MOBILE COMMUNICATION

Module 1

Introduction

Communication is one of the integral parts of science that has always been a focus point for exchanging information among parties at locations physically apart. After its discovery, telephones have replaced the telegrams and letters. Similarly, the term `mobile' has completely revolutionized the communication by opening up innovative applications that are limited to one's imagination. Today, mobile communication has become the backbone of the society. All the mobile system technologies have improved the way of living. Its main plus point is that it has privileged a common mass of society. In this chapter, the evolution as well as the fundamental techniques of the mobile communication is discussed. The first wireline telephone system was introduced in the year 1877. Mobile communication systems as early as 1934 were based on Amplitude Modulation (AM) schemes and only certain public organizations maintained such systems. With the demand for newer and better mobile radio communication systems during the World War II and the development of Frequency Modulation (FM) technique by Edwin Armstrong, the mobile radio communication systems began to witness many new changes. Mobile telephone was introduced in the year 1946. However, during its initial three and a half decades it found very less market penetration owing to high costs and numerous technological drawbacks. But with the development of the cellular concept in the 1960s at the Bell Laboratories, mobile communications began to be a promising field of expanse which could serve wider populations. Initially, mobile communication was restricted to certain official users and the cellular concept was never even dreamt of being made commercially available. Moreover, even the growth in the cellular networks was very slow. However, with the development of newer and better technologies starting from the 1970s and with the mobile users now connected to the Public Switched Telephone Network (PSTN), there has been an astronomical growth in the cellular radio and the personal communication systems. Advanced Mobile Phone System (AMPS) was the first U.S. cellular telephone system and it was deployed in 1983. Wireless services have since then been experiencing a 50% per year growth rate. The number of cellular telephone users grew from 25000 in 1984 to around 3 billion in the year 2007 and the demand rate is increasing day by Day.

Mobile Telephony Developement

The first wireline telephone system was introduced in the year 1877. Mobile communication systems as early as 1934 were based on Amplitude Modulation (AM) schemes and only certain public organizations maintained such systems. With the demand for newer and better mobile radio communication systems during the World War II and the development of Frequency Modulation (FM) technique by Edwin Armstrong, the mobile radio communication systems began to witness many new changes. Mobile telephone was introduced in the year 1946. However, during its initial three and a half decades it found very less market penetration owing to high 1 Figure The worldwide mobile subscriber chart. costs and numerous technological drawbacks. But with the development of the cellular concept in the 1960s at the Bell Laboratories, mobile communications began to be a promising field of expanse which could serve wider populations. Initially, mobile communication was restricted to certain official users and the cellular concept was never

even dreamt of being made commercially available. Moreover, even the growth in the cellular networks was very slow. However, with the development of newer and better technologies starting from the 1970s and with the mobile users now connected to the Public Switched Telephone Network (PSTN), there has been an astronomical growth in the cellular radio and the personal communication systems. Advanced Mobile Phone System (AMPS) was the first U.S. cellular telephone system and it was deployed in 1983. Wireless services have since then been experiencing a 50% per year growth rate. The number of cellular telephone users grew from 25000 in 1984 to around 3 billion in the year 2007 and the demand rate is increasing day by
day. A schematic of the subscribers is shown in Fig.

Cellular Concept:

The power of the radio signals transmitted by the BS decay as the signals travel away from it. A minimum amount of signal strength (let us say, x dB) is needed in order to be detected by the MS or mobile sets which may the hand-held personal units or those installed in the vehicles. The region over which the signal strength lies above this threshold value x dB is known as the coverage area of a BS and it must be a circular region, considering the BS to be isotropic radiator. Such a circle, which gives this actual radio coverage, is called the foot print of a cell (in reality, it is amorphous). It might so happen that either there may be an overlap between any two such side by side circles or there might be a gap between the

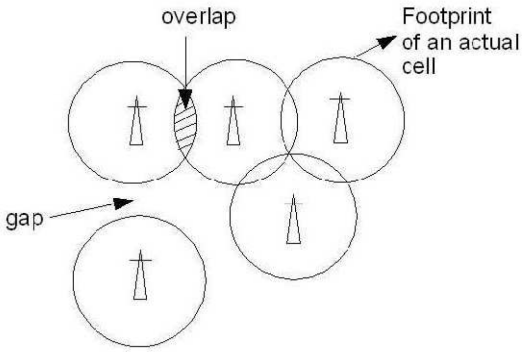


Fig 1:Footprint of cells showing the overlaps and gaps.

coverage areas of two adjacent circles. This is shown in Figure 1. Such a circular geometry, therefore, cannot serve as a regular shape to describe cells. We need a regular shape for cellular design over a territory which can be served by 3 regular polygons, namely, equilateral triangle, square and regular hexagon, which can cover the entire area without any overlap and gaps. Along with its regularity, a cell must be designed such that it is most reliable too, i.e., it supports even the weakest mobile with occurs at the edges of the cell. For any distance between the center and the farthest point in the cell from it, a regular hexagon covers the maximum area. Hence regular hexagonal geometry is used as the cells in mobile communication.

Frequency Reuse:

Frequency reuse, or, frequency planning, is a technique of reusing frequencies and channels within a communication system to improve capacity and spectral efficiency. Frequency reuse is one of the fundamental concepts on which commercial wireless systems are based that involve the partitioning of an RF radiating area into cells. The increased capacity in a commercial wireless network, compared with a network with a single transmitter, comes from the fact that the same radio frequency can be reused in a different area for a completely different transmission. Frequency reuse in mobile cellular systems means that frequencies allocated to

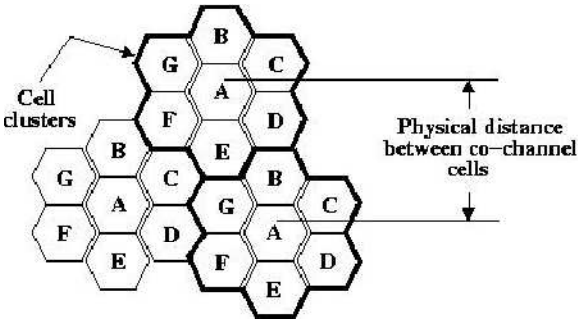


Fig 2: Frequency reuse technique of a cellular system.

the services are reused in a regular pattern of cells, each covered by one base station. The repeating regular pattern of cells is called cluster. Since each cell is designed to use radio frequencies only within its boundaries, the same frequencies can be reused in other cells not far away without interference, in another cluster. Such cells are called `co-channel' cells. The reuse of frequencies enables a cellular system to handle a huge number of calls with a limited number of channels. Figure 3.2 shows a frequency planning with cluster size of 7, showing the co-channels cells in different clusters by the same letter. The closest distance between the co-channel cells (in different clusters) is determined by the choice of the cluster size and the layout of the cell cluster. Consider a cellular system with S duplex channels available for use and let N be the number of cells in a cluster. If each cell is allotted K duplex channels with all being allotted unique and disjoint channel groups we have S = KN under normal circumstances. Now, if the cluster are repeated M times within the total area, the total number of duplex channels, or, the total number of users in the system would be T = MS = KMN. Clearly, if K and N remain constant, then

$T \propto M$ and, if T and K remain constant, then

$$N \propto \frac{1}{M}.$$

Hence the capacity gain achieved is directly proportional to the number of times a cluster is repeated, as shown in (3.1), as well as, for a fixed cell size, small N =25 decreases the size of the cluster with in turn results in the increase of the number of clusters and hence the capacity. However for small N, co-channel cells are located much closer and hence more interference. The value of N is determined by calculating the amount of interference that can be tolerated for a sufficient quality communication. Hence the smallest N having interference below the tolerated limit is used. However, the cluster size N cannot take on any value and is given only by the following equation

$$N = i^2 + ij + j^2, \qquad i \geq 0, j \geq 0,$$

Where i and j are integer numbers.

Channel Assignment Strategies

With the rapid increase in number of mobile users, the mobile service providers had to follow strategies which ensure the effective utilization of the limited radio spectrum. With increased capacity and low interference being the prime objectives, a frequency reuse scheme was helpful in achieving these objectives. A variety of channel assignment strategies have been followed to aid these objectives. Channel assignment strategies are classified into two types: fixed and dynamic, as discussed below.

Fixed Channel Assignment (FCA)

In fixed channel assignment strategy each cell is allocated a fixed number of voice channels. Any communication within the cell can only be made with the designated unused channels of that particular cell. Suppose if all the channels are occupied, then the call is blocked and subscriber has to wait. This is simplest of the channel assignment strategies as it requires very simple circuitry but provides worst channel utilization. Later there was another approach in which the channels were borrowed from adjacent cell if all of its own designated channels were occupied. This was named as borrowing strategy. In such cases the MSC supervises the borrowing process and ensures that none of the calls in progress are interrupted.

Dynamic Channel Assignment (DCA)

In dynamic channel assignment strategy channels are temporarily assigned for use in cells for the duration of the call. Each time a call attempt is made from a cell the corresponding BS requests a channel from MSC. The MSC then allocates a channel to the requesting the BS. After the call is over the channel is returned and kept in a central pool. To avoid co-channel interference any channel that in use in one cell can only be reassigned simultaneously to another cell in the system if the distance between the two cells is larger than minimum reuse distance. When compared to the FCA, DCA has reduced the likelihood of blocking and even increased the trunking capacity of the network as all of the channels are available to all cells, i.e., good quality of service. But this type of assignment strategy results in heavy load on switching center at heavy traffic condition.

Handoff Process

When a user moves from one cell to the other, to keep the communication between the user pair, the user channel has to be shifted from one BS to the other without interrupting the call, i.e., when a MS moves into another cell, while the conversation is still in progress, the MSC automatically transfers the call to a new FDD channel without disturbing the conversation. This process is called as handoff. A schematic diagram of handoff is given in Figure Processing of handoff is an important task in any cellular system. Handoffs must be performed successfully and be imperceptible to the users. Once a signal

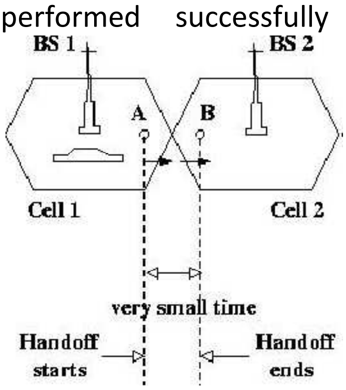


Fig 3:Handoff scenario at two adjacent cell boundary.

level is set as the minimum acceptable for good voice quality (Prmin), then a slightly stronger level is chosen as the threshold (PrH)at which handoff has to be made, as shown in Figure 3.4. A parameter, called power margin, defined as

$$\Delta = P_{r_H} \quad P_{r_{min}}$$

is quite an important parameter during the handoff process since this margin can neither be too large nor too small. If Δis too small, then there may not be enough time to complete the handoff and the call might be lost even if the user crosses the cell boundary. If Δ is too high o the other hand, then MSC has to be burdened with unnecessary handoffs. This is because MS may not intend to enter the other cell. Therefore Δ should be judiciously chosen to ensure imperceptible handoffs and to meet other objectives

Interference & System Capacity

Susceptibility and interference problems associated with mobile communications equipment are because of the problem of time congestion within the electromagnetic spectrum. It is the limiting factor in the performance of cellular systems. This interference can occur from clash with another mobile in the same cell or because of a call in the adjacent cell. There can be interference between the base stations operating at same frequency band or any other non-cellular system's energy leaking inadvertently into the frequency band of the cellular system. If there is an interference in the voice channels, cross talk is heard will appear as noise between the users. The interference in the control channels leads to missed and error calls because of digital signaling. Interference is more severe in urban areas because of the greater RF noise and greater density of mobiles and base stations. The interference can be
divided into 2 parts: co-channel interference and adjacent channel interference.

Co-channel interference (CCI)

For the efficient use of available spectrum, it is necessary to reuse frequency bandwidth over relatively small geographical areas. However, increasing frequency reuse also increases interference, which decreases system capacity and service quality. The cells where the same set of frequencies is used are call co-channel cells. Co-channel interference is the cross talk between two different radio transmitters using the same radio frequency as is the case with the co-channel cells. The reasons of CCI can be because of either adverse weather conditions or poor frequency planning or overly crowded radio spectrum. If the cell size and the power transmitted at the base stations are same then CCI will become independent of the transmitted power and will depend on radius of the cell (R) and the distance between the interfering co-channel cells (D). If D/R ratio is increased, then the effective distance between the co-channel cells will increase 34 and interference will decrease. The parameter Q is called the frequency reuse ratio and is related to the cluster size. For hexagonal geometry

$$Q - D/R - \sqrt{3N}$$

From the above equation, small of \`Q' means small value of cluster size \`N' and increase in cellular capacity. But large \`Q' leads to decrease in system capacity but increase in transmission quality. Choosing the options is very careful for the selection of \`N', the proof of which is given in the first section. The Signal to Interference Ratio (SIR) for a mobile receiver which monitors the forward channel can be calculated as

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

where i0 is the number of co-channel interfering cells, S is the desired signal power from the baseband station and Ii is the interference power caused by the i-th interfering co-channel base station. In order to solve this equation from power calculations, we need to look into the signal power characteristics. The average power in the mobile radio channel decays as a power law of the distance of separation between transmitter and receiver. The expression for the received power Pr at a distance d can be approximately calculated as

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

and in the dB expression as

$$P_r(dB) = P_0(dB) - 10n\log\left(\frac{d}{d_0}\right)$$

where P0 is the power received at a close-in reference point in the far field region at a small distance do from the transmitting antenna, and `n' is the path loss exponent. Let us calculate the SIR for this system. If Di is the distance of the i-th interferer from the mobile, the received power at a given mobile due to i-th interfering cell is proportional to $(D_i)^{-n}$ (the value of 'n' varies between 2 and 4 in urban cellular systems). Let us take that the path loss exponent is same throughout the coverage area and the transmitted power be same, then SIR can be approximated as

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

where the mobile is assumed to be located at R distance from the cell center. If we consider only the first layer of interfering cells and we assume that the interfering base stations are equidistant from the reference base station and the distance between the cell centers is 'D' then the above equation can be converted as

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

which is an approximate measure of the SIR. Subjective tests performed on AMPS cellular system which uses FM and 30 kHz channels show that sufficient voice quality can be obtained by SIR being greater than or equal to 18 dB. If we take n=4 , the value of 'N' can be calculated as 6.49. Therefore minimum N is 7. The above equations are based on hexagonal geometry and the distances from the closest interfering cells can vary if different frequency reuse plans are used. We can go for a more approximate calculation for co-channel SIR. This is the example of a 7 cell reuse case. The mobile is at a distance of D-R from 2 closest interfering cells and approximately D+R/2, D, D-R/2 and D+R distance from other interfering cells in the first tier. Taking n = 4 in the above equation, SIR can be approximately calculated as

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

which can be rewritten in terms frequency reuse ratio Q as

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

Using the value of N equal to 7 (this means Q = 4.6), the above expression yields that worst case SIR is 53.70 (17.3 dB). This shows that for a 7 cell reuse case the worst case SIR is slightly less than 18 dB. The worst case is when the mobile is at the corner of the cell i.e., on a vertex as shown in the Figure 3.6. Therefore N = 12 cluster size should be used. But this reduces the capacity by 7/12 times. Therefore, co-channel interference controls link performance, which in a way controls frequency reuse plan and the overall capacity of the cellular system. The effect of co-channel interference can be minimized by optimizing the frequency assignments of the base stations and their transmit powers. Tilting the base-station antenna to limit the spread of the signals in the system can also be done.

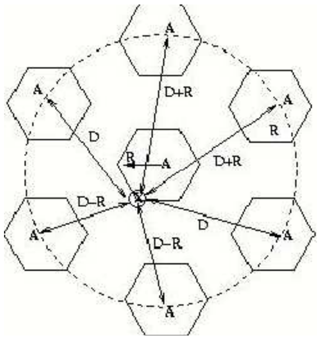


Fig 4:First tier of co-channel interfering cells

Adjacent Channel Interference (ACI)

This is a different type of interference which is caused by adjacent channels i.e. channels in adjacent cells. It is the signal impairment which occurs to one frequency due to presence of another signal on a nearby frequency. This occurs when imperfect receiver filters allow nearby frequencies to leak into the pass band. This problem is enhanced if the adjacent channel user is transmitting in a close range compared to the subscriber's receiver while the receiver attempts to receive a base station on the channel. This is called near-far effect. The more adjacent channels are packed into the channel block, the higher the spectral efficiency, provided that the performance degradation can be tolerated in the system link budget. This effect can also occur if a mobile close to a base station transmits on a channel close to one being used by a weak mobile. This problem might occur if the base station has problem in discriminating the mobile user from the "bleed over" caused by the close adjacent channel mobile. Adjacent channel interference occurs more frequently in small cell clusters and heavily used cells. If the frequency separation between the channels is kept large this interference can be reduced to some extent. Thus assignment of channels is given such that they do not form a contiguous band of frequencies within a particular cell and frequency separation is maximized. Efficient assignment strategies are very much important in making the interference as less as possible. If the frequency factor is small then distance between the adjacent channels cannot put the interference level within tolerance limits. If a mobile is 10 times close to the base station than other mobile and has energy spill out of its pass band, then SIR for weak mobile is approximately

$$\frac{S}{I} = 10^{-n}$$

which can be easily found from the earlier SIR expressions. If n = 4, then SIR is 52 dB. Perfect base station filters are needed when close-in and distant users share the same cell. Practically, each base station receiver is preceded by a high Q cavity filter in order to remove adjacent channel interference. Power control is also very much important for the prolonging of the battery life for the subscriber unit but also reduces reverse channel SIR in the system.

Power control is done such that each mobile transmits the lowest power required to maintain a good quality link on the reverse channel.

Cell-Splitting

Cell Splitting is based on the cell radius reduction and minimizes the need to modify the existing cell parameters. Cell splitting involves the process of sub-dividing a congested cell into smaller cells, each with its own base station and a corresponding reduction in antenna size and transmitting power. This increases the capacity of a cellular system since it increases the number of times that channels are reused. Since the new cells have smaller radii than the existing cells, inserting these smaller cells, known as microcells, between the already existing cells results in an increase of capacity due to the additional number of channels per unit area. There are few challenges in increasing the capacity by reducing the cell radius. Clearly, if cells are small, there would have to be more of them and so additional base stations will be needed in the system. The challenge in this case is to introduce the new base stations without the need to move the already existing base station towers. The other challenge is to meet the generally increasing demand that may vary quite rapidly between geographical areas of the system. For instance, a city may have highly populated areas and so the demand must be supported by cells with the smallest radius. The radius of cells will generally increase as we move from urban to sub urban areas, because the user density decreases on moving towards sub-urban areas. The key factor is to add as minimum number of smaller cells as possible

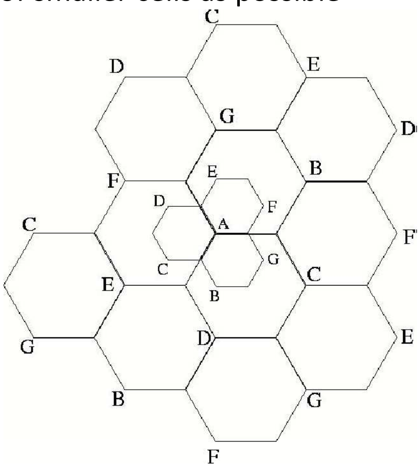


Fig 5:Splitting of congested seven-cell clusters.

wherever an increase in demand occurs. The gradual addition of the smaller cells implies that, at least for a time, the cellular system operates with cells of more than one size. Figure 5 shows a cellular layout with seven-cell clusters. Consider that the cells in the center of the diagram are becoming congested, and cell A in the center has reached its maximum capacity. Figure also shows how the smaller cells are being superimposed on the original layout. The new smaller cells have half the cell radius of the original cells. At half the radius, the new cells will have one-fourth of the area and will consequently need to support one-fourth the number of subscribers. Notice that one of the new smaller cells lies in the center of each of the larger cells. If we assume that base stations are located in the cell centers, this allows the original base stations to be maintained even in the new system layout. However, new base stations will have to be added for new cells that do not lie in the center of the larger cells. The organization of cells into clusters is independent of the cell radius, so that the cluster size can be the same in the small-cell layout as it was in the large-cell layout.

Also the signal-to-interference ratio is determined by cluster size and not by cell radius. Consequently, if the cluster size is maintained, the signal-to-interference ratio will be the same after cell splitting as it was before. If the entire system is 41 replaced with new half-radius cells, and the cluster size is maintained, the number of channels per cell will be exactly as it was before, and the number of subscribers per cell will have been reduced. When the cell radius is reduced by a factor, it is also desirable to reduce the transmitted power. The transmit power of the new cells with radius half that of the old cells can be found by examining the received power PR at the new and old cell boundaries and setting them equal. This is necessary to maintain the same frequency re-use plan in the new cell layout as well. Assume that PT1 and PT2 are the transmit powers of the larger and smaller base stations respectively. Then, assuming a path loss index n=4, we have power received at old cell boundary = PT1=R4 and the power received at new cell boundary = PT2=(R=2)4. On equating the two received powers, we get PT2 = PT1 / 16. In other words, the transmit power must be reduced by 12 dB in order to maintain the same S/I with the new system lay-out. At the beginning of this channel splitting process, there would be fewer channels in the smaller power groups. As the demand increases, more and more channels need to be accommodated and hence the splitting process continues until all the larger cells have been replaced by the smaller cells, at which point splitting is complete within the region and the entire system is rescaled to have a smaller radius per cell. If a cellular layout is replaced entirety by a new layout with a smaller cell radius, the signal-to-interference ratio will not change, provided the cluster size does not change. Some special care must be taken, however, to avoid co-channel interference when both large and small cell radii coexist. It turns out that the only way to avoid interference between the large-cell and small-cell systems is to assign entirely different sets of channels to the two systems. So, when two sizes of cells co-exist in a system, channels in the old cell must be broken down into two groups, one that corresponds to larger cell reuse requirements and the other which corresponds to the smaller cell reuse requirements. The larger cell is usually dedicated to high speed users as in the umbrella cell approach so as to minimize the number of hand-offs.

Sectoring

Sectoring is basically a technique which can increase the SIR without necessitating an increase in the cluster size. Till now, it has been assumed that the base station is located in the center of a cell and radiates uniformly in all the directions behaving as an omni-directional antenna. However it has been found that the co-channel interference in a cellular system may be decreased by replacing a single omni-directional antenna at the base station by several directional antennas, each radiating within a specified sector. In the Figure 3.8, a cell is shown which has been split into three 120o sectors. The base station feeds three 120o directional antennas, each of which radiates into one of the three sectors. The channel set serving this cell has also been divided, so that each sector is assigned one-third of the available number cell of channels. This technique for reducing co-channel interference wherein by using suit-

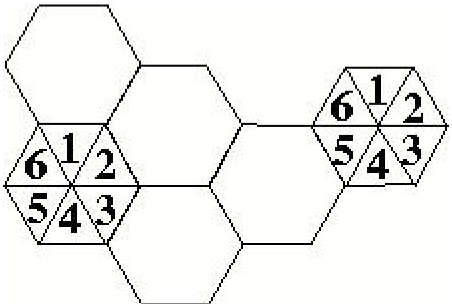


Fig 6:A seven-cell cluster with 60o sectors

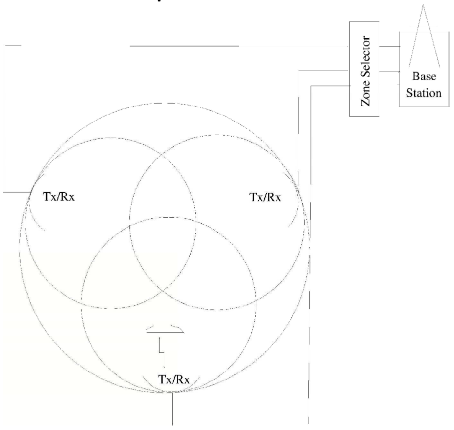
able directional antennas, a given cell would receive interference and transmit with a fraction of available co-channel cells is called 'sectoring'. In a seven-cell-cluster layout with 120o sectored cells, it can be easily understood that the mobile units in a particular sector of the center cell will receive co-channel interference from only two of the first-tier co-channel base stations, rather than from all six. Likewise, the base station in the center cell will receive co-channel interference from mobile units in only two of the co-channel cells. Hence the signal to interference ratio is now modified to

$$\frac{S}{I} = \frac{(\sqrt{3N})^n}{2}$$

where the denominator has been reduced from 6 to 2 to account for the reduced number of interfering sources. Now, the signal to interference ratio for a seven-cell cluster layout using 120o sectored antennas can be found from equation above to be 23.4 dB which is a significant improvement over the Omni-directional case where the worst-case S/I is found to be 17 dB (assuming a path-loss exponent, n=4). Some cellular systems divide the cells into 60o sectors. Similar analysis can be performed on them as well.

Microcell Zone Concept:

The increased number of handoffs required when sectoring is employed results in an increased load on the switching and control link elements of the mobile system. To overcome this problem, a new microcell zone concept has been proposed. As shown in Figure 7 this scheme has a cell divided into three microcell zones, with each of the three zone sites connected to the base station and sharing the same radio equipment. It is necessary to note that all the microcell zones, within a cell, use the same frequency used by that cell; that is no handovers occur between microcells. Thus when a mobile user moves between two microcell zones of the cell, the BS simply switches the channel to a different zone site and no physical re-allotment of channel takes place. Locating the mobile unit within the cell: An active mobile unit sends a signal to all zone sites, which in turn send a signal to the BS. A zone selector at the BS uses that signal to select a suitable zone to serve the mobile unit - choosing the zone with the strongest signal. Base Station Signals: When a call is made to a cellular phone, the system already knows the cell location of that phone. The base station of that cell knows in which zone, within that cell, the cellular phone is located. Therefore when it receives the signal, the base station transmits it to the suitable zone site. The zone site receives the cellular signal from the base station and transmits that signal to the mobile phone after amplification. By confining the power transmitted to the mobile phone, co-channel interference is reduced between the zones and the capacity of system is increased. Benefits of the micro-cell zone concept: 1) Interference is reduced in this case as compared to the scheme in which the cell size is reduced. 2) Handoffs are reduced (also compared to decreasing the cell size) since the microcells within the cell operate at the same frequency; no handover occurs when the mobile unit moves between the microcells. 3) Size of the zone apparatus is small. The zone site equipment being small can be mounted on the side of a building or on poles. 4) System capacity is increased. The new microcell knows where to locate the mobile unit in a particular zone of the cell and deliver the power to that zone. Since the signal power is reduced, the microcells can be closer and result in an increased system capacity. However, in a microcellular system, the transmitted power to a mobile phone within a microcell has to be precise; too much power results in interference between microcells, while with too little power the signal might not reach the mobile phone. This is a drawback of microcellular systems, since a change in the surrounding (a new building, say, within a microcell) will require a change of the transmission power



.

Fig 7: The micro-cell zone concept

Trunked Radio System

In the previous sections, we have discussed the frequency reuse plan, the design trade-offs and also explored certain capacity expansion techniques like cell-splitting and sectoring. Now, we look at the relation between the number of radio channels a cell contains and the number of users a cell can support. Cellular systems use the concept of trunking to accommodate a large number of users in a limited radio spectrum. It was found that a central office associated with say, 10,000 telephones 47 requires about 50 million connections to connect every possible pair of users. However, a worst case maximum of 5000 connections need to be made among these telephones at any given instant of time, as against the possible 50 million connections. In fact, only a few hundreds of lines are needed owing to the relatively short duration of a call. This indicates that the resources are shared so that the number of lines is much smaller than the number of possible connections. A line that connects switching offices and that is shared among users on an as-needed basis is called a trunk. The fact that the number of trunks needed to make connections between offices is much smaller than the maximum number that could be used suggests that at times there might not be sufficient facilities to allow a call to be completed. A call that cannot be completed owing to a lack of resources is said to be blocked. So one important to be answered in mobile cellular systems is: How many channels per cell are needed in a cellular telephone system to ensure a reasonably low probability that a call will be blocked?

In a trunked radio system, a channel is allotted on per call basis. The performance of a radio system can be estimated in a way by looking at how efficiently the calls are getting connected and also how they are being maintained at handoffs. Some of the important factors to take into consideration are (i) Arrival statistics, (ii)Service statistics, (iii)Number of servers/channels.

Module 2

Mobile Radio Propagation :

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 refection, diffractions and scattering, due to the presence of buildings, mountains and other such obstructions. Refection 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.

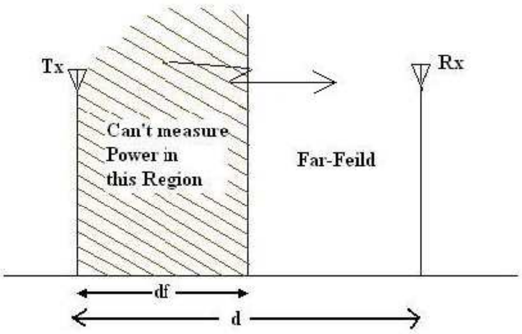


Fig 8: Free space propagation model, showing the near and far fields.

Basic Methods of Propagation:

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 diffractions 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 wave front acts as point sources for the production of secondary wavelets, and they combine to produce a new wave front in the direction of propagation. The propagation of secondary wavelets in the shadowed region results in diffractions. 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 refection and diffractions. The cause is that the trees, buildings and lampposts scatter energy in all directions. This provides extra energy at the receiver. Roughness is tested by a Rayleigh criterion, which defines a critical height hc of surface protuberances for a given angle of incidence θi, given by,

$$h_c = \frac{\lambda}{8 sin\theta_i}.$$

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

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

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

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

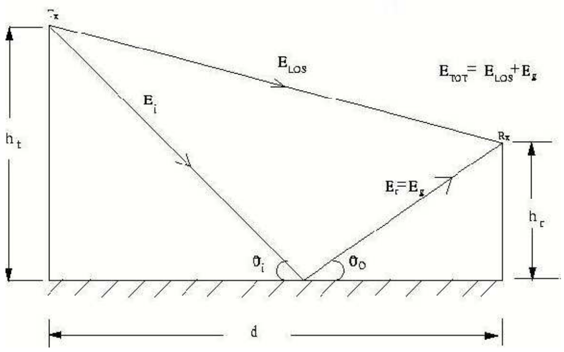


Fig 9: Two-ray refection model

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

$$\Gamma_{rough} = \rho_S \Gamma.$$

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 receiver 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 (ht) ranging from 30 m to 1000 m. Okumura used extensive measurements of base station-to-mobile signal attenuation throughout Tokyo to develop a set of curves giving median attenuation relative to free space (Amu) of signal propagation in irregular terrain. The empirical path loss formula of Okumura at distance d parameterized by the carrier frequency fc is given by

$$P_L(d)dB = L(f_c, d) + A_{mu}(f_c, d) - G(h_t) - G(h_r) - G_{AREA}$$

where L(fc; d) is free space path loss at distance d and carrier frequency fc, Amu(fc; d) is the median attenuation in addition to free-space path loss across all environments(ht) is the base station antenna height gain factor,G(hr) is the mobile antenna height gain factor,GAREA is the gain due to type of environment. The values of Amu(fc; d) and GAREA are obtained from Okumura's empirical plots. Okumura derived empirical formulas for G(ht) and G(hr) as follows:

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

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

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

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.

Multipath & Small-Scale Fading

Multipath signals are received in a terrestrial environment, i.e., where diffierent 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 multipath fading. Adding the effiect 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 uctuations 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 results 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 motion, they induce a time varying Doppler shift on multipath components. If the surrounding objects move at a greater rate than the mobile, then this effect dominates fading.

(4) Transmission Bandwidth of the signal { If the transmitted radio signal bandwidth is greater than the \bandwidth" of the multipath channel (quanti- ffied 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 at fading. So we can say that at fading occurs when

$$B_S \ll B_C$$

where BS is the signal bandwidth and BC is the coherence bandwidth. Also

$T_S \gg \sigma_\tau$

where TS is the symbol period and σ_T 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 coherence bandwidth of the mobile radio channel or equivalently the symbols duration of the signal is less than the rms delay spread.

$$B_S \gg B_C$$

$$T_S \ll \sigma_\tau$$

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 at 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

$$T_S \gg T_C$$

where TC is the coherence time and

$$B_S \gg B_D$$

where BD 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 at least one symbol duration. Hence

$$T_S \ll T_C$$

$$B_S \gg B_D$$

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

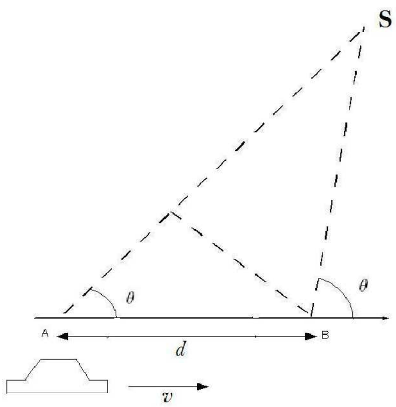


Figure10 : Illustration of Doppler effect.

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 fo 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, vs is the velocity of the source relative to the medium and vr 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

$\pm \frac{v_r}{c} f_0$ or, $\pm \frac{v_r}{\lambda}$.

As the BS is located at an elevated place, a cos factor would also be multiplied with this. The exact scenario, as given in Figure10 , 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

$$\Delta l = d \cos\theta = v \Delta t \cos\theta, \text{ where } \Delta t$$

Where Δt is the time required for the mobile

to travel from A to B, and ff 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

$$\Delta\varphi = \frac{2\pi \Delta l}{\lambda} = \frac{2\pi v \Delta t}{\lambda} \cos\theta$$

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

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

Different Modulation Techniques:

BPSK

In binary phase shift keying (BPSK), the phase of a constant amplitude carrier signal is switched between two values according to the two possible signals m1 and m2 corresponding to binary 1 and 0, respectively. Normally, the two phases are separated by 180o. If the sinusoidal carrier has an amplitude A, and energy per bit $E_o = \frac{1}{2} A_c^2 T_b$ then the transmitted BPSK signal is

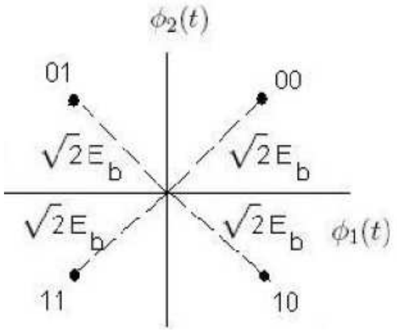
$$s_{BPSK}(t) = m(t) \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t + \theta_c).$$

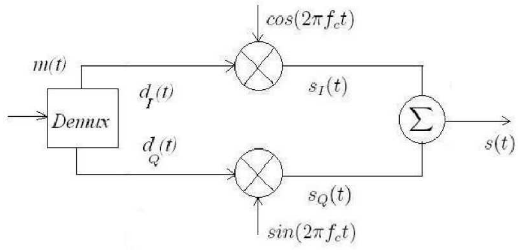
A typical BPSK signal constellation diagram is shown in Figure. The probability of bit error for many modulation schemes in an AWGN channel is found using the Q-function of the distance between the signal points. In case of BPSK,

$$P_{eBPSK} = Q(\sqrt{\frac{2E_b}{N_0}}).$$

QPSK

The Quadrature Phase Shift Keying (QPSK) is a 4-ary PSK signal. The phase of the carrier in the QPSK takes 1 of 4 equally spaced shifts. Although QPSK can be viewed as a quaternary modulation, it is easier to see it as two independently modulated quadrature carriers. With this interpretation, the even (or odd) bits are





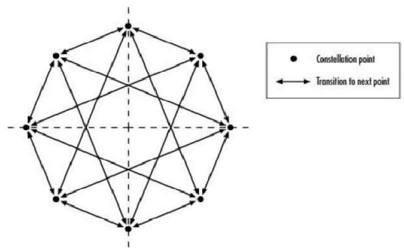
used to modulate the in-phase component of the carrier, while the odd (or even) bits are used to modulate the quadrature-phase component of the carrier. The QPSK transmitted signal is defined by:

$$s_i(t) = A\cos(\omega t + (i-1)\pi/2), i = (1, 2, 3, 4)$$

and the constellation disgram is shown in Figure

Offset-QPSK

As in QPSK, as shown in Figure 6.3, the NRZ data is split into two streams of odd and even bits. Each bit in these streams has a duration of twice the bit duration,



Tb, of the original data stream. These odd (d1(t)) and even bit streams (d2(t)) are then used to modulate two sinusoidals in phase quadrature,and hence these data streams are also called the in-phase and and quadrature phase components. After modulation they are added up and transmitted. The constellation diagram of Offset- QPSK is the same as QPSK. Offset-QPSK differs from QPSK in that the d1(t) and d2(t) are aligned such that the timing of the pulse streams are offset with respect to each other by Tb seconds. From the constellation diagram it is observed that a signal point in any quadrant can take a value in the diagonally opposite quadrant only when two pulses change their polarities together leading to an abrupt 180 degree phase shift between adjacent symbol slots. This is prevented in O-QPSK and the allowed phase transitions are 90 degree. Abrupt phase changes leading to sudden changes in the signal amplitude in the time domain corresponds to significant out of band high frequency components in the frequency domain. Thus to reduce these sidelobes spectral shaping is done at baseband. When high effciency power ampliffers, whose non-linearity increases as the effciency goes high, are used then due to distortion, harmonics are generated and this leads to what is known as spectral regrowth.

Since sudden 180 degree phase changes cannot occur in OQPSK, this problem is reduced to a certain extent.
DQPSK like QPSK can be thought to be carried in the phase of a single modulated carrier or on the amplitudes of a pair of quadrature carriers. The
modulated signal during the time slot of kT < t < (k + 1)T given by:

$$s(t) = cos(2\pi f_c t + \psi_{k+1})$$

Here

$\psi_{k+1} = \psi_k + \Delta\psi_k$ and $\Delta\psi_k$ can take values $\pi/4$ for 00, $3\pi/4$ for 01, $-3\pi/4$

This corresponds to eight points in the signal constellation but at any instant of time only one of the four points are possible: the four points on axis or the four points axis. The constellation diagram along with possible transitions are shown in Figure

Angle Modulation (FM and PM)

There are a number of ways in which the phase of a carrier signal may be varied in accordance with the baseband signal; the two most important classes of angle modulation being frequency modulation and phase modulation. Frequency modulation (FM) involves changing of the frequency of the carrier signal according to message signal. As the information in frequency modulation is in the frequency of modulated signal, it is a nonlinear modulation technique. In this method, the amplitude of the carrier wave is kept constant (this is why FM is called constant envelope). FM is thus part of a more general class of modulation known as angle modulation. Frequency modulated signals have better noise immunity and give better performance in fading scenario as compared to amplitude modulation.Unlike AM, in an 114 FM system, the modulation index, and hence bandwidth occupancy, can be varied to obtain greater signal to noise performance.This ability of an FM system to trade bandwidth for SNR is perhaps the most important reason for its superiority over AM. However, AM signals are able to occupy less bandwidth as compared to FM signals, since the transmission system is linear. An FM signal is a constant envelope signal, due to the fact that the envelope of the carrier does not change with changes in the modulating signal. The constant envelope of the transmitted signal allows effcient Class C power ampliffers to be used for RF power ampliffcation of FM. In AM, however, it is critical to maintain linearity between the applied message and the amplitude of the transmitted signal, thus linear Class A or AB amplifiers, which are not as power efficient, must be used. FM systems require a wider frequency band in the transmitting media (generally several times as large as that needed for AM) in order to obtain the advantages of reduced noise and capture effect. FM transmitter and receiver equipment is also more complex than that used by amplitude modulation systems. Although frequency modulation systems are tolerant to certain types of signal and circuit nonlinearities, special attention must be given to phase characteristics. Both AM and FM may be demodulated using inexpensive non-coherent detectors. AM is easily demodulated using an envelope detector whereas FM is demodulated using a discriminator or slope detector. In FM the instantaneous frequency of the carrier signal is varied linearly with the baseband message signal m(t), as shown in following equation:

$$s_{FM}(t) = A_c \cos[2\pi f_c t + \theta(t)] = A_c \cos[2\pi f_c t + 2\pi k_f \int m(\eta) d\eta]$$

where Ac, is the amplitude of the carrier, fc is the carrier frequency, and kf is the frequency deviation constant (measured in units of Hz/V). Phase modulation (PM) is a form of angle modulation in which the angle _(t) of the carrier signal is varied linearly with the baseband message signal m(t), as shown in equation below.

$$s_{PM}(t) = A_c \cos(2\pi f_c t + k_\theta m(t))$$

The frequency modulation index Δf, defines the relationship between the message amplitude and the bandwidth of the transmitted signal, and is given by

$$\beta_f = \frac{k_f A_m}{W} = \frac{\Delta}{W}$$

where Am is the peak value of the modulating signal, Δf is the peak frequency deviation of the transmitter and W is the maximum bandwidth of the modulating signal.
The phase modulation index Δp is given by

$$\beta_p = k_\theta A_m = \Delta\theta$$

where, Δθ is the peak phase deviation of the transmitter.

BFSK

In Binary Frequency Shift keying (BFSK),the frequency of constant amplitude carrier signal is switched between two values according to the two possible message states (called high and low tones) corresponding to a binary 1 or 0. Depending on how the frequency variations are imparted into the transmitted waveform,the FSK signal will have either a discontinuous phase or continuous phase between bits. In general, an FSK signal may be represented as

$$S(t) = \sqrt{(2E_b/T)}\cos(2\pi f_i t).$$

where T is the symbol duration and Eb is the energy per bit

$$S_i = \sqrt{(E_b)}\phi(t).$$

$$\phi(t) = \sqrt{(2/T)}\cos(2\pi f_i t).$$

There are two FSK signals to represent 1 and 0, i.e.,

$$S_1(t) = \sqrt{(2E_b/T)}\cos(2\pi f_1 t + \theta(0)) \qquad \rightarrow 1$$

$$S_2(t) = \sqrt{(2E_b/T)}\cos(2\pi f_2 t + \theta(0)) \qquad \rightarrow 0$$

where θ(0) sums the phase up to t = 0. Let us now consider a continuous phase FSK as

$$S(t) = \sqrt{(2E_b/T)}\cos(2\pi f_c t + \theta(t))$$

Expressing θ(t) in terms of θ(0) with a new unknown factor h, we get

$$\theta(t) = \theta(0) \pm \pi h t/T \qquad\qquad 0 \le t \le T$$

$$S(t) = \sqrt{\frac{2E_b}{T}}\cos(2\pi f_c t \pm \pi h t/T + \theta(0)) = \sqrt{\frac{2E_b}{T}}\cos(2\pi(f_c \pm h/2T)t + \theta(0))$$

It shows that we can choose two frequencies f1 and f2 such that

$$f_1 = f_c + h/2T$$

$$f_2 = f_c - h/2T$$

for which the expression of FSK conforms to that of CPFSK. On the other hand, fc and h can be expressed in terms of f1 and f2 as

$$f_c = [f_1 + f_2]/2$$

$$h = \frac{(f_1 - f_2)}{1/T}.$$

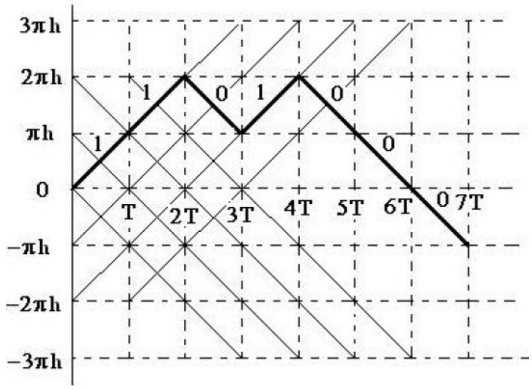
Therefore, the unknown factor h can be treated as the difference between f1 and f2, normalized with respect to bit rate 1=T . It is called the deviation ratio. We know that

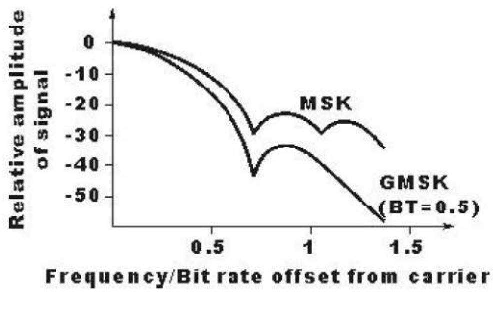
$\theta(t) - \theta(0) = \pm\pi ht/T,\ 0 \leq t \leq T$

If we substitute t = T, we have

$$\begin{aligned}\theta(T) - \theta(0) &= \pm\pi h && \text{where}\\ &= \pi h && \rightarrow 1\\ &= -\pi h && \rightarrow 0\end{aligned}$$

This type of CPFSK is advantageous since by looking only at the phase, the transmitted bit can be predicted. In Figure 6.8, we show a phase tree of such a CPFSK signal with the transmitted bit stream of 1101000. A special case of CPFSK is achieved with h = 0:5, and the resulting scheme is called Minimum Shift Keying (MSK) which is used in mobile communications. In this case, the phase differences reduce to only θ_=2 and the phase tree is called the phase trellis. An MSK signal can also be thought as a special case of OQPSK where the baseband rectangular pulses are replaced by half sinusoidal pulses. Spectral characteristics of an MSK signal is shown in Figure 6.9 from which it is clear that ACI is present in the spectrum. Hence a pulse shaping technique is required. In order to have a compact signal spectrum as well as maintaining the constant envelope property, we use a pulse shaping filter with



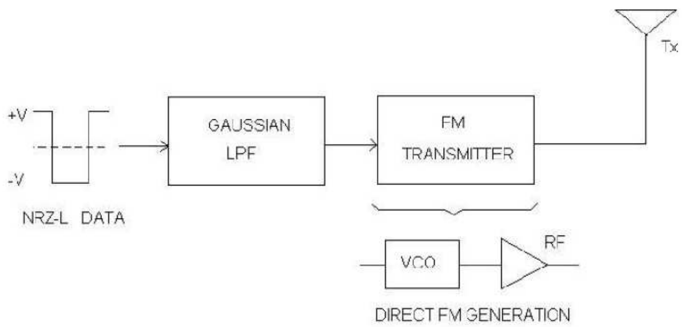


1. a narrow BW frequency and sharp cutoff characteristics (in order to suppress the high frequency component of the signal);
2. an impulse response with relatively low overshoot (to limit FM instant frequency deviation;

GMSK Scheme

GMSK is a simple modulation scheme that may be taken as a derivative of MSK. In GMSK, the sidelobe levels of the spectrum are further reduced by passing a non- return to zero (NRZ-L) data waveform through a premodulation Gaussian pulse shaping filter. Baseband

Gaussian pulse shaping smoothes the trajectory of the MSK signals and hence stabilizes instantaneous frequency variations over time. This has the effect of considerably reducing the sidelobes in the transmitted spectrum. A GMSK generation scheme with NRZ-L data is shown in Figure and a receiver of the same scheme with some MSI gates is shown in Figure



Module 3

Spread Spectrum Techniques :

Spread Spectrum techniques have some powerful properties which make them an excellent candidate for networking applications. To better understand why, we will take a closer look at this fascinating area, and its implications for networking.
The first major application of Spread Spectrum Techniques (SST) arose during the mid-sixties, when NASA employed the method to precisely measure the range to deep space probes. In the following years, the US military became enamoured of SST due to its ability to withstand jamming (ie intentional interference), and it ability to resist eavesdropping. Today this technology forms the basis for the ubiquitous NavStar Global Positioning System (GPS), the soon to become ubiquitous JTIDS (Joint Tactical Information Distribution System/Link-16) datalink (used between aircraft, ships and land vehicles), and last but not least, the virtually undetectable bombing and navigation radar on the bat-winged B-2 bomber. if you ever get asked what your mobile networked laptop shares in common with a stealth bomber (excluding astronomical cost), you can state without fear of contradiction that it uses the same class of modulation algorithm.How is this black magic achieved ? The starting point is Claude Shannon's information theory, a topic beloved by diehard communications engineers. Shannon's formula for channel capacity is a relationship between achievable bit rate, signal bandwidth and signal to noise ratio. Channel capacity is proportional to bandwidth and the logarithm to the base of two of one plus the signal to noise ratio, or:*Capacity = Bandwidth*log2 (1 + SNR).*What this means is that the more bandwidth and the better the signal to noise ratio, the more bits per second you can push through a channel. This is indeed common sense. However, let us consider a situation where

the signal is weaker than the noise which is trashing it. Under these conditions this relationship becomes much simpler, and can be approximated by a ratio of

Capacity/Bandwidth = 1.44 SNR.*

What this says is that we can trade signal to noise ratio for bandwidth, or vice versa. If we can find a way of encoding our data into a large signal bandwidth, then we can get error free transmission under conditions where the noise is much more powerful than the signal we are using. This very simple idea is the secret behind spread spectrum techniques. Consider the example of a 3 kHz voice signal which we wish to send through a channel with a noise level 100 times as powerful as the signal. Manipulating the preceding equation, we soon find that we require a bandwidth of 208 kHz, which is about 70 times greater than the voice signal we wish to carry. Readers with a knowledge of radio will note here that this idea of spreading is a central part of FM radio and the reason why it produces good sound quality compared to the simpler AM scheme. Other than punching through large levels of background noise, why would we otherwise consider using spread spectrum techniques ? There are a number of good practical reasons why spread spectrum modulation is technically superior to the intuitively more obvious techniques such as AM and FM, and all of the hybrids which lie in between.

- The Ability to Selectively Address. If we are clever about how we spread the signal, and use the proper encoding method, then the signal can only be decoded by a receiver which knows the transmitter's code. Therefore by setting the transmitter's code, we can target a specific receiver in a group, or vice versa. This is termed *Code Division Multiple Access*.
- Bandwidth Sharing. If we are clever about selecting our modulation codes, it is entirely feasible to have multiple pairs of receivers and transmitters occupying the same bandwidth. This would be equivalent to having say ten TV channels all operating at the same frequency. In a world where the radio spectrum is being busily carved up for commercial broadcast users, the ability to share bandwidth is a valuable capability.
- Security from Eavesdropping. If an eavesdropper does not know the modulation code of a spread spectrum transmission, all the eavesdropper will see is random electrical noise rather than something to eavesdrop. If done properly, this can provide almost perfect immunity to interception.
- Immunity to Interference. If an external radio signal interferes with a spread spectrum transmission, it will be rejected by the demodulation mechanism in a fashion similar to noise. Therefore we return to the starting point of this discussion, which is that spread spectrum methods can provide excellent error rates even with very faint signals.
- Difficulty in Detection. Because a spread spectrum link puts out much less power per bandwidth than a conventional radio link, having spread it over a wider bandwidth, and a knowledge of the link's code is required to demodulate it, spread spectrum signals are extremely difficult to detect. This means that they can coexist with other more conventional signals without causing catastrophic interference to narrowband links.

These characteristics endeared spread spectrum comes to the military community, who are understandably paranoid about being eavesdropped and jammed. However, the same properties are no less useful for local area networking over radio links. Indeed these are the reasons why the current IEEE draft specification for radio LANs is written around spread spectrum modulations. To better understand the inner workings of this fascinating area, we will now more closely examine the various choices we have for spread spectrum designs. The two basic methods are indeed both used in LAN equipment.

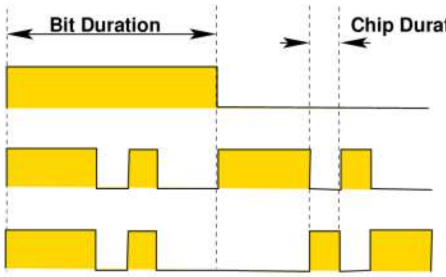


Fig 11: DIRECT SPREADING (DS)

Direct Sequence Systems : Direct Sequence (DS) methods are the most frequently used spread spectrum technique, and also the conceptually simplest to understand. DS modulation is achieved by modulating the carrier wave with a digital code sequence which has a bit rate much higher than that of the message to be sent. This code sequence is typically a pseudorandom binary code (often termed "pseudo-noise" or PN), specifically chosen for desirable statistical properties. In effect we are transmitting a wideband noise like signal which contains embedded message data. The time period of a single bit in the PN code is termed a *chip*, and the bit rate of the PN code is termed the *chip rate*. A wide range of pseudorandom codes exist which can be applied to this task. These codes should ideally be balanced, with an equal number of ones and zeroes over the length of the sequence (also termed the code run), as well a good code should be cryptographically secure. A spread spectrum system which uses a cryptographically insecure code will still possess the properties previously discussed, but if an eavesdropper can synchronise on to the signal they should be able to eventually crack it and extract the data. Using a secure code prevents this. The mechanics of generating pseudorandom codes is a fascinating area within itself. The most commonly used approach for producing a wide range of code types is the use of a tapped register with feedback, very simple to implement in hardware. A PN code generator of this type uses a register with taps between selected stages. These taps are logically ORed and then fed back in to the input stage of the register. The state machine produced in this fashion will periodically cycle through the same PN sequence as the clock is applied. Significantly, code sequence lengths of up to thousands of bits in length can be produced with about a dozen register stages. With modern VLSI techniques it is feasible to build generators with clock speeds up to hundreds of MHz on any die, moreover recent high speed Emitter Coupled Logic devices allow the creation of generators with clock speeds into the GHz region.

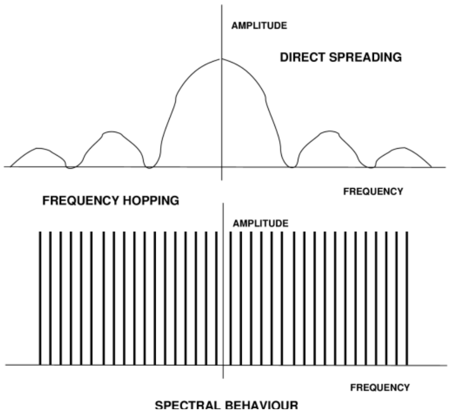


Fig 12:

Having produced a black box which generates a PN code with the required characteristics, the process of combining the PN modulation with the data to be transmitted, and modulating this upon a carrier is not technically difficult at all. The simplest technique, one of many, is to invert the PN code when a '0' bit of message data is to be sent, and to transmit the PN code unchanged when a '1' bit of message data is to be sent. This technique is termed *Bit Inversion Modulation*. The result is a PN code with an embedded data message. The simplest form of carrier modulation which can be used is AM, however in practice one or another form of *Phase Shift Keying (PSK)* is usually employed. PSK schemes are commonly used in modems, and involve the modulation of the carrier phase with the data signal. In a DS transmitter using Binary PSK, the carrier wave is phase shifted back and forth 180 degrees with each 1 or 0 in the PN code chip stream being sent. The process of modulating the carrier with the PN code is often termed *spreading*.
The internals of a DS receiver are somewhat more complex than those of the transmitter, but not vastly so. The central idea in all SST receivers is the use of the correlation operation. Correlation, a favourite method of our friends in the statistics community, is a mathematical operation which determines a measure of likeness or similarity between two sets of data or two time processes. In an SST receiver, the correlation operation is use to measure the similarity of a received PN code sequence to an internally generated PN code sequence. Ideally, if these PN sequences are the same, a high correlation will be detected, whereas if the codes are different, a low correlation is detected.
Mathematically the correlation operation, in its simplest form, is the integral of the product of two time varying functions. In a DS receiver of the simplest kind, the hardware maps directly onto the basic maths. The *correlator* is built by combining a multiplier with a low pass filter (ie integrator in a control engineer's language).
One of the two time varying functions is the received PN modulated signal, the other is the PN sequence produced by a PN generator internal to the receiver. In the simplest situation, the receiver's PN generator is a clone of the PN generator in the transmitter.
The multiplier can be one of many designs, importantly it multiplies in effect two single numbers and is therefore trivially simple. Classical textbooks cite the analogue doubly balanced mixer as the standard multiplier. The output from the multiplier is a time varying measure of the similarity between the two codes, blended with the remnants of uncorrelated (ie real) noise and interfering signals.
The integration operation disposes of the latter, and we are then left with the data which we intended to extract. This series of operations is often termed *despreading*. In practice,

we often need to synchronise our receiver's PN generator to the incoming SST signal, therefore there is often much additional complexity required to produce an internal reference PN sequence in proper lockstep with the incoming message PN sequence.
At this point it is worth reflecting upon what we have. We can generate either cryptographically secure or insecure codes. We can embed a digital data stream in one or another fashion into the code stream. All of this can be performed with pure digital logic. Once we have a combined data/code stream, we can use a very simple analogue modulation to put the message upon a carrier.
The resulting radio signal looks like white noise to a third party who doesn't know out code. Our receiver shares similar hardware design with our transmitter. It uses a trivial demodulation scheme, and extracts digital data from the incoming PN data/code stream. Other radio signals occupying our bandwidth are largely ignored. Whilst an SST transmitter-receiver pair may be conceptually more complex to understand than most classical analogue schemes, it is well suited to implementation in digital logic because most of the smarts at either end of the link are purely digital. This means that such hardware can be made much more compact than many classical narrowband analogue schemes, which often require a lot of analogue hardware which may or may not be easy to squeeze into Silicon.
Consider a narrowband 16 or 64 level QAM scheme, which is not only vulnerable to interference and noise, but also requires a digital signal processing chip to demodulate. For those readers with a bent toward radio engineering, the spectral envelope of a DS system is typically a sinc function, with suppressed outer sidebands beyond the first null, and often a suppressed carrier. A parameter which radio types will appreciate is *process gain*, a measure of signal to noise ratio improvement achieved by despreading the received signal. For a DS system it is typically about twice the ratio of RF bandwidth to message bandwidth.
Therefore to improve your ability to reject interference by 20 dB, you need to increase your chip rate by a factor of 100.

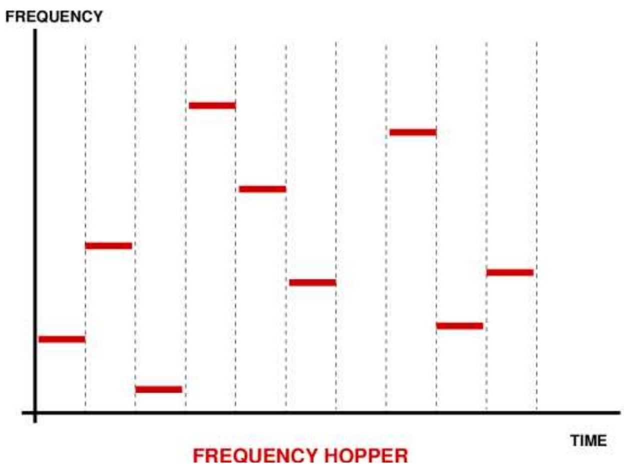


Fig 13

Frequency Hopping Systems : Frequency Hoppers (FH) are a more sophisticated and arguably better family of spread spectrum techniques than the simpler DS systems. However, performance comes with a price tag here, and FH systems are significantly more complex than DS systems. The central idea behind a FH system is to retune the transmitter RF carrier frequency to a pseudorandomly determined frequency value. In this fashion the carrier keeps popping up a different frequencies, in a pseudorandom pattern. The carrier itself amy be modulated directly with the data using one of many possible schemes. The available radio spectrum is thus split up into a discrete number of frequency channels, which are occupied by the RF carrier pseudorandomly in time. Unless you know the PN code used, you have no idea where the carrier wave is likely to pop up next, therefore eavesdropping will be quite difficult. Frequency hoppers are typically divided into fast and slow hoppers. A slow frequency hopper will change carrier frequency pseudorandomly at a frequency which is much slower than the data bit rate on the carrier. A fast frequency hopper will do so at a frequency which is faster than that of the data message.

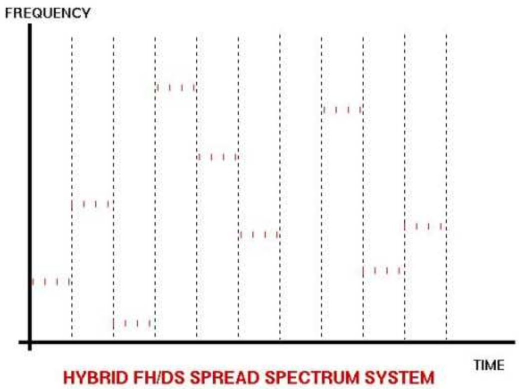


Fig 14

Hybrid(FH/DS)Systems

If we are *really* paranoid about being eavesdropped, we can take further steps to make our signal difficult to find. A commonly used example is that of a hybrid spread spectrum system using both FH and DS techniques. Such schemes will typically employ frequency hopping of the carrier wave, while concurrently using a DS modulation technique to modulate the data upon the carrier. In this fashion an essentially DS modulated message is hopped about the spectrum. To successfully intercept such a signal you must first crack the FH code, and then crack the DS code. If you want to be further secure, you encrypt your data stream with a very secure crypto code before you feed it into your DS modulator, and employ cryptographically secure PN codes for the DS and FH operations. Your eavesdropper then has to chew his way through three levels of encoding. Such a scheme is used in the military JTIDS/Link 16 datalink.

Multiple Access Techniques : In wireless communication systems it is often desirable to allow the subscriber to send simultaneously information to the base station while receiving information from the base station. A cellular system divides any given area into cells where a mobile unit in each cell communicates with a base station. The main aim in the cellular system design is to be able to increase the capacity of the channel i.e. to handle as many calls as possible in a given bandwidth with a sufficient level of quality of service. There are several different ways to allow access to the channel. These includes mainly the following:

1) Frequency division multiple-access (FDMA)
2) Time division multiple-access (TDMA)
3) Code division multiple-access (CDMA)
4) Space Division Multiple access (SDMA)

Frequency Division Multiple Access

This was the initial multiple-access technique for cellular systems in which each individual user is assigned a pair of frequencies while making or receiving a call as shown in Figure below. One frequency is used for downlink and one pair for uplink. This is called frequency division duplexing (FDD). That allocated frequency pair is not used in the same cell or adjacent cells during the call so as to reduce the co-channel interference. Even though the user may not be talking, the spectrum cannot be reassigned as long as a call is in place. Different users can use the same frequency in the same cell except that they must transmit at different times. The features of FDMA are as follows: The FDMA channel carries only one phone circuit at a time. If an FDMA channel is not in use, then it sits idle and it cannot be used by other users to increase share capacity. After the assignment of the voice channel the BS and the MS transmit simultaneously and continuously. The bandwidths of FDMA systems are generally narrow i.e. FDMA is usually 159 implemented in a narrow band system The symbol time is large compared to the average delay spread. The complexity of the FDMA mobile systems is lower than that of TDMA mobile systems. FDMA requires tight filtering to minimize the adjacent channel interference.

Time Division Multiple Access

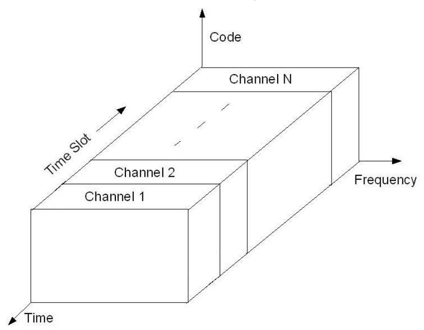


Fig 15: Time division Multiplexing

In digital systems, continuous transmission is not required because users do not use the allotted bandwidth all the time. In such cases, TDMA is a complimentary access technique to FDMA. Global Systems for Mobile communications (GSM) uses the TDMA technique. In TDMA, the entire bandwidth is available to the user but only for a finite period of time. In most cases the available bandwidth is divided into fewer channels compared to FDMA and the users are allotted time slots during which they have the entire channel bandwidth at their disposal, as shown in Figure above. TDMA requires careful time synchronization since users share the bandwidth in the frequency domain. The number of channels are less, inter channel interference is almost negligible. TDMA uses different time slots for transmission and reception.

This type of duplexing is referred to as Time division duplexing(TDD). The features of TDMA includes the following: TDMA shares a single carrier frequency with several users where each users makes use of non overlapping time slots. The number of time slots per frame depends on several factors such as modulation technique, available bandwidth etc. Data transmission in TDMA is not continuous but occurs in bursts. This results in low battery consumption since the subscriber transmitter can be turned OFF when not in use. Because of a discontinuous transmission in TDMA the handoff process is much simpler for a subscriber unit, since it is able to listen to other base stations during idle time slots. TDMA uses different time slots for transmission and reception thus duplexers are not required. TDMA has an advantage that is possible to allocate different numbers of time slots per frame to different users. Thus bandwidth can be supplied on demand to different users by concatenating or reassigning time slot based on priority.

Code Division Multiple Access: In CDMA, the same bandwidth is occupied by all the users, however they are all assigned separate codes, which differentiates them from each other (shown in Figure). CDMA utilize a spread spectrum technique in which a spreading signal (which is uncorrelated to the signal and has a large bandwidth) is used to spread the narrow band message signal. The most commonly used technology for CDMA is Direct Sequence Spread Spectrum (DS-SS) . In DS-SS, the message signal is multiplied by a Pseudo Random Noise Code. Each user is given his own codeword which is orthogonal to the codes of other users and in order to detect the user, the receiver must know the codeword used by the transmitter. There are, however, two problems in such systems which are discussed in the sequel.

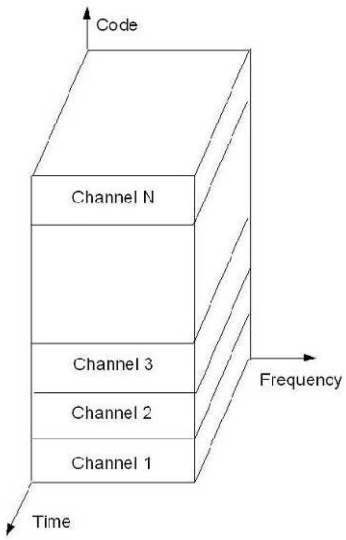


Fig 16 Code Division Multiple Access

Space Division Multiple Access

SDMA utilizes the spatial separation of the users in order to optimize the use of the frequency spectrum. A primitive form of SDMA is when the same frequency is reused in different cells in a cellular wireless network. The radiated power of each user is controlled by Space division multiple access. SDMA serves different users by using spot beam antenna. These areas may be served by the same frequency or different frequencies. However for limited co-channel interference it is required that the cells be sufficiently separated. This limits the number of cells a region can be divided into and hence limits the frequency re-use factor. A more advanced approach can further increase the capacity of the network. This technique would enable frequency re-use within the cell. In a practical cellular environment it is improbable to have just one transmitter fall within the receiver beam width. Therefore it becomes imperative to use other multiple access techniques in conjunction with SDMA. When different areas are covered by the antenna beam, frequency can be re-used, in which case TDMA or CDMA is employed, for different frequencies FDMA can be used.

Module 4

Various Generations of Wireless Networks:

At the initial phase, mobile communication was restricted to certain official users and the cellular concept was never even dreamt of being made commercially available. Moreover, even the growth in the cellular networks was very slow. However, with the development of newer and better technologies starting from the 1970s and with the mobile users now connected to the PSTN, there has been a remarkable growth in the cellular radio. However, the spread of mobile communication was very fast in the 1990s when the government throughout the world provided radio spectrum licenses for Personal Communication Service (PCS) in 1.8 - 2 GHz frequency band.

First Generation Networks

The first mobile phone system in the market was AMPS. It was the first U.S. cellular telephone system, deployed in Chicago in 1983. The main technology of this first generation mobile system was FDMA/FDD and analog FM.

Second Generation Networks

Digital modulation formats were introduced in this generation with the main technology as TDMA/FDD and CDMA/FDD. The 2G systems introduced three popular TDMA standards and one popular CDMA standard in the market. These are as follows:

TDMA/FDD Standards

(a) Global System for Mobile (GSM): The GSM standard, introduced by Groupe Special Mobile, was aimed at designing a uniform pan-European mobile system. It was the first fully digital system utilizing the 900 MHz frequency band. The initial GSM had 200 KHz radio channels, 8 full-rate or 16 half-rate TDMA channels per carrier, encryption of speech, low speed data services and support for SMS for which it gained quick popularity.

(b) Interim Standard 136 (IS-136): It was popularly known as North American Digital Cellular (NADC) system. In this system, there were 3 full-rate TDMA users over each 30 KHz channel. The need of this system was mainly to increase the capacity over the earlier analog (AMPS) system.

(c) Pacific Digital Cellular (PDC): This standard was developed as the counterpart of NADC in Japan. The main advantage of this standard was its low transmission bit rate which led to its better spectrum utilization.

CDMA/FDD Standard

Interim Standard 95 (IS-95): The IS-95 standard, also popularly known as CDMAOne,
uses 64 orthogonally coded users and codewords are transmitted simultaneously on each of 1.25 MHz channels. Certain services that have been standardized as a part of IS-95 standard are: short messaging service, slotted paging, over-the-air activation (meaning the mobile can be activated by the service provider without any third party intervention), enhanced mobile station identities etc.

2.5G Mobile Networks

In an effort to retrofft the 2G standards for compatibility with increased throughput rates to support modern Internet application, the new data centric standards were developed to be overlaid on 2G standards and this is known as 2.5G standard. Here, the main upgradation techniques are:

ff supporting higher data rate transmission for web browsing

ff supporting e-mail traffic

ff enabling location-based mobile service

2.5G networks also brought into the market some popular application, a few of which are: Wireless Application Protocol (WAP), General Packet Radio Service (GPRS), High Speed Circuit Switched Dada (HSCSD), Enhanced Data rates for GSM Evolution (EDGE) etc.

3G: Third Generation Networks

3G is the third generation of mobile phone standards and technology, superseding 2.5G. It is based on the International Telecommunication Union (ITU) family of standards under the International Mobile Telecommunications-2000 (IMT-2000). ITU launched IMT-2000 program, which, together with the main industry and standardization bodies worldwide, targets to implement a global frequency band that would support a single, ubiquitous wireless communication standard for all countries, to provide the framework for the definition of the 3G mobile systems. Several radio access technologies have been accepted by ITU as part of the IMT-2000 framework. 3G networks enable network operators to offer users a wider range of more advanced services while achieving greater network capacity through improved spectral efficiency. Services include wide-area wireless voice telephony, video calls, and broadband wireless data, all in a mobile environment. Additional features also include HSPA data transmission capabilities able to deliver speeds up to 14.4Mbit/s on the
down link and 5.8Mbit/s on the uplink. 3G networks are wide area cellular telephone networks which evolved to incorporate high-speed internet access and video telephony. IMT-2000 defines a set of technical requirements for the realization of such targets, which can be summarized as follows:

_ high data rates: 144 kbps in all environments and 2 Mbps in low-mobility and
indoor environments

_ symmetrical and asymmetrical data transmission

_ circuit-switched and packet-switched-based services
_ speech quality comparable to wire-line quality
_ improved spectral efficiency
_ several simultaneous services to end users for multimedia services
_ seamless incorporation of second-generation cellular systems
_ global roaming
_ open architecture for the rapid introduction of new services and technology.

Beyond 3G networks, or 4G (Fourth Generation), represent the next complete evolution in wireless communications. A 4G system will be able to provide a comprehensive IP solution where voice, data and streamed multimedia can be given to users at higher data rates than previous generations. There is no formal definition for 4G ; however, there are certain objectives that are projected for 4G. It will be capable of providing between 100 Mbit/s and 1 Gbit/s speeds both indoors and outdoors, with premium quality and high security. It would also support systems like multicarrier communication, MIMO and UWB.

GSM Architecture

Figure below depicts the original GSM architecture that supported circuit switching services only. Mobile Stations (MS) (handheld phones) have a radio connection with a cell. One or a set of cells are supported by a Base Transceiver Station (BTS). A BTS may have an antenna high above the ground on top of a mast in order to increase the coverage of the cells. Cells cover the area where GSM users are reachable. Cells may be organized into several layers. Large cells are useful for fast moving Mobile Stations and small cells are needed to increase network capacity – the number of MSs that can be served simultaneously in a particular spot. A call may start in one cell, the MS may traverse through a number of cells while the call is on and the MS may be located in a cell under a different BTS or even a different MSC when the call ends. The action of changing a cell during a call is called a handover. Several BTSs are controlled by a Base Station Controller (BSC). There may be many BSCs under the control of one Mobile Switching Center (MSC). An MSC controlling BSCs is in the same position GSM as a Local Exchange in PSTN or ISDN. The difference is that an MSC does "own" its subscribers. Rather, all the MSs it is controlling are visitors. A Visitor Location Register (VLR) for storing information about the users and MSs visiting this MSC is separa tely specified but always resides in the MSC. Recall that also a wire line Local Exchange contains a subscriber database.

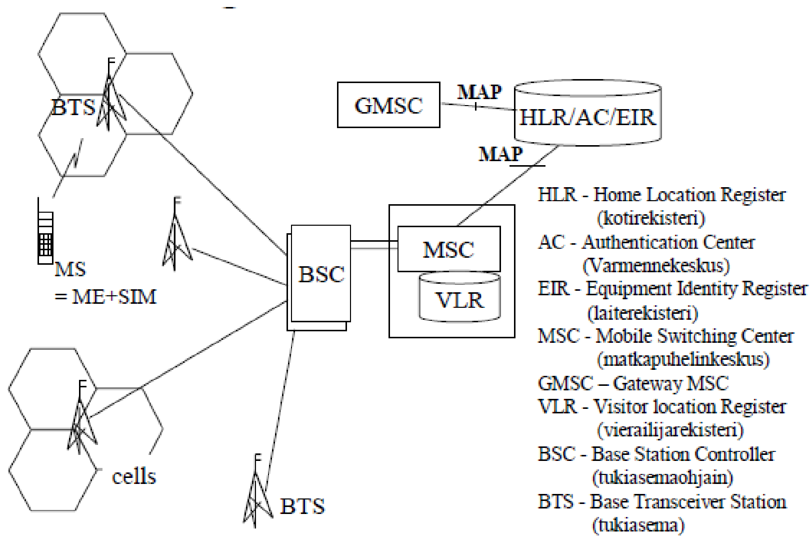


Fig 17: GSM Architecture

Cells normally forming a continuous geographical area are organized into a Location Area. A location area is a logical concept and gives the accuracy with which the location of MSs is maintained in the HLR. Each MS has a MSISDN number. It is a logical or a directory number. An MSISDN number maps to a particular HLR based on its leading digits. Thus MSISDN is routable to the HLR that is supposed to know where the MS is located. The HLR also house the Authentication Center (AC) for storing authentication information about the users and the Equipment Identity Register (EIR). Each MS has an Equipment Identity on silicon and thus stolen equipment can be identified, traced and removed from the network. MSs are reachable because they regularly make location updates to the VLR. The VLR will authenticate the user and establish that the user can be billed for the use of the network resources by contacting the HLR. When the MS first establishes a connection with the network or changes the Location Area, VLR will update the location of the MS in the user's HLR. For call setup the GSM system uses ISUP between exchanges and BSSAP as the access signalling between BSS and the Circuit Core. Due to mobility additional signalling functionality is required in the Core network. This is mainly provided by the Mobile Application Part. Making a call from an MS (Mobile Originated call) is quite similar to the wire line case. Terminating a call to a Mobile (Mobile Terminated call) is different. When an exchange sees that it has a mobile number for the callee, it will route the call to a gateway MSC. The GMSC will send a request using the SS7 signalling network and the Mobile Application Part signaling to the HLR. The HLR will return the Mobile Station Routing Number (MSRN) to the GMSC. The MSRN like its name tells is routable in all switching systems i.e. based on leading digits only. It was dynamically assigned by the VLR for the call or for the duration of the visit and given to the HLR to be used for terminating calls. When the call arrives to the visited MSC, the MSC finds out the location of MS in the VLR with the accuracy of several cells. The MSC will page the MS in all those cells. Paging means that the MSC sends a call signal to the mobile using signaling channels over the air in several cells simultaneously. The MS will respond using a signaling channel and the cell that it sees best. There is one more significant difference in call establishment in GSM as compared to wire line networks. Radio resources are seen as very expensive and should be preserved as much as possible. Therefore, when an MS makes a call, no radio resources are allocated at the originating cell until it is known that the callee is not busy, out of coverage and is also willing

to take the call. So, the reservation of timeslots of the air takes place later in the call establishment process.
From the very beginning GSM supported roaming. This means that subscribers of one operator can move into the area of another and make the use of the other operator's network. Usually, only international roaming is supported for business reasons. I.e. the two operators that allow their subscribers to use each other's networks cover different countries. International roaming has been a major benefit of GSM to the users and has helped to consolidate the GSM operators.

GSM frame structure

In GSM frequency band of 25 MHz is divided into 200 KHz of smaller bands, each carry one RF carrier, this gives 125 carriers. As one carrier is used as guard channel between GSM and other frequency bands 124 carriers are useful RF channels. This division of frequency pool is called FDMA. Now each RF carrier will have eight time slots. This division time wise is called TDMA. Here each RF carrier frequency is shared between 8 users hence in GSM system, the basic radio resource is a time slot with duration of about 577 microsec. As mentioned each time slot has 15/26 or 0.577ms of time duration. This time slot carries 156.25 bits which leads to bit rate of 270.833 kbps. This is explained below in TDMA gsm frame structure. For E-GSM number of ARFCNs are 174, for DCS1800 ARFNCs are 374. The GSM frame structure is designated as hyperframe, superframe, multiframe and frame. The minimum unit being frame (or TDMA frame) is made of 8 time slots. One GSM hyperframe composed of 2048 superframes. Each GSM superframe composed of multiframes (either 26 or 51 as described below). Each GSM multiframe composed of frames (either 51 or 26 based on multiframe type). Each frame composed of 8 time slots. Hence there will be total of 2715648 TDMA frames available in GSM and the same cycle continues. As shown in the figure 2 below, there are two varients to multiframe structure. 1. 26 frame multiframe - Called traffic multiframe, composed of 26 bursts in a duration of 120ms, out of these 24 are used for traffic, one for SACCH and one is not used. 2. 51 frame multiframe- Called control multiframe, composed of 51 bursts in a duration of 235.4 ms. This type of multiframe is divided into logical channels. These logical channels are time scheduled by BTS. Always occur at beacon frequency in time slot 0, it may also take up other time slots if required by system for example below

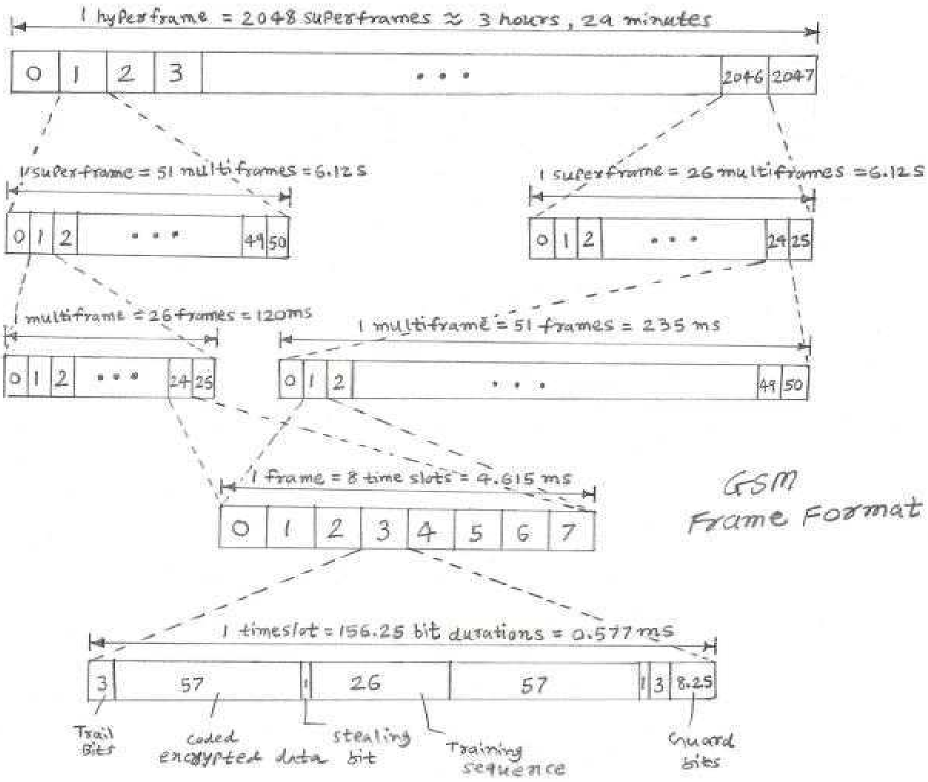


Fig 18. GSM Frame Structure

As shown in fig above each ARFCN or each channel in GSM will have 8 time slots TS0 to TS7. During network entry each GSM mobile phone is allocated one slot in downlink and one slot in uplink. Here in the figure GSM Mobile is allocated 890.2 MHz in the uplink and 935.2 MHz in the downlink. As mentioned TS0 is allocated which follows either 51 or 26 frame multiframe structure. Hence if at start 'F' is depicted which is FCCH after 4.615 ms (which is 7 time slot duration) S(SCH) will appear then after another 7 slots B(BCCH) will appear and so on till end of 51 frame Multiframe structure is completed and cycle continues as long as connection between Mobile and base station is active. similarly in the uplink, 26 frame multiframe structure follow, where T is TCH/FS (Traffic channel for full rate speech), and S is SACCH. The gsm frame structure can best be understood as depicted in the figure below with respect to downlink(BTS to MS) and uplink (MS to BTS) directions.

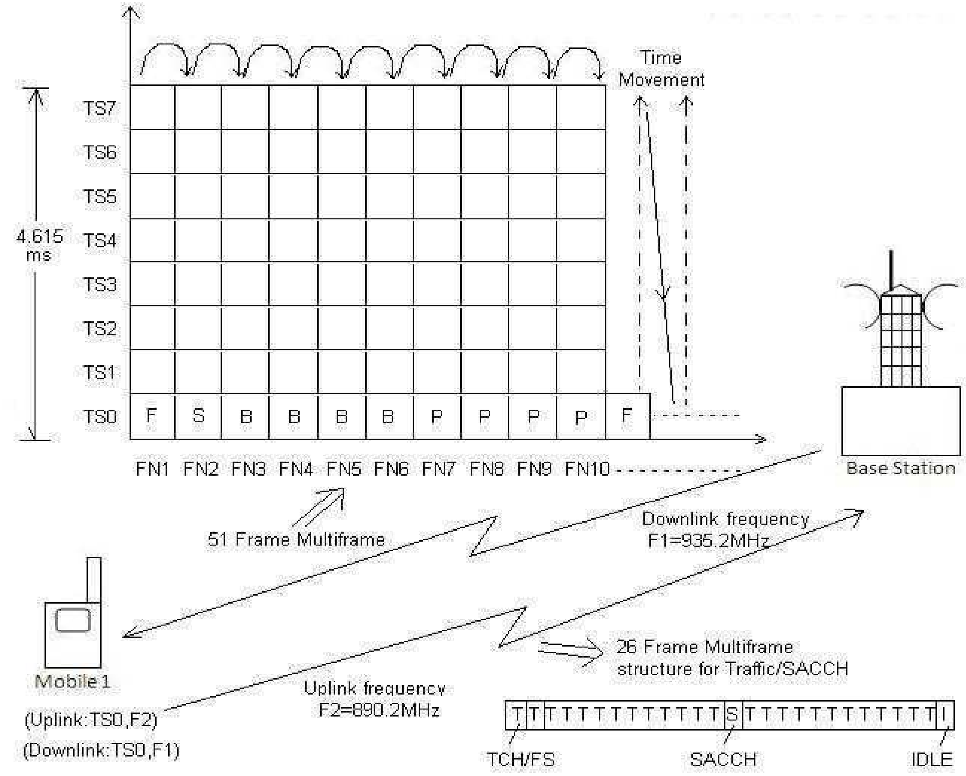


Fig 19. GSM Physical and logical channel concept

Frequencies in the uplink = 890.2 + 0.2 (N-1) MHz
Frequencies in the downlink = 935.2 + 0.2 (N-1) MHz
where, N is from 1 to 124 called ARFCN
As same antenna is used for transmit as well as receive, there is 3 time slots delay introduced between TS0 of uplink and TSO of downlink frequency. This helps avoid need of imultaneous transmission and reception by GSM mobile phone. The 3 slot time period is used by the Mobile subscriber to perform various functions e.g. processing data, measuring signal quality of neighbour cells etc. Engineers working in GSM should know GSM frame structure for both the downlink as well as uplink. They should also understand mapping of different channels to time slots in these GSM frame structures.

Traffic Routing in Wireless Networks

When there are many devices, it is necessary to develop suitable echanism for communication between any two devices. One a lternative is to establish point-to-point communication between each pair of devices using mesh topology. However, mesh topology is impractical for larg

e number of devices, because the number of links increases exponentially (n(n-1)/2, where n is the number of devices) with the number of devices. A better alternative is to use switching techniques leading to switched communication network. In the switched network methodology, the network consists of a set of interconnected nodes, among which information is transmitted from source to destination via different routes, which is controlled by the switching mechanism. A basic model of a

switched communication is shown in Fig. below. The end devices that wish to communicate with each other are called stations . The switching evices are called nodes . Some nodes

connect to other nodes and some are to connected to some stations . Key features of a switched communication network are given below:
Network Topology is not regular.
Uses FDM or TDM for node-to-node communication.
There exist multiple paths between a source-destination pair for better network reliability. The switching nodes are not concerned with the contents of data. Their purpose is to provide a switching facility that will move data from node to node until they reach the destination. The switching performed by different nodes can be categorized into the following three types:

Circuit Switching

Packet Switching

Message Switching

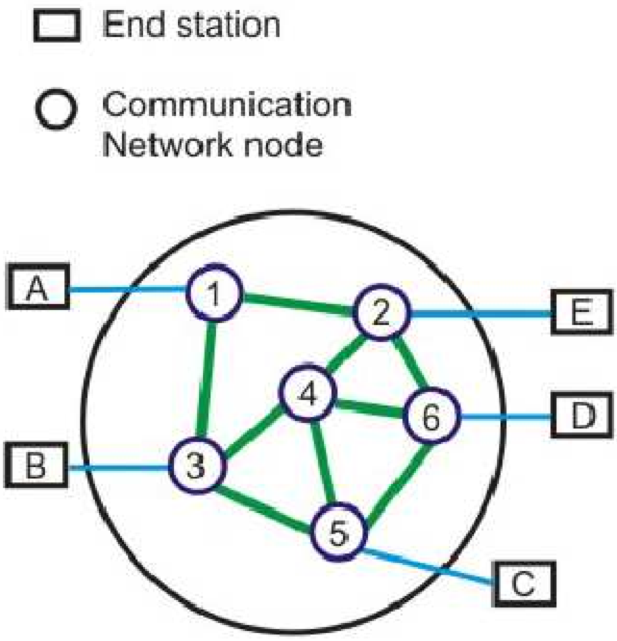


Figure20 : Basic model of a switched communication network

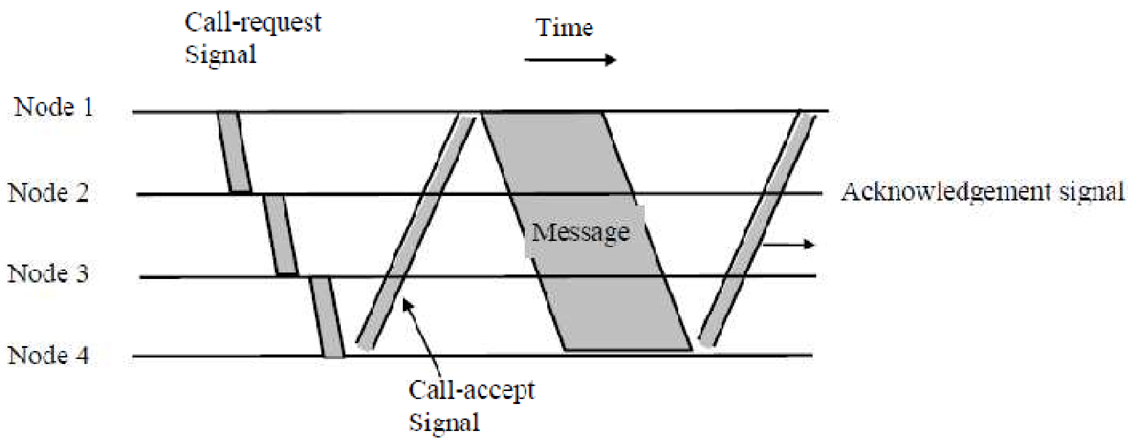
Circuit switching Technique

Communication via circuit switching implies that there is a dedicated communication path between the two stations. The path is a connected through a sequence of links between network nodes. On each physical link, a logical channel is dedicated to the connection. Circuit switching is commonly used technique in telephony, where the caller sends a special message with the address of the callee (i.e. by dialling a number) to state its destination. It involved the following three distinct steps, as shown in Fig

Circuit Establishment: To establish an end-to-end connection before any transfer of data.

Some segments of the circuit may be a dedicated link, while some other segments may be shared. *Data transfer:* Transfer data is from the source to the destination. The data may be analog or digital, depending on the nature of the network. The connection is generally full-duplex. Circuit disconnect: Terminate connection at the end of data transfer.

• Signals must be propagated to deallocate the dedicated resources.



Thus the actual physical electrical path or circuit between the source and destination host must be established before the message is transmitted. This connection, once established, remains exclusive and continuous for the complete duration of information exchange and the circuit becomes disconnected only when the source wants to do so.

Switching Node

Let us consider the operation of a single circuit switched node comprising a collection of stations attached to a central switching unit, which establishes a dedicated path between any two devices that wish to communicate.

Major elements of a single-node network are summarized below:

Digital switch: That provides a transparent (full-duplex) signal path between any pair of attached devices.

Network interface: That represents the functions and hardware needed to connect digital devices to the network (like telephones).

Control unit: That establishes, maintains, and tears down a connection.

The simplified schematic diagram of a switching node is shown in Fig. An important characteristic of a circuit-switch node is whether it is *blocking* or *non-blocking*. A blocking network is one, which may be unable to connect two stations because all possible paths between them are already in use. A non-blocking network permits all stations to be connected (in pairs) at once and grants all possible connection requests as long as the called party is free. For a network that supports only voice traffic, a blocking configuration may be acceptable, since most phone calls are of short duration. For data applications, where a connection may remain active for hours, non-blocking configuration is desirable.

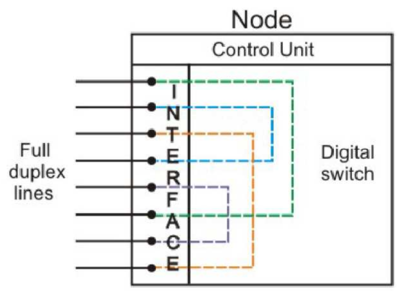
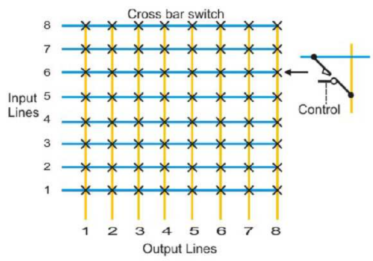


Fig 22:

Circuit switching uses any of the three technologies: Space-division switches, Time-division switches or a combination of both. In Space-division switching, the paths in the circuit are separated with each other spatially, i.e. different ongoing connections, at a same instant of time, uses different switching paths, which are separated spatially. This was originally developed for the analog environment, and has been carried over to the digital domain. Some of the space switches are crossbar switches, Multi-stage switches (e.g. Omega Switches). A crossbar switch is shown in Fig. 4.1.4. Basic building block of the switch is a metallic crosspoint or semiconductor gate that can be enabled or disabled by a control unit.



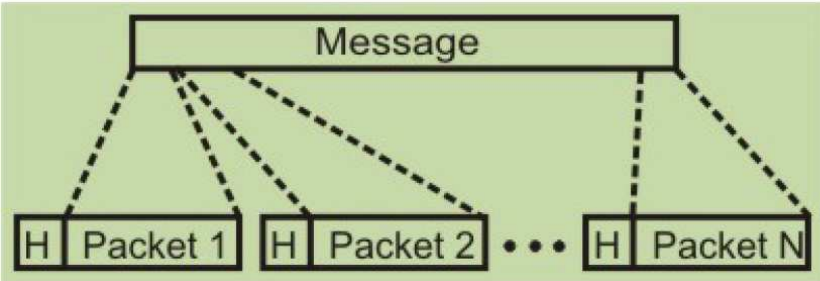
In circuit switching, network resources are dedicated to a particular connection. Although this satisfies the requirement of voice communication, it suffers from the following two shortcomings for data communication:

o In a typical user/host data connection, line utilization is very low.
o Provides facility for data transmission at a constant rate.

However, for information transmission applications, the circuit switching method is very slow, relatively expensive and inefficient. First of all, the need to establish a dedicated connection before sending the message itself inserts a delay time, which might become significant for the total message transfer time. Moreover, the total channel remains idle and unavailable to the other users once a connection is made. On the other hand once a connection is established, it is guaranteed and orderly delivery of message is ensured. Unfortunately, the data transmission pattern may not ensure this, because data transmission is bursty in nature. As a consequence, it limits the utility of the method To overcome the limitations of message switching, another switching technique, known as packet switching was invented

Packet Switching

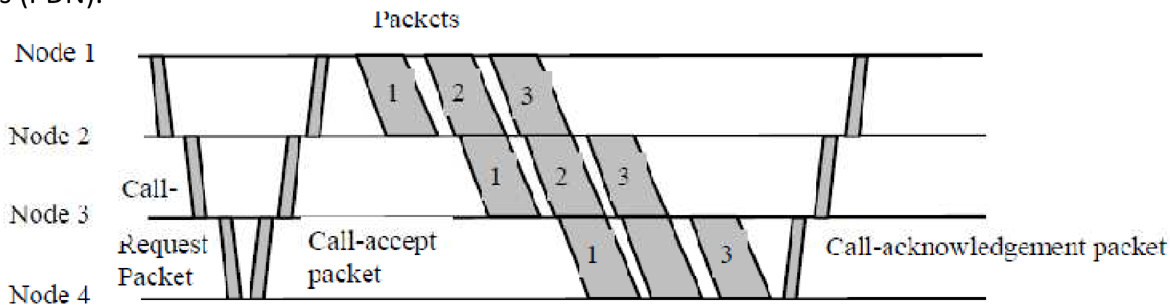
The basic approach is not much different from message switching. It is also based on the same ‘store-and-forward’ approach. However, to overcome the limitations of message switching, messages are divided into subsets of equal length called packets. This approach was developed for long-distance data communication (1970) and it has evolved over time. In packet switching approach, data are transmitted in short packets (few Kbytes). A long message is broken up into a series of packets as shown in Fig. 4.2.2. Every packet contains some control information in its header, which is required for routing and other purposes.



Main difference between Packet switching and Circuit Switching is that the communication lines are not dedicated to passing messages from the source to the destination. In Packet Switching, different messages (and even different packets) can pass through different routes, and when there is a "dead time" in the communication between the source and the destination, the lines can be used by other sources. There are two basic approaches commonly used to packet Switching: virtual-circuit packet switching and datagram packet switching. In virtual-circuit packet switching a virtual circuit is made before actual data is transmitted, but it is different from circuit switching in a sense that in circuit switching the call accept signal comes only from the final destination to the source while in case of virtual-packet switching this call accept signal is transmitted between each adjacent intermediate node as shown in Fig. Other features of virtual circuit packet switching are discussed in the following subsection.

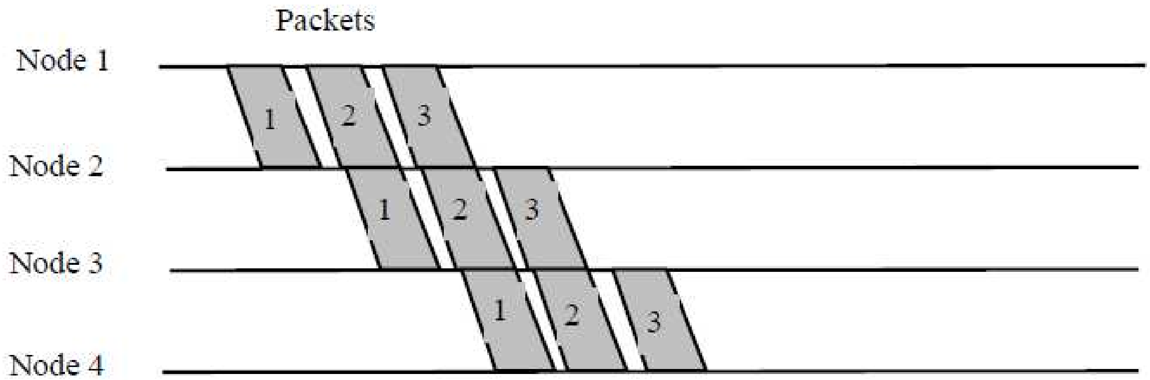
Virtual Circuit Packet Switching Networks

An initial setup phase is used to set up a route between the intermediate nodes for all the packets passed during the session between the two end nodes. In each intermediate node, an entry is registered in a table to indicate the route for the connection that has been set up. Thus, packets passed through this route, can have short headers, containing only a *virtual circuit identifier* (VCI), and not their destination. Each intermediate node passes the packets according to the information that was stored in it, in the setup phase. In this way, packets arrive at the destination in the correct sequence, and it is guaranteed that essentially there will not be errors. This approach is slower than Circuit Switching, since different virtual circuits may compete over the same resources, and an initial setup phase is needed to initiate the circuit. As in Circuit Switching, if an intermediate node fails, all virtual circuits that pass through it are lost. The most common forms of Virtual Circuit networks are X.25 and Frame Relay, which are commonly used for public data networks (PDN).



Datagram Packet Switching Networks

This approach uses a different, more dynamic scheme, to determine the route through the network links. Each packet is treated as an independent entity, and its header contains full information about the destination of the packet. The intermediate nodes examine the header of the packet, and decide to which node to send the packet so that it will reach its destination. In the decision two factors are taken into account: The shortest ways to pass the packet to its destination - protocols such as RIP/OSPF are used to determine the shortest path to the destination. Finding a free node to pass the packet to - in this way, bottlenecks are eliminated, since packets can reach the destination in alternate routes. Thus, in this method, the packets don't follow a pre-established route, and the intermediate nodes (the routers) don't have pre-defined knowledge of the routes that the packets should be passed through.



Packets can follow different routes to the destination, and delivery is not guaranteed (although packets usually do follow the same route, and are reliably sent). Due to the nature of this method, the packets can reach the destination in a different order than they were sent, thus they must be sorted at the destination to form the original message. This approach is time consuming since every router has to decide where to send each packet. The main implementation of Datagram Switching network is the Internet, which uses the IP network protocol.

The X .25 Protocol :

X.25 is an ITU-T standard protocol suite for packet switched wide area network (WAN) communication. An X.25 WAN consists of packet-switching exchange (PSE) nodes as the networking hardware, and leased lines, plain old telephone service connections or ISDN connections as physical links. X.25 is a family of protocols that was popular during the 1980s with telecommunications companies and in financial transaction systems such as automated teller machines. X.25 was originally defined by the International Telegraph and Telephone Consultative Committee (CCITT, now ITU-T) in a series of drafts and finalized in a publication known as The Orange Book in 1976. While X.25 has, to a large extent, been replaced by less complex protocols, especially the Internet protocol (IP), the service is still used and available in niche and legacy applications

Architecture

The general concept of X.25 was to create a universal and global packet-switched network. Much of the X.25 system is a description of the rigorous error correction needed to achieve this, as well as more efficient sharing of capital-intensive physical resources. The X.25 specification defines only the interface between a subscriber (DTE) and an X.25 network (DCE). X.75, a very similar protocol to X.25, defines the interface between two X.25 networks to allow connections to traverse two or more networks. X.25 does not specify how the network operates internally—many X.25 network implementations used something very similar to X.25 or X.75 internally, but others used quite different protocols internally. The ISO equivalent protocol to X.25, ISO 8208, is compatible with X.25, but additionally includes provision for two X.25 DTEs to be directly connected to each other with no network in between. By separating the Packet-Layer Protocol, ISO 8208 permits operation over additional networks such as ISO 8802 LLC2 (ISO LAN) and the OSI data link layer. X.25 originally defined three basic protocol levels or architectural layers. In the original specifications these were referred to as levels and also had a level number, whereas all ITU-T X.25 recommendations and ISO 8208 standards released after 1984 refer to them as layers. The layer numbers were dropped to avoid confusion with the OSI Model layers. Physical layer: This layer specifies the physical, electrical, functional and procedural characteristics to control the physical link between a DTE and a DCE. Common implementations use X.21, EIA-232, EIA-449 or other serial protocols.

- Data link layer: The data link layer consists of the link access procedure for data interchange on the link between a DTE and a DCE. In its implementation, the Link Access Procedure, Balanced (LAPB) is a data link protocol that manages a communication session and controls the packet framing. It is a bit-oriented protocol that provides error correction and orderly delivery.
- Packet layer: This layer defined a packet-layer protocol for exchanging control and user data packets to form a packet-switching network based on virtual calls, according to the Packet Layer Protocol.

The X.25 model was based on the traditional telephony concept of establishing reliable circuits through a shared network, but using software to create "virtual calls" through the network. These calls interconnect "data terminal equipment" (DTE) providing endpoints to users, which looked like point-to-point connections. Each endpoint can establish many separate virtual calls to different endpoints. For a brief period, the specification also included a connectionless datagram service, but this was dropped in the next revision. The "fast select with restricted response facility" is intermediate between full call establishment and connectionless communication. It is widely used in query-response transaction applications involving a single request and response limited to 128 bytes of data carried each way. The data is carried in an extended call request packet and the response is carried in an extended field of the call reject packet, with a connection never being fully established. Closely related to the X.25 protocol are the protocols to connect asynchronous devices (such as dumb terminals and printers) to an X.25 network: X.3, X.28 and X.29. This functionality was performed using a Packet Assembler/Disassembler or PAD (also known as a Triple-X device, referring to the three protocols used).

Error control

Error recovery procedures at the packet layer assume that the data link layer is responsible for retransmitting data received in error. Packet layer error handling focuses on resynchronizing the information flow in calls, as well as clearing calls that have gone into unrecoverable states: Level 3 Reset packets, which re-initializes the flow on a virtual call (but does not break the virtual call).Restart packet, which clears down all virtual calls on the data link and resets all permanent virtual circuits on the data link.

Addressing and virtual circuits

An X.25 Modem once used to connect to the German Datex-P network.

X.25 supports two types of virtual circuits, virtual calls (VC) and permanent virtual circuits (PVC). Virtual calls are established on an as-needed basis. For example, a VC is established when a call is placed and turn down after the call is complete. VCs are established through a call establishment and clearing procedure. On the other hand, permanent virtual circuits are preconfigured into the network.[20] PVCs are seldom turn down and thus provide a dedicated connection between end points. VC may be established using X.121 addresses. The X.121 address consists of a three-digit data country code (DCC) plus a network digit, together forming the four-digit data network identification code (DNIC), followed by the national terminal number (NTN) of at most ten digits. Note the use of a single network digit, seemingly allowing for only 10 network carriers per country, but some countries are assigned more than one DCC to avoid this limitation. Networks often used fewer than the full NTN digits for routing, and made the spare digits available to the subscriber (sometimes called the sub-address) where they could be used to identify applications or for further routing on the subscribers networks. NSAP addressing facility was added in the X.25(1984) revision of the specification, and this enabled X.25 to better meet the requirements of OSI Connection Oriented Network Service (CONS).[21] Public X.25 networks were not required to make use of NSAP addressing, but, to support OSI CONS, were required to carry the NSAP addresses and other ITU-T specified DTE facilities transparently from DTE to DTE. Later revisions allowed multiple addresses in addition to X.121 addresses to be carried on the same DTE-DCE interface: Telex addressing (F.69), PSTN addressing (E.163), ISDN addressing (E.164), Internet Protocol addresses (IANA ICP), and local IEEE 802.2 MAC addresses. VCs are permanently established in the network and therefore do not require the use of addresses for call setup. PVCs are identified at the subscriber interface by their logical channel identifier (see below). However, in practice not many of the national X.25 networks supported PVCs.One DTE-DCE interface to an X.25 network has a maximum of 4095 logical channels on which it is allowed to establish virtual calls and permanent virtual circuitsalthough networks are not expected to support a full 4095 virtual circuits. For identifying the channel to which a packet is associated, each packet contains a 12 bit logical channel identifier made up of an 8-bit logical channel number and a 4-bit logical channel group number. Logical channel identifiers remain assigned to a virtual circuit for the duration of the connection. Logical channel identifiers identify a specific logical channel between the DTE (subscriber appliance) and the DCE (network), and only has local significance on the link between the subscriber and the network. The other end of the connection at the remote DTE is likely to have assigned a different logical channel identifier. The range of possible logical channels is split into 4 groups: channels assigned to permanent virtual circuits, assigned to incoming virtual calls, two-way (incoming or outgoing) virtual calls, and outgoing virtual calls. (Directions refer to the direction of virtual call initiation as viewed by the DTE—they all carry data in both directions.). The ranges allowed a subscriber to be configured to handle significantly differing numbers of calls in each direction while reserving some channels for calls in one direction. All International networks are required to implement support for permanent virtual circuits, two-way logical channels and one-way logical channels outgoing; one-way logical channels incoming is an additional optional facility. DTE-DCE interfaces are not required to support more than one logical channel. Logical channel identifier zero will not be assigned to a permanent virtual circuit or virtual call. The logical channel identifier of zero is used for packets which don't relate to a specific virtual circuit (e.g. packet layer restart, registration, and diagnostic packets).lower price-per-segment than VCs, making them cheaper only where large volumes of data are passed.

X.25 packet types

Packet Type	DCE → DTE	DTE → DCE	Service VC	PVC
Call setup and Clearing	Incoming Call	Call Request	X	
	Call Connected	Call Accepted	X	
	Clear Indication	Clear Request	X	
	Clear Confirmation	Clear Confirmation	X	
Data and Interrupt	Data	Data	X	X
	Interrupt	Interrupt	X	X
	Interrupt Confirmation	Interrupt Confirmation	X	X
Flow Control and Reset	RR	RR	X	X

	RNR	RNR		X	X
	REJ	REJ		X	X
	Reset Indication	Reset Request		X	X
	Reset Confirmation	Reset Confirmation		X	X
Restart	Restart Indication	Restart Request	X		
	Restart Confirmation	Restart Confirmation	X		
Diagnostic	Diagnostic		X		
Registration	Registration Confirmation	Registration Request	X		

X.25 details

The network may allow the selection of the maximal length in range 16 to 4096 octets (2^n values only) per virtual circuit by negotiation as part of the call setup procedure. The maximal length may be different at the two ends of the virtual circuit.

- Data terminal equipment constructs control packets which are encapsulated into data packets. The packets are sent to the data circuit-terminating equipment, using LAPB Protocol.
- Data circuit-terminating equipment strips the layer-2 headers in order to encapsulate packets to the internal network protocol.

X.25 facilities

X.25 provides a set of user facilities defined and described in ITU-T Recommendation X.2.[35] The X.2 user facilities fall into five categories:

- essential facilities;
- additional facilities;
- conditional facilities;
- mandatory facilities; and,
- optional facilities.

X.25 also provides X.25 and ITU-T specified DTE optional user facilities defined and described in ITU-T Recommendation X.7.[36] The X.7 optional user facilities fall into four categories of user facilities that require:

- subscription only;
- subscription followed by dynamic invocation;
- subscription or dynamic invocation; and,
- dynamic invocation only.