved because they do not have the overhead involved with spread spectrum systems. RadioLAN is an example of systems with microwave technology.

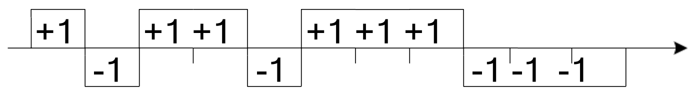
Radio

Radio frequency systems must use spread spectrum technology in the United States. This spread spectrum technology currently comes in two types: direct sequence spread spectrum (DSSS) and frequency hopping spread spectrum (FHSS). There is a lot of overhead involved with spread spectrum and so most of the DSSS and FHSS systems have historically had lower data rates than IR or MW.

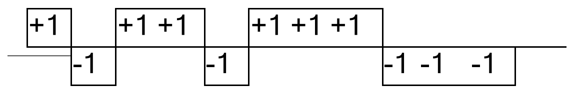
Direct Sequence Spread Spectrum (DSSS) Scheme

Direct Sequence Spread Spectrum (DSSS) represents each bit in the frame by multiple bits in the transmitted frame. DSSS represents each data 0 and 1 by the symbol –1 and +1 and then multiplies each symbol by a binary pattern of +1's and –1's to obtain a digital signal that varies more rapidly occupying larger band. The IEEE 802.11 uses a simple 11- chip Barker sequence B11 [-1, +1, -1, -1, +1, -1, -1, -1, +1, +1, +1] with QPSK or BPSK modulation as shown in Figure 5.7.2. The DSSS transmission system takes 1 Mbps data, converts it into 11 Mbps signal using differential binary phase shift keying (DBPSK) modulation.

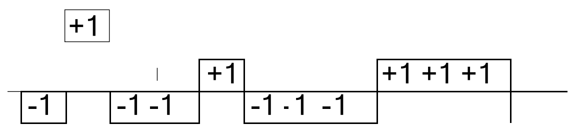
The Barker sequence provides good immunity against interference and noise as well as some protection against multi-path propagation. In both cases of spread spectrum transmission, the signal look like noise to any receiver that does not know the pseudorandom sequence. The third transmission media is based on infrared signal in the near visible range of 850 to 950 nanometers. Diffused transmission is used so that the transmitter and receivers do not have to point to each other and do not require a clear line of sight communication. The transmission distance is limited to 10 to 20 meters and is limited to inside the buildings only.



(a) 11-chip Barker sequence



(b) Transmission of -1



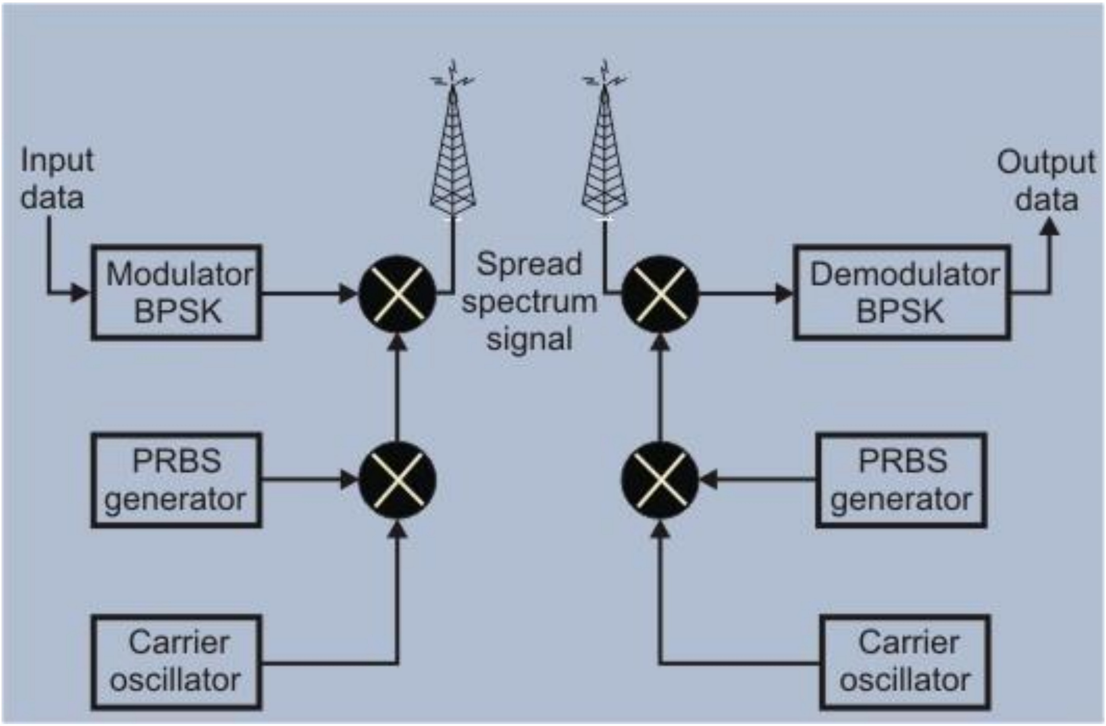
(c) Transmission of +1

Figure 5.2 Direct-sequence spread spectrum technique using Barker sequence

With direct sequence spread spectrum the transmission signal is spread over an allowed band (for example 25MHz). A random binary string is used to modulate the transmitted signal. This random string is called the ***spreading code***. The data bits are mapped to into a pattern of "chips" and mapped back into a bit at the destination. The number of chips that represent a bit is the ***spreading ratio***. The higher the spreading ratio, the more the signal is resistant to interference. The lower the spreading ratio, the more bandwidth is available to the user. The FCC dictates that the spreading ratio must be more than ten. Most products have a spreading ratio of less than 20 and the new IEEE
standard requires a spreading ratio of eleven. The transmitter and the receiver must be synchronized with the same spreading code. If orthogonal spreading codes are used then more than one LAN can share the same band. However, because DSSS systems use wide sub channels, the number of co-located LANs is limited by the size of those sub channels. Recovery is faster in DSSS systems because of the ability to spread the signal over a wider band. Current DSSS products include Digital's RoamAbout and NCR's WaveLAN.

Figure 5.3 shows a typical DSSS implementation. Here, the data stream and pseudo-random sequence are both converted into analog signals before combining, rather than performing the exclusive-OR of the two streams and than modulating. Eleven channels have been defined to operate in the 2.4 GHz ISM band in US. Channels can operate without interference with each other if their center frequencies are separated by at least 30MHz. The

802.11 DSSS physical layer also defines an option for 2 Mbps operation using Differential Quadrature PSK (DQPSK).



(a) Transmitter (b) Receiver

Figure 5.3 Direct Sequence Spread Spectrum (DSSS) system,

Frequency Hopping Spread Spectrum (FHSS)

The idea behind spread spectrum is to *spread the signal over a wider frequency band*, so as to make jamming and interception more difficult and to minimize the effect of interference from other devices In FH it is done by transmitting the signal over a random sequence of frequencies; that is, first transmitting at one frequency, then second, then a third and so on. The random sequence of frequencies is generated with the help of a pseudorandom number generator. As both the receiver and sender use the same algorithm to generate random sequence, both the devices hop frequencies in a synchronous manner and frames transmitted by the sender are received correctly by the receiver. This is somewhat similar to sending different parts of one song over several FM channels. Eavesdroppers hear only unintelligible blips and any attempt to jam the signal results in damaging a few bits only.

Typical block diagram of a frequency-hopping system is shown in Figure 5.7.4.

As shown in Figure 5.4(a) the digital data is first encoded to analog signal, such as frequency-shift keying (FSK) or Binary-phase shift keying (BPSK). At any particular instant, a carrier frequency is selected by the pseudo-random sequence. The carrier frequency is modulated by the encoder output a then transmitted after band pass filtering At the receiving end, the spread-spectrum signal is demodulated using the same sequence of carrier frequencies generated with the help of same pseudo-random sequence in synchronization with the transmitter, and the demodulated signal filtered using a band- pass filter before decoding as shown in Fig. 5.4(b).

This technique splits the band into many small sub channels (each of 1MHz). The signal then hops from sub channel to sub channel transmitting short bursts of data on each channel for a set period of time, called ***dwell time***. The hopping sequence must be synchronized at the sender and the receiver or information is lost.

The 802.11 frequency hopping physical layer uses 79 non-overlapping 1 MHz Channels to transmit 1 Mbps data signal over 2.4 GHz ISM band. There is option to transmit at the rate of 2 Mbps. A channel hop occurs every 224 μsec. The standard defines 78 hopping patterns that are divided into three sets of 26 patterns each. Each hopping pattern jumps a minimum of six channels in each hop and the hopping sequences are derived via a simple modulo 79 calculation. The hopping patterns from each set collide three times on the average and five times in the worst case over a hopping cycle. Each 802.11 network must use a particular hopping pattern. The hopping patterns allow up to 26 networks to be collocated and still operate simultaneously.

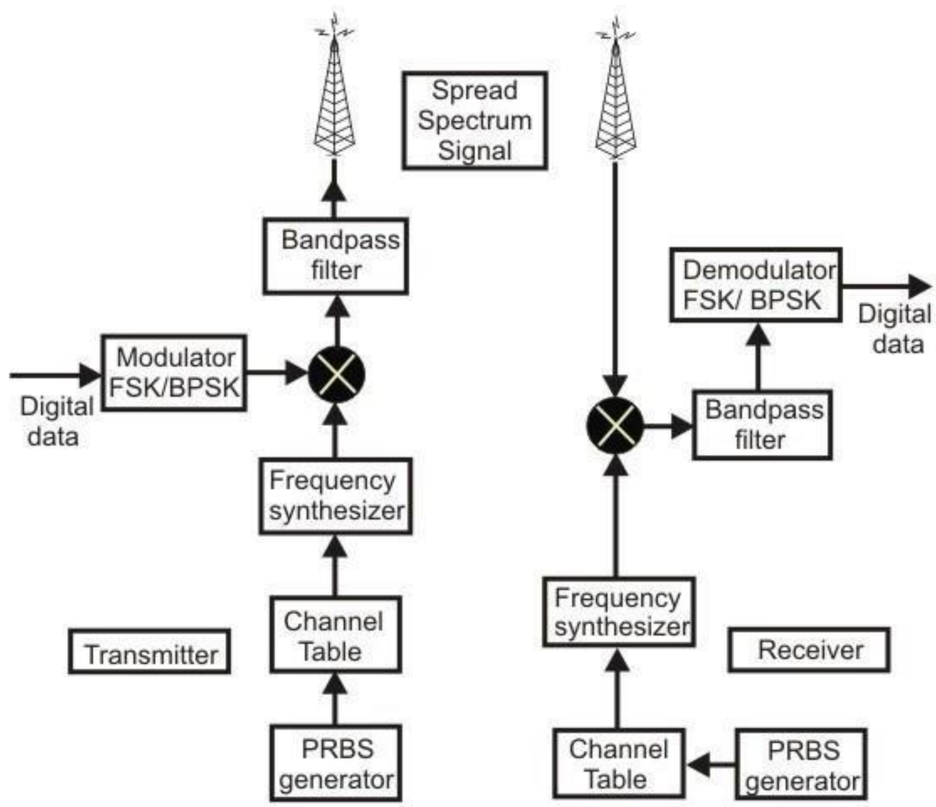


Figure 5.4 Frequency Hopping system, (a) Transmitter (b) Receiver

This feature gives FH systems a *high degree of security*. In order to jam a frequency hopping system the whole band must be jammed. These features are very attractive to agencies involved with law enforcement or the military. Many FHSS LANs can be co-located if an orthogonal hopping sequence is used. Because the sub channels are smaller than in DSSS, the number of co-located LANs can be greater with FHSS systems. Most new products in wireless LAN technology are currently being developed with FHSS technology. Some examples are WaveAccess's Jaguar, Proxim RangeLAN2, and BreezeCom's BreezeNet Pro.

Multipath Interference

Interference caused by signals bouncing off of walls and other barriers and arriving at the receiver at different times is called ***multipath interference***. Multipath interference affects IR, RF, and MW systems. FHSS inherently solves the multipath problem by simply hopping to other frequencies. Other systems use anti-multipath algorithms to avoid this interference. A subset of multipath is Rayleigh fading. This occurs when the difference in path length is arriving from different directions and is a multiple of half the wavelength. Rayleigh fading has the effect of completely cancelling out the signal. IR systems are not affected by Rayleigh fading, because the wavelengths used in IR are very small.

Topology

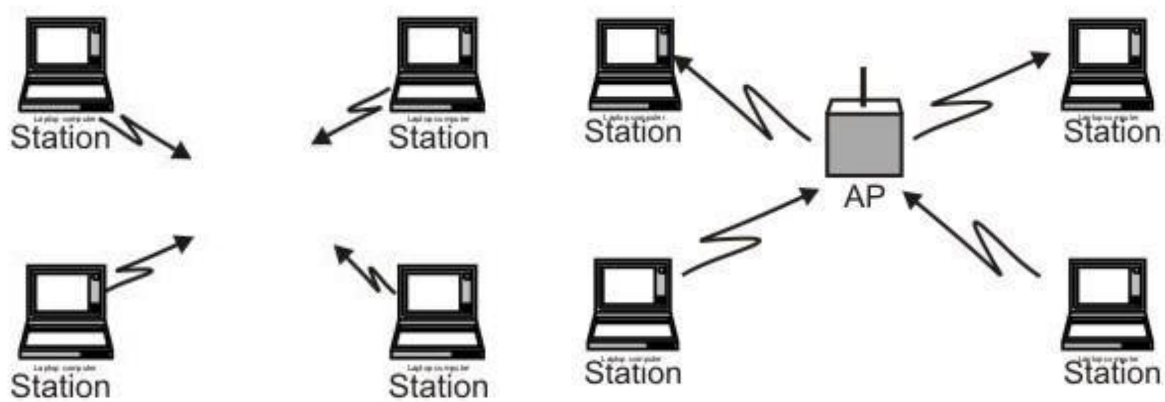
Each computer, mobile, portable or fixed, is referred to as a *station* in 802.11. The difference between a portable and mobile station is that a portable station moves from point to point but is only used at a fixed point. Mobile stations access the LAN during movement. Fundamental to the IEEE 802.11 architecture is the concept of *Basic Service Set (BSS) or wireless LAN cell*. A **BSS** is defined as a group of stations that coordinate their access to the medium under a given instance of medium access control. The geographic area covered by a BSS is known as the *Basic Service Area (BSA)*, which is very similar to a cell in a cellular communication network. All stations with in a BSA with tens of meters in diameter may communicate with each other directly. The 802.11 standard support the formation of two distinct types of BSSs: ad hoc network and Infrastructure BSS.

Two or more BSS's are interconnected using a *Distribution System or DS*. This concept of DS increases network coverage. Each BSS becomes a component of an extended, larger network. Entry to the DS is accomplished with the use of *Access Points (AP)*. An access point is a station, thus addressable. So data moves between the BSS and the DS with the help of these access points.

Creating large and complex networks using BSS's and DS's leads us to the next level of hierarchy, the *Extended Service Set or ESS*. The beauty of the ESS is the entire network looks like an independent basic service set to the Logical Link Control layer (LLC). This means that stations within the ESS can communicate or even move between BSS's transparently to the LLC.

The first type of BSS is known as *ad hoc network*, which consists of a group of stations within the range of each other. As its name implies, ad hoc networks are temporary in nature, which are typically created and maintained as needed without prior administrative arrangement. Ad hoc networks can be formed anywhere spontaneously and can be disbanded after a limited period of time. A typical ad hoc network is shown in Figure 5.5(a).

The second type of BSS is known as *infrastructure BSS (IBSS),* which is commonly used in practice. An ESS is shown in Fig. 5.7.6 Here, several BSSs are interconnected by a distribution system to form an extended service set (ESS) as shown in Fig. 5.5(b). The BSSs are like cells in a cellular communications network. Each BSS is provided with an Access point (AP) that has station functionality and provides access to the distribution system. APs operate on a fixed channel and remain stationary like *base stations* in a cellular communication system. APs are located such that the BSSs they serve overlap slightly to provide continuous service to all the stations.



(b)

Figure 5.5 (a) Basic Service set (BSS), (b) Infrastructure BSS (ESS)

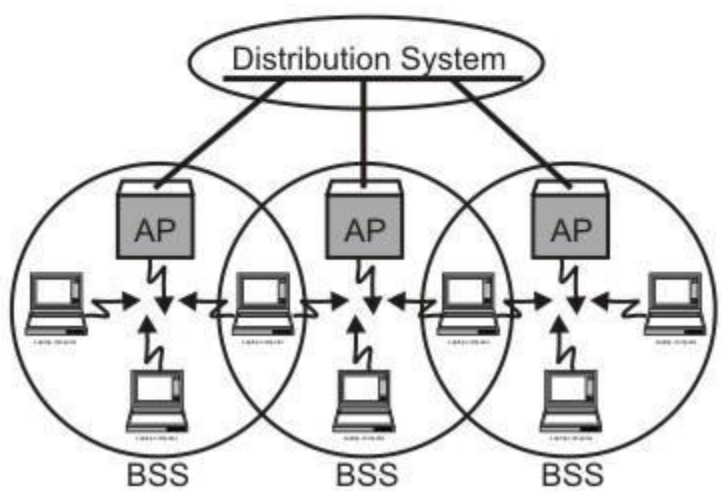


Figure 5.6 Extended service set (ESS)

An ESS can also provide gateway access for wireless users into a wired network. Each end station associates itself with one access point. Figure 5.7.6 shows three BSSs interconnected through three APs to a distribution system. If station A associated with AP-1 wants to send a frame to another station associated with AP-2, the first sends a frame to its access point (AP-1), which forwards the frame across the distribution system to the access point AP-2. AP-2 finally delivers it to the destination station. For forwarding frames across the APs, bridging protocol may be used, which is beyond the scope of IEEE 802.11 standard. However, the 802.11 standard specifies how stations select their access points. The technique used for this purpose is known as *scanning*, which involves the following steps:

- A station sends a *probe frame*.
- All APs within reach reply with a *probe response frame*.
- The station selects one of the access points, and sends the AP an *Association Request frame*.
- The AP replies with an *Association Response frame*.

The above protocol is used when a station joins a network or when it wants to discontinue association with the existing AP because of weakened signal strength or some other reason. The discontinuation of association takes place whenever a station acquires a new AP and the new AP announces it in step 4 mentioned above. For example, assume that station B is moving away from the BSS of AP-1 towards the BSS of AP-2. As it moves closer to the BSS of AP-2, it sends probe frames, which is responded eventually by AP-2. As some of point of time station B prefers AP-2 over AP-1 and associates itself with the access point AP-2. The above mechanism is known as *active scanning*, as the node is actively searching for an access point. An access point also periodically sends Beacon frame that advertises the capabilities of the access point. In response, a station can associate to the AP simply by sending it an Association request frame. This is known as *passive scanning*.

Medium Access Control

Most wired LANs products use Carrier Sense Multiple Access with Collision Detection (CSMA/CD) as the MAC protocol. Carrier Sense means that the station will listen before it transmits. If there is already someone transmitting, then the station waits and tries again later. If no one is transmitting then the station goes ahead and sends what it has. But when more than one station tries to transmit, the transmissions will collide and the information will be lost. This is where Collision Detection comes into play. The station will listen to ensure that its transmission made it to the destination without collisions. If a collision occurred then the stations wait and try again later. The time the station waits is determined by the back off algorithm. This technique works great for wired LANs but wireless topologies can create a problem for CSMA/CD. However, the wireless medium presents some unique challenges not present in wired LANs that must be dealt with by the MAC used for IEEE 802.11.

Some of the challenges are:

- The wireless LAN is prone to more interference and is less reliable.
- The wireless LAN is susceptible to unwanted interception leading to security problems.
- There are so called *hidden station* and *exposed station* problems.

In the discussion of both the problem, we shall assume that all radio transmitters have fixed range. When the receiver is in the range of two active transmitters then the signal will be garbled. It is important to note that not all stations are in range of two transmitters.

The Hidden Station Problem

Consider a situation when A is transmitting to B, as depicted in the Fig. 5.7. If C senses the media, it will not hear anything because it is out of range, and thus will falsely conclude that no transmission is going on and will start transmit to B. the transmission will interfere at B, wiping out the frame from A. The problem of a station not been able to detect a potential competitor for the medium because the competitor is too far away is referred as *Hidden Station Problem.* As in the described scenario C act as a hidden station to A, which is also competing for the medium.

Error!

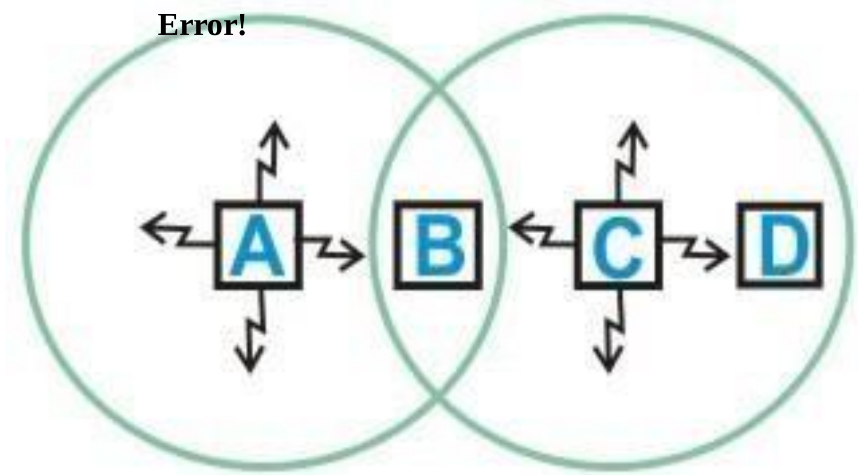


Figure 5.7 Hidden Station Problem

Exposed Station problem

Now consider a different situation where B is transmitting to A, and C sense the medium and detects the ongoing transmission between B and A. C falsely conclude that it can not transmit to D, when the fact is that such transmission would cause on problem. A transmission could cause a problem only when the destination is in zone between B and

C. This problem is referred as *Exposed station Problem.* In this scenario as B is exposed to C, that's why C assumes it cannot transmit to D. So this problem is known as *Exposed station problem* (i.e. problem caused due to exposing of a station). The problem here is that before transmission, a station really wants to know that whether or not there is any activity around the receiver. CSMA merely tells whether or not there is any activity around the station sensing the carrier.

Carrier Sense Multiple Access with Collision Avoidance (CSMA-CA)

Main steps can be summarized as:

- Sender sends a short frame called *Request to send* RTS (20bytes) to the destination. RTS also contains the length of the data frame.
- Destination station responds with a short (14 bytes) *clear to send* (CTS) frame.
- After receiving the CTS, the sender starts sending the data frame.
- If collision occurs, CTS frame is not received within a certain period of time.

CSMA/CA works as follows: the station listens before it sends. If someone is already transmitting, wait for a random period and try again. If no one is transmitting then it sends a short message. This message is called the *Ready To Send* message *(RTS)*. This message contains the destination address and the duration of the transmission.
Other stations now know that they must wait that long before they can transmit. The destination then sends a short message, which is the *Clear To Send message (CTS)*. This message tells the source that it can send without fear of collisions. Each packet is acknowledged. If an acknowledgement is not received, the MAC layer retransmits the data. This entire sequence is called the 4- way handshake protocol.

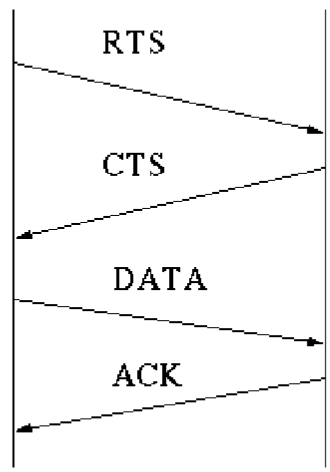


Figure 5.8 Four-Way handshake protocol

Carrier Sensing

In IEEE 802.11, carrier sensing is performed in two levels known as *physical carrier sensing* and *virtual carrier sensing*.

Physical carrier sensing is performed at the radio interface by sensing the presence of other IEEE 802.11 stations by analyzing all detected packets and relative strength from other sources.

Virtual carrier sensing is used by a source station to inform all other stations in the BSS about the length of the data frame that it intends to send. The headers of the RTS and CTS control frames contain the duration field (in µsec). Stations detecting a duration field adjust their Network Allocation Vector (NAV), which indicates the duration the station must wait before channel can be sampled again for sensing status of the medium. The protocol may be considered as a 4-way handshake protocol is shown in Figure 6.39.

The above protocol known as *Multiple Access Carrier Avoidance (MACA)* was subsequently extended to improve its performance and the new protocol, with the following three additions, was renamed as *MACAW*. First, the receiver sends an ACK frame after receiving a frame and all stations must wait for this ACK frame before trying to transmit. Second, the back-off algorithm is to run separately for each data stream, rather than for each station. This change improves the fairness of the protocol. Finally, some mechanism was added for stations to exchange information about configuration, and way to make the back-off algorithm react less violently to temporary problem.

The IEEE 802.11 protocol is specified in terms of coordination function that determine when a station in a BSS is allowed to transmit and when it may be able to receive data over the wireless medium. The distributed coordination function (DCF) provides support for asynchronous data transfer on a best-effort basis. Four following types of inter frame spaces (IFSs) are used:

- Short IFS (SIFS): This is the period between the completion of packet transmission and the start of ACK frame.
- Point coordination IFS (PIFS): This is SIFS plus a slot time.
- Distributed IFS (DIFS): This PIFS Plus a slot time.
- Extended IFS (EIFS): This is longer than IFS used by a station that has received a packet that it could not understand. This is needed to prevent collisions. The sequence of events that take place at the source, destination and other stations is shown in Figure 5.7.9.

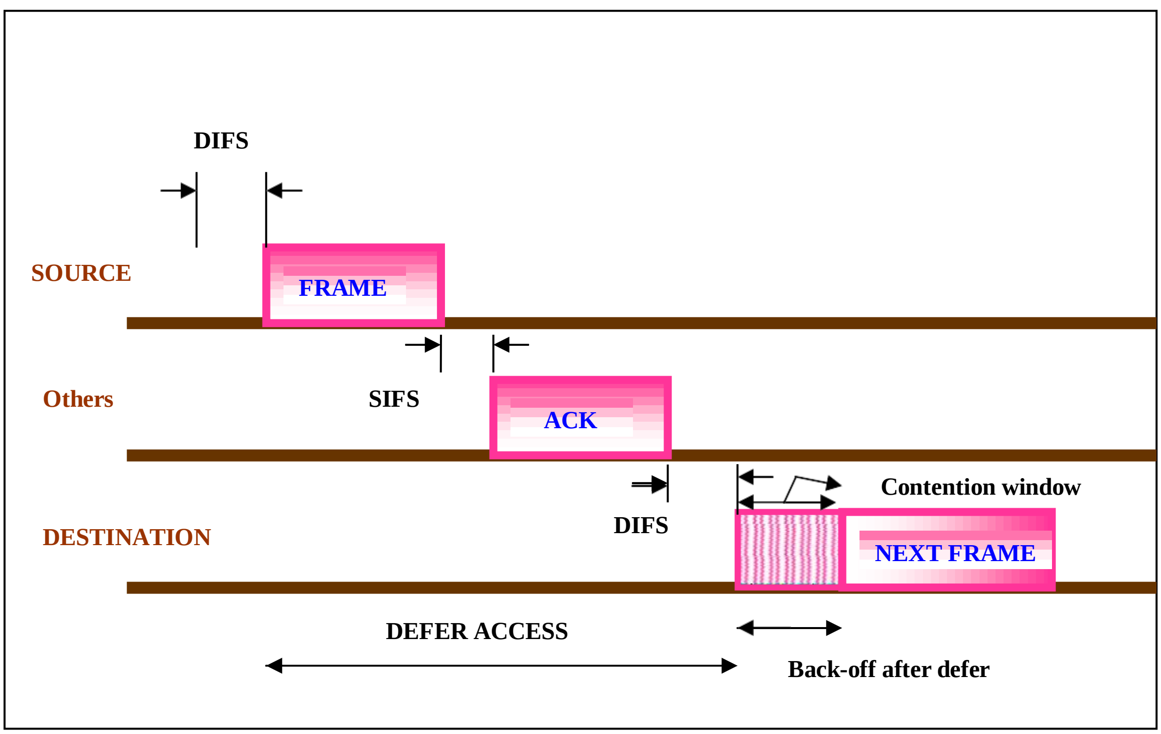


Figure 5.9 CSMA/CA Back-off algorithm timing sequence

Framing

The frame format of the IEEE 802.11 is shown in Figure 5.7.10(a). The frames can be categorized into three types; management frame, control frame and data frame. The management frames are used for association and disassociation of stations with at the AP, authentication and de-authentication, and timing and synchronization. The detailed Frame Format is shown in Fig. 5.10.

MAC HEADER

2	2	6	6	6	2	6	0-2312	4
Frame control	nation / ID	ddr. 1	ddr. 2	ddr. 3	equence Control	ddr. 4	rame body	CRC

(a)

Protocol version	Type	Subtype	To DS	rom DS	More Frag	Retry	Pwr. mgt.	More data	WEP	Rsvd

(b)

To DS	rom DS	ddr. 1	ddr. 2	ddr. 3	ddr. 4	Meaning
0	0	Desti. Addr.	ource Addr.	SS ID	N/A	Data from station to station within a BSS
0	1	Desti. Addr.	SS ID	ource Addr.	N/A	e existing in DS
1	0	SS ID	ource Addr.	Desti. Addr.	N/A	frame destined for the DS
1	1	eceiver Addr.	Trans Addr.	Desti. Addr.	ource Addr.	WDS frame being distributed from AP to AP

(c)

Figure 5.10 Frame format for 802.11

Each frame consists of a MAC header, a frame body and a frame check sequence (FCS). The basic frame can be seen in Figure 5.11 below.

MAC Header

Frame Control	Duration ID	Address 1	Address 2	Address 3	Sequence Control	Address 4	Frame Body	FCS

2 **2** **6** **6** **6** **2** **6** **0-2312** **4**

Bytes

Figure 5.11 802.11 Frame (also shown in 5.10(a))

AC header will be described in a little while. Frame Body varies from 0-2312 bytes. At last is the FCS field. The *frame check sequence* is a 32-bit cyclic redundancy check which ensures there are no errors in the frame. For the standard generator polynomial see IEEE P802.11.

The MAC header consists of seven fields and is 30 bytes long. The fields are frame control, duration, address 1, address 2, address 3, sequence control, and address 4. The frame control field is 2 bytes long and is comprised of 11 subfields as shown in Fig.
below.

Protocol Version	Type	Subtype	To DS	From DS	More Frag	Retry	Power Mgmt.	More Data	WEP	Order

2 **2** **4** **1** **1** **1** **1** **1** **1** **1** **1**

Bits

Figure 5.12 802.11 MAC Header

Frame Control Field (in MAC header)

- The protocol version field is 2 bits in length and will carry the version of the 802.11 standard. The initial value of 802.11 is 0; all other bit values are reserved.

- **Type** and **subtype** fields are 2 and 4 bits, respectively. They work together hierarchically to determine the function of the frame.
- The remaining 8 fields are all 1 bit in length.
- The **To DS** field is set to 1 if the frame is destined for the distribution system.
- **From DS** field is set to 1 when frames exit the distribution system. Note that frames which stay within their basic service set have both of these fields set to 0.

- The More **Frag field** is set to 1 if there is a following fragment of the current MSDU.
- **Retry** is set to 1 if this frame is a retransmission.
- **Power Management** field indicates if a station is in power save mode (set to 1) or active (set to 0).
- **More data** field is set to 1 if there is any MSDUs are buffered for that station.
- The **WEP** field is set to 1 if the information in the frame body was processed with the WEP algorithm.
- The **Order** field is set to 1 if the frames must be strictly ordered.
- **The Duration/ID field** is 2 bytes long. It contains the data on the duration value for each field and for control frames it carries the associated identity of the transmitting station.
- The **address fields** identify the basic service set, the destination address, the source address, and the receiver and transmitter addresses. Each address field is 6 bytes long.
- The **sequence control field** is 2 bytes and is split into 2 subfields, fragment number and sequence number.
- **Fragment number** is 4 bits and tells how many fragments the MSDU is broken into.
- The **sequence number field** is 12 bits that indicates the sequence number of the MSDU. The frame body is a variable length field from 0 - 2312. This is the payload.

Security

Wireless LANs are subjected to possible breaches from unwanted monitoring. To overcome this problem, IEEE 802.11 specifies an optional MAC layer security system known as *Wired Equivalent Privacy* (WEP). The objective is to provide a level of privacy to the wireless LAN similar to that enjoyed by wired Ethernets. It is achieved with the help of a 40-bit shared key authentication service. By default each BSS supports up to four 40-bit keys that are shared by all the clients in the BSS. Keys unique to a pair of communicating clients and direction of transmission may also be used. Advanced Encryption Standard (AES) (802.11i) for authentication and encryption is recommended as a long-term solution.

IEEE 802.11 extensions

As the first standard was wrapping up, the creation of a new standards activity begun in the 802.11 standards body. The new activity gave rise to two more standards; IEEE
b and IEEE 802.11a.

- **802.11b:** This standard was developed by IEEE with the support from the consortium Wireless Ethernet Compatibility Alliance (WECA). This standard is backward compatible with the original standard that added two new data rates 5.5 mbps and 11 Mbps using two coding techniques; the mandatory coding mode known as Complementary Coding Keying (CCK) modulation and Packet Binary Convolution Coding (PBCC). Because of backward compatibility with the 802.11, this standard has gained wide popularity with millions of installed base, which is growing rapidly.

- **802.11a:** The successor to 802.11b is 802.11a with greater speed and at a different frequency. It operates at radio frequencies between 5 GHz incorporating a coded multi-carrier scheme known as Orthogonal Frequency Division Multi-carrier (OFDM). The 5 GHz band is currently unlicensed and less congested than the 2.4 GHz ISM band. The 802.11a specifies data speed as high as 54 mbps, also supports 6, 12, 24, and 34 mbps. There is trade off between bandwidth and range - lower bandwidth cases offering increases range. For 54 mbps, the typical range is 20-30 meters. The 802.11a and 802.11b devices can coexist without interference or reduced performance.

- **802.11g:** The success of 802.11b has led to another extension that provides 22 Mbps transmission. It retains backward compatibility with the popular 802.11b standard. This standard will become 802.11g.

Upper Layers				
802.11 FHSS	802.11 DSSS	802.11a OFDM	802.11b HR-DSSS	802.11g OFDM

WiFi: Any of the above wireless LAN standards are referred to by the brand name **"WiFi"**. It essentially denotes a set of Wireless LAN standards developed by the working group 11 of the IEEE LAN/MAN Standards Committee (IEEE 802).

WiMAX: The story of wireless LAN cannot be complete without the mention of WiMAX, which stands for **W**orldwide **I**nteroperability for **M**icrowave **Acc**ess by the WiMAX Forum. The forum was formed in June 2001 to promote conformance and interoperability of the IEEE 802.16 standard, officially known as Wireless (Metropoliton Area Network) MAN. The Forum describes WiMAX as "a standards-based technology enabling the delivery of last mile wireless broadband access as an alternative to cable and DSL". It supports point to multi-point (PMP) broadband wireless access. WiMAX can deliver a maximum of 70 Mbit/s, over a maximum distance of 70 miles (112.6 kilometers). It has some similarities to DSL in this respect, where one can either have high bandwidth or long range, but not both simultaneously. The other feature to consider with WiMAX is that available bandwidth is shared between users in a given radio sector, so if there are many active users in a single sector, each will get reduced bandwidth.

1. **Development of IEEE 802.11**

The Physical layer (PHY) and medium access control (MAC) layer were mainly targeted by the IEEE 802 project. When the idea of wireless local area network (WLAN) was first conceived, it was just thought of another PHY of one of the available standards. The first candidate which was considered for this was IEEE's most prominent standard 802.3.

However later findings showed that the radio medium behaved quite different than the conventional well behaved wire. As there was attenuation even over short distances, collisions could not be detected. Hence, 802.3's carrier sense multiple access with collision detection (CSMA/CD) could not be applied.

The next candidate standard considered was

802.4. At that point of time, its coordinated medium access i.e. the token bus concept was believed to be superior to 802.3's contention-based scheme. Hence, WLAN began as 802.4L. Later in 1990 it became obvious that token handling in radio networks was rather difficult. The standardization body realized the need of a wireless communication standard that would have its own very unique MAC. Finally, on March 21, 1991, the project 802.11 was approved (fig. 1).

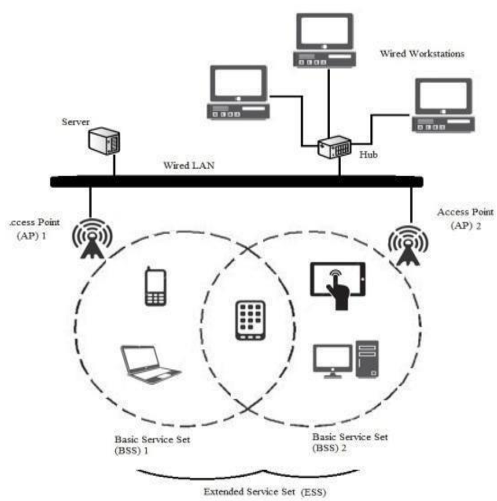


Figure 1 WLAN Network Architecture

2. **IEEE 802.11 family**

The most widely deployed 802.11 standard has a lot of extension and many more are currently under development. First introduced in 1999,the IEEE 802.11 standards was primarily developed keeping in mind the home and the office environment for wireless local area connectivity. The Initial standards gave a maximum data rate of 2Mbps per AP which increased to 11 Mbps per AP with the deployment of IEEE 802.11b.Newer extensions like IEEE 802.11g and IEEE 802.11a provided maximum data rate of 54Mbps per AP using various methods to boost up the maximum data rates. WLAN devices based on IEEE 802.11g currently offer data rate 100-125Mbps. Similarly, a relatively newer IEEE 802.11n gives a maximum data rate of about 540Mbps .Furthermore, in addition to these, several other standards were deployed which solved many QoS and security issues related with the earlier standards. Additional mechanisms were introduced to remedy QoS support and security problems in IEEE 802.11e [12] and IEEE 802.11i.The IEEE 802.11n standard which we earlier talked about also introduced MAC enhancements to overcome MAC layer limitations in the current standards. The IEEE 802.11s standard added mesh topology support to the IEEE 802.11. The IEEE 802.11u improved internetworking with external non-802.11 networks. The IEEE 802.11w was an added onto 802.11i covering management frame security.

The IEEE 802.11ad standard adds a "fast session transfer" feature, enabling the wireless devices to seamlessly make transition between the legacy 2.4 GHz and 5 GHz bands and the 60 GHz frequency band. The IEEE 802.11ac standard, still under development is expected to provide a multi-station WLAN throughput of at least 1 Gbps and a single link throughput of at least 500 Mbps.

The IEEE 802.11a extension employs a number of channels ranging from 36-161 depending on the frequency band (5.15-5.825 GHz) although it works with a fixed channel centre frequency of 5 GHz. There are 12 non overlapping channels in the frequency band for the IEEE standard in the U.S. and 19 non-overlapping channels in Europe. In contrast, there are only 3 out 14 non-overlapping in case of 802.11b [2]. IEEE 802.11n uses overlapping channels with channel bandwidth 20 and 40MHz [19]. The 20MHz channel bandwidth is incorporated in every 802.11n device, the 40MHz channel is optional.

Peer to Peer (P2P) WLAN links can be established with the help of directional antennas for a few km ranges. A typical WLAN Access Point (AP) uses omnidirectional antennas with a range of 30-50m (indoors) and 100m (outdoors).This range is greatly affected by the obstacles between the AP and the STA.IEEE 802.11a suffer from increased range and attenuation compared to IEEE 802.11b/g networks, because it operates on the higher frequency range of 5MHz.Use of sectored antennas instead of omnidirectional antennas increases the aggregate WLAN data rate in a given area to 2-3 times .

Medium Access Control (MAC) Layer

IEEE 802.11 uses a contention based scheme known as Distributed Coordinated Function (DCF). In this method the STA linked with the AP scans the air interface for channel availability. If the interface is idle, the STA sends it data to the destination through the AP. If however the air interface is busy or more than one STA tries to access the same AP; a collision occurs. The IEEE
802.11 uses a Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) to avoid the collisions. IEEE 802.11 uses another MAC technique known as Point Coordination Function (PCF) [18]. This mechanism is divided in to two parts. In the first part, the AP scans all its STA in a round robin fashion and checks to see if any of the STAs has any packets to send. If any of the STAs is not polled during the current period, it will be queued up for polling during the next polling period. The scanned part uses the contention based scheme and it same as DCF.

Moreover, due to polling mechanism in PCF the aggregate throughput of an IEEE 802.11 network decreases. DCF is the default MAC technique used in the IEEE 802.11 standard. While the standard includes both the MAC techniques, PCF is included in the Wi-fi alliance standard and hence not quite as popular as DCF [17]. In both the MAC techniques an automatic response request mechanism is used in this method. Any device in the network receiving data will send an acknowledgement signal (ACK) back to the sender. In case the receiver receives a corrupt data packet, it issues a NAK (Negative Acknowledgement) and the sender resends the data packet. There is a round trip delay as the sender has to wait for the ACK to transmit the next data packet in the queue.

Request to Send/Clear to Send (RTS/CTS)

In the contention based scheme called DCF if more than two STAs simultaneously try to access the air interface, a collision occurs. To avoid such collision CSMA/CA may result in incorrect medium information. This is called Hidden Node Problem in which collision in the some part of the network cannot be detected [15]. If any two STAs cannot directly communicate, the AP invokes a RTS/CTS mechanism. For each transmission, the source STA issues a RTS message. The destination STA replies to this by sending a CTS message. Upon receiving the CTS message, the source STA starts its data transmission. The medium is assumed to be in use given in the message when they receive RTS and/or CTS message. In PCF using RTS/CTS reduces the network throughput.

Authentication & Encryption

Security is also handled in the MAC layer. To avoid unauthorized access from other STAs, several encryption methods have been used. One of earlier encryption mechanism was Wired Equivalent Piracy (WEP). But the encryption method had security vulnerabilities and the Wi-fi Alliance developed another encryption technique named Wi-Fi protection Access (WPA).

The IEEE 802.11i standard incorporated an enhanced version of WPA (WPA2) [20]. IEEE 802.11i also addressed security issues associated with authentication methods like open standard and shared key authentication and incorporated IEEE 802.1X authentication method which is now used in all the later versions of IEEE 802 family standards. In this method, users can authenticate their identities by a RADIUS or diameter server.

Management Frame

The current 802.11 standards define "frame" types for use in management and control of wireless links. The TGw implemented the IEEE 802.11w standard to implement the Protected Management Frames. The TGw is still working on improving the IEEE 802.11 MAC layer. Security can be enhanced by providing data confidentiality of management frames. These extensions will have interactions with IEEE 802.11r as well as IEEE 802.11u

Operating Modes

The IEEE 802.11 supports two operating modes. They are namely

1. Infrastructure operating mode and
2. Independent operating mode.

In the infrastructure operating mode, the STAs communicate with each other through the Access Point. In this scheme, an STA needs to be connected to an Access Point in the network in order to talk to another STA (fig.2).

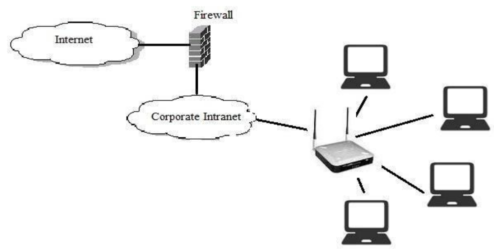


Figure 2 Infrastructure Mode

However, in the independent mode or ad hoc mode, the STAs can directly communicate with each other. In this mode, the source STA forms an ad hoc link directly with the destination STA (fig.3).

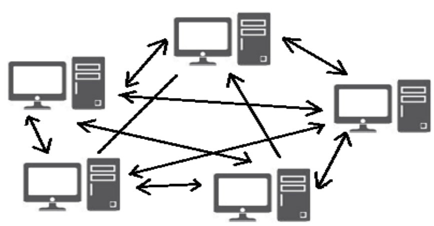


Figure 3 Independent Mode