Cellular Systems

Lal C. Godara
Australian Defence Force Academy, School of Electrical Engineering University College, University of New South Wales, Canberra, Australian Capital Territory, Australia

Abstract
Cellular technology is a burgeoning field in wireless communications. Each cell contains a base station that communicates with mobiles in the cell by transmitting and receiving signals on radio links.

CELLULAR FUNDAMENTALS
The area served by mobile phone systems is divided into small areas known as cells. Each cell contains a base station that communicates with mobiles in the cell by transmitting and receiving signals on radio links. The transmission from the base station to a mobile is typically referred to as downstream, forwardlink, or downlink. The corresponding terms for the transmission from a mobile to the base are upstream, reverse-link, and uplink. Each base station is associated with a mobile switching center (MSC) that connects calls to and from the base to mobiles in other cells and the public switched telephone network. A typical setup depicting a group of base stations to a switching center is shown in Fig. 1. In this section, terminology associated with cellular systems is introduced with a brief description to understand how these systems work.

COMMUNICATION USING BASE STATIONS
A base station communicates with mobiles using two types of radio channels, control channels to carry control information and traffic channels to carry messages. Each base station continuously transmits control information on its control channels. When a mobile is switched on, it scans the control channels and tunes to a channel with the strongest signal. This normally would come from the base station located in the cell in which the mobile is located. The mobile exchanges identification information with the base station and establishes the authorization to use the network. At this stage, the mobile is ready to initiate and receive a call.

A Call from a Mobile
When a mobile wants to initiate a call, it sends the required number to the base. The base station sends this information to the switching center that assigns a traffic channel to this call because the control channels are only used for control information. Once the traffic channel is assigned, this information is relayed to the mobile via the base station. The mobile switches itself to this channel. The switching center then completes the rest of the call.

A Call to a Mobile
When someone calls a mobile, the call arrives at the MSC. It then sends a paging message through several base stations. A mobile tuned to a control channel detects its number in the paging message and responds by sending a response signal to the nearby base station. The base station informs the switching center about the location of the desired mobile. The switching center assigns a traffic channel to this call and relays this information to the mobile via the base. The mobile switches itself to the traffic channel and the call is complete.

Registration
A mobile is normally located by transmitting a paging message from various base stations. When a large number of base stations are involved in the paging process, it becomes impractical and costly. It is avoided by a registration procedure where a roaming phone registers with an MSC closer to itself. This information may be stored with the switching center of the area as well as the home switching center of the phone. The home base of the phone is the one where it is permanently registered. Once a call is received for this phone, its home switching center contacts the switching center where the phone is roaming. Paging in the vicinity of the previous known location helps to locate the phone. Once it responds, the call may be connected as discussed previously.
An understanding of propagation conditions and channel characteristics is important for an efficient use of a transmission medium. Attention is being given to understanding the propagation conditions where a mobile is to operate and many experiments have been conducted to model the channel characteristics. Many of these results could be found in review articles\cite{2–4} and references therein.

### Fading Channels

The signal arriving at a receiver is a combination of many components arriving from various directions as a result of multipath propagation. This depends on terrain conditions and local buildings and structures, causing the received signal power to fluctuate randomly as a function of distance. Fluctuations on the order of 20 dB are common within the distance of one wavelength ($\lambda$). This phenomenon is called fading. One may think this signal as a product of two variables.

The first component, also referred to as the short term fading component, changes faster than the second one and has a Rayleigh distribution. The second component is a long-term or slow-varying quantity and has lognormal distribution.\cite{5,6} In other words, the local mean varies slowly with lognormal distribution and the fast variation around the local mean has Rayleigh distribution.

A movement in a mobile receiver causes it to encounter fluctuations in the received power level. The rate at which this happens is referred to as the fading rate in mobile communication literature\cite{7} and it depends on the frequency of transmission and the speed of the mobile. For example, a mobile on foot operating at 900 MHz would cause a fading rate of about 4.5 Hz, whereas a typical vehicle mobile would produce the fading rate of about 70 Hz.

### Doppler Spread

The movement in a mobile causes the received frequency to differ from the transmitted frequency because of the Doppler shift resulting from its relative motion. As the received signals arrive along many paths, the relative velocity of the mobile with respect to various components of the signal differs, causing the different components to yield a different Doppler shift. This can be viewed as spreading of the transmitted frequency and is referred to as the Doppler spread. The width of the Doppler spread in frequency domain is closely related to the rate of fluctuations in the observed signal.\cite{2}

### Delay Spread

Because of the multipath nature of propagation in the area where a mobile is being used, it receives multiple and delayed copies of the same transmission, resulting in spreading of the signal in time. The rms delay spread may range from a fraction of a microsecond in urban areas to on the order of 100 $\mu$s in a hilly area, and this restricts the maximum signal bandwidth between 40 and 250 kHz. This bandwidth is known as coherence bandwidth. The coherence bandwidth is inversely proportional to the rms delay spread. This is the bandwidth over which the channel is flat; i.e., it has a constant gain and linear phase.

For a signal bandwidth above the coherence bandwidth, the channel loses its constant gain and linear phase characteristic and becomes frequency selective. Roughly speaking, a channel becomes frequency selective when the rms delay spread is larger than the symbol duration and causes inter-symbol interference (ISI) in digital communications. Frequency-selective channels are also known as dispersive channels, whereas the nondispersive channels are referred to as flat-fading channels.

### Link Budget and Path Loss

Link budget is a name given to the process of estimating the power at the receiver site for a microwave link taking into account the attenuation caused by the distance between the transmitter and the receiver. This reduction is referred to as the path loss. In free space the path loss is proportional to the second power of the distance; i.e., the distance power gradient is two. In other words, by doubling the distance between the transmitter and the receiver, the received power at the receiver reduces to one-fourth of the original amount.
For a mobile communication environment utilizing fading channels the distance power gradient varies and depends on the propagation conditions. Experimental results show that it ranges from a value lower than two in indoor areas with large corridors to as high as six in metal buildings. For urban areas the path loss between the base and the cell site is often taken to vary as the fourth power of the distance between the two.\textsuperscript{[3]}

Normal calculation of link budget is done by calculating carrier to noise ratio (CNR), where noise consists of background and thermal noise, and the system utility is limited by the amount of this noise. However, in mobile communication systems, the interference resulting from other mobile units is a dominant noise compared with the background and man made noise. For this reason these systems are limited by the amount of total interference present instead of the background noise as in the other case. In other words, the signal to interference ratio (SIR) is the limiting factor for a mobile communication system instead of the signal to noise ratio (SNR) as is the case for other communication systems. The calculation of link budget for such interference-limited systems involves calculating the carrier level, above the interference level contributed by all sources.\textsuperscript{[8]}

### MULTIPLE ACCESS SCHEMES

The available spectrum bandwidth is shared in a number of ways by various wireless radio links. The way in which this is done is referred to as a multiple access scheme. There are basically four principle schemes. These are frequency division multiple access (FDMA), time division multiple access (TDMA), CDMA, and space division multiple access (SDMA).\textsuperscript{[9–20]}

#### FDMA Scheme

In an FDMA scheme the available spectrum is divided into a number of frequency channels of certain bandwidth and individual calls use different frequency channels. All first generation cellular systems use this scheme.

#### TDMA Scheme

In a TDMA scheme, several calls share a frequency channel.\textsuperscript{[9]} The scheme is useful for digitized speech or other digital data. Each call is allocated a number of time slots based on its data rate within a frame for upstream as well as downstream. Apart from the user data, each time slot also carries other data for synchronization, guard times, and control information.

The transmission from base station to mobile is done in time division multiplexmode, whereas in the upstream direction, each mobile transmits in its own time slot. The overlap between different slots resulting from different propagation delay is prevented by using guard times and precise slot synchronization schemes.

The TDMA scheme is used along with the FDMA scheme because there are several frequency channels used in a cell. The traffic in two directions is separated either by using two separate frequency channels or by alternating in time. The two schemes are referred to as frequency division duplex (FDD) and time division duplex (TDD), respectively. The FDD scheme uses less bandwidth than TDD schemes use and does not require as precise synchronization of data flowing in two directions as that in the TDD method. The latter, however, is useful when flexible bandwidth allocation is required for upstream and downstream traffic.\textsuperscript{[9]}

#### CDMA Scheme

The CDMA scheme is a direct sequence (DS), spread-spectrum method. It uses linear modulation with wideband pseudonoise (PN) sequences to generate signals. These sequences, also known as codes, spread the spectrum of the modulating signal over a large bandwidth, simultaneously reducing the spectral density of the signal. Thus, various CDMA signals occupy the same bandwidth and appear as noise to each other. More details on DS spread-spectrum may be found in Pickholz et al.\textsuperscript{[16]}

In the CDMA scheme, each user is assigned an individual code at the time of call initiation. This code is used both for spreading the signal at the time of transmission and despreading the signal at the time of reception. Cellular systems using CDMA schemes use FDD, thus employing two frequency channels for forward and reverse links.

On forward-link a mobile transmits to all users synchronously and this preserves the orthogonality of various codes assigned to different users. The orthogonality, however, is not preserved between different components arriving from different paths in multipath situations.\textsuperscript{[14]} On reverse links each user transmits independently from other users because of their individual locations. Thus, the transmission on reverse link is asynchronous and the various signals are not necessarily orthogonal.

It should be noted that these PN sequences are designed to be orthogonal to each other. In other words, the cross correlation between different code sequences is zero and thus the signal modulated with one code appears to be orthogonal to a receiver using a different code if the orthogonality is preserved during the transmission. This is the case on forward-link and in the
absence of multipath the signal received by a mobile is not affected by signals transmitted by the base station to other mobiles.

On reverse link the situation is different. Signals arriving from different mobiles are not orthogonalized because of the asynchronous nature of transmission. This may cause a serious problem when the base station is trying to receive a weak signal from a distant mobile in the presence of a strong signal from a nearly mobile. This situation where a strong DS signal from a nearby mobile swamps a weak DS signal from a distant mobile and makes its detection difficult is known as the “near–far” problem. It is prevented by controlling the power transmitted from various mobiles such that the received signals at the base station are almost of equal strength. The power control is discussed in a later section.

The term wideband CDMA (WCDMA) is used when the spread bandwidth is more than the coherence bandwidth of the channel. Thus, over the spread bandwidth of DS-CDMA, the channel is frequency selective. On the other hand, the term narrowband CDMA is used when the channel encounters flat fading over the spread bandwidth. When a channel encounters frequency selective fading, over the spread bandwidth, a RAKE receiver may be employed to resolve the multipath component and combine them coherently to combat fading.

A WCDMA signal may be generated using multicarrier (MC) narrowband CDMA signals, each using different frequency channels. This composite MC-WCDMA scheme has a number of advantages over the single carrier WCDMA scheme. It not only is able to provide diversity enhancement over multipath fading channels but also does not require a contiguous spectrum as is the case for the single carrier WCDMA scheme. This helps to avoid frequency channels occupied by narrowband CDMA, by not transmitting MC-WCDMA signals over these channels. More details on these and other issues may be found in Milstein and references therein.

Comparison of Different Multiple Access Schemes

Each scheme has its advantages and disadvantages such as complexities of equipment design and robustness of system parameter variations. For example, a TDMA scheme not only requires complex time synchronization of different user data but also presents a challenge to design portable RF units that overcome the problem of a periodically pulsating power envelope caused by short duty cycles of each user terminal. It should be noted that when a TDMA frame consists of \( N \) users transmitting equal bit rates, the duty cycles of each user is \( 1/N \). TDMA also has a number of advantages.\(^9\)

1. A base station communicating with a number of users sharing a frequency channel only requires one set of common radio equipment.
2. The data rate, to and from each user, can easily be varied by changing the number of time slots allocated to the user as per the requirements.
3. It does not require as stringent power control as that of CDMA because its interuser interference is controlled by time slot and frequency channel allocations.
4. Its time slot structure is helpful in measuring the quality of alternative slots and frequency channels that could be used for mobile-assisted handoffs. Handoff is discussed in a later section.

It is argued in Kohno et al.\(^{14}\) that though there does not appear to be a single scheme that is the best for all situations, CDMA possesses characteristics that give it distinct advantages over others.

1. It is able to reject delayed multipath arrivals that fall outside the correlation interval of the PN sequence in use and thus reduces the multipath fading.
2. It has the ability to reduce the multipath fading by coherently combing different multipath components using a RAKE receiver.
3. In TDMA and FDMA systems, a frequency channel used in a cell is not used in adjacent cells to prevent cochannel interference. In a CDMA system it is possible to use the same frequency channel in adjacent cells and thus increase the system capacity.
4. The speech signal is inherently bursty because of the natural gaps during conversation. In FDMD and TDMA systems, once a channel (frequency and/or time slot) is allocated to a user, that channel cannot be used during non-activity periods. However, in CDMA systems the background noise is roughly the average of transmitted signals from all other users and thus a nonactive period in speech reduces the background noise. Hence, extra users may be accommodated without the loss of signal quality. This in turn increases the system capacity.

SDMA

The SDMA scheme, also referred to as space diversity, uses an array of antennas to provide control of space by providing virtual channels in angle domain.\(^{18}\) This scheme exploits the directivity and beam-shaping capability of an array of antennas to reduce cochannel interference. Thus, it is possible that by using this scheme, simultaneous calls in a cell could be established at the same carrier frequency. This helps to increase the capacity of a cellular system.
The scheme is based on the fact that a signal arriving from a distant source reaches different antennas in an array at different times as a result of their spatial distribution, and this delay is utilized to differentiate one or more users in one area from those in another area. The scheme allows an effective transmission to take place between a base station and a mobile without disturbing the transmission to other mobiles. Thus, it has the potential such that the shape of a cell may be changed dynamically to reflect the user movement instead of used fixed size cells. This arrangement then is able to create an extra dimension by providing dynamic control in space.\cite{19,20}

### CHANNEL REUSE

The generic term, \textit{channel}, is normally used to denote a frequency in FDMA system, a time slot in TDMA system, and a code in CDMA system or a combination of these in a mixed system. Two channels are different if they use different combinations of these at the same place. For example, two channels in a FDMA system use two different frequencies. Similarly, in TDMA system two separate time slots using the same frequency channel is considered two different channels. In that sense, for an allocated spectrum the number of channels in a system is limited. This limits the capacity of the system to sustain simultaneous calls and may only be increased by using each traffic channel to carry many calls simultaneously. Using the same channel again and again is one way of doing it. This is the concept of channel reuse.

The concept of channel reuse can be understood from Fig. 2 that shows a cluster of three cells. These cells use three separate sets of channels. This set is indicated by a letter. Thus, one cell uses set A, the other uses set B, and so on. In Fig. 2 this cluster of three cells is being repeated to indicate that three sets of channels are being reused in different cells. Fig. 3 shows a similar arrangement with cluster size of seven cells. Now let us see how this helps to increase the system capacity.

Assume there are a total of \(F\) channels in a system to be used over a given geographic area. Also assume that there are \(N\) cells in a cluster that use all the available channels. In the absence of channel reuse this cluster covers the whole area and the capacity of the system to sustain simultaneous calls is \(F\). Now if the cluster of \(N\) cells is repeated \(M\) times over the same area, then the system capacity increases to \(MF\) as each channel is used \(M\) times.

The number of cells in a cluster is referred to as the cluster size, the parameter \(1/N\) is referred to as the frequency reuse factor, and a system using a cluster size of \(N\) is sometimes also referred to as a system using \(N\) frequency reuse plan. The cluster size is an important parameter. For a given cell size, as the cluster size decreases, more clusters are required to cover the given area leading to more reuse of channels, and hence the system capacity increases. Theoretically, the maximum capacity is attained when cluster size is one, i.e., when all the available channels are reused in each cell. For hexagonal cell geometry, the cluster size can only have certain values. These are given by \(N i^2 + j^2 + ij\), where \(i\) and \(j\) are nonnegative integers.

The cells using the same set of channels are known as cochannel cells. For example, in Fig. 2, the cells using channels A are cochannel cells. The distance between cochannel cells is known as cochannel distance and the interference caused by the radiation from these cells is referred to as cochannel interference. For proper functioning of the system, this needs to be minimized by decreasing the power transmitted by mobiles and base stations in cochannel cells and increasing the cochannel distance. Because the transmitted power normally depends on the cell size, the minimization of cochannel interference requires a minimum cochannel distance; i.e., the distance cannot be smaller than this minimum distance.
In a cellular system of equal cell size, the cochannel interference is a function of a dimensionless parameter known as cochannel reuse ratio $Q$. This is a ratio of the cochannel distance $D$ and the cell radius $R$, i.e.,

$$Q = \frac{D}{R}$$

For hexagonal geometry,

$$Q = \sqrt{3}N$$

It follows from these equations that an increase in $Q$ increases the cochannel distance and thus minimizes the cochannel interference. On the other hand, a decrease in $Q$ decreases the cluster size $N$ and hence maximizes the system capacity. Thus, the selection of $Q$ is a tradeoff between the two parameters, namely, the system capacity and cochannel interferences. It should be noted that for proper functioning of the system, the signal to co-channel interference ratio should be above a certain minimum value.\[^{[21]}\]

### CELLULAR CONFIGURATION

A cellular system may be referred to as a macrocell, a microcell, or a picocell system depending on the size of cells. Some characteristics of these cellular structures are now described.

#### Macrocell System

A cellular system with its cell size of several kilometer is referred to as macrocell systems. Base stations of these systems transmit several watts of power from antennas mounted on high towers. Normally there is no line of sight (LOS) between the base station and mobiles and thus a typical received signal is a combination of various signals arriving from different directions. The received signal in these systems experience spreading of several microsecond because of the nature of propagation conditions.

#### Microcell Systems

As cells are split and their boundaries are redefined, their size becomes very small. At a radius less than about a kilometer, the system is referred to as a microcell system. In these systems a typical base station transmits less than 1 W of power from an antenna mounted at a few meter above the ground and normally an LOS exists between the base and a mobile. Cell radius in microcell systems is less than a kilometer giving rms delay spread on the order of few tens of nanosecond compared with a few microseconds for macrocell systems. This impacts the maximum data rate a channel could sustain. For microcell systems maximum bit rate is about 1 Mbps compared with that of about 300 Kbps for macrocell systems.\[^{[8]}\]

Microcell systems are also useful in providing coverage along roads and highways. Because the antenna height is normally lower than the surrounding buildings the propagation is along the streets and an LOS exists between the base and a mobile. When a mobile turns a corner, sometimes a sudden drop in received signal strength is experienced because of loss of LOS. Depending on how antennas are mounted on intersections and corners, various cell plans are possible. More details on these aspects may be found in Tripathi et al.\[^{[22]}\] and references therein.

#### Picocell Systems

When cell sizes are reduced below about 100 m covering areas such as large rooms, corridors, underground stations, large shopping centers, cellular systems are sometimes referred to as picocell systems with antennas mounted below rooftop levels or in buildings. These in-building areas have different propagation conditions than those covered by macrocell and microcell systems, and thus require different considerations for developing channel models. Details on various models to predict propagation conditions may be found in Fleury and Leuthold.\[^{[4]}\] Sometimes the picocell and microcell systems are also referred to as cordless communication systems with the term cellular identifying a macrocell system. Mobiles within these smaller cell systems are called cordless terminals or cordless phones.\[^{[23–25]}\]

Providing in-building communication services using wireless technology based on cell shapes dictated by floors and walls, is a feasible alternative and offers many advantages. It is argued in Pandya\[^{[24]}\] that RPs in 18-GHz band are ideal for such services because these do not penetrate concrete and steel structures, eliminating the problem of cochannel interferences. These frequencies offer huge bandwidth and require millimeter size antennas that are easy to manufacture and install.

#### Overlaid System

Small cell systems make very efficient use of the spectrum, allowing large frequency reuse resulting in an increased capacity of a system. However, these are not suitable for all conditions because of their large handoff requirement. A system of mixed cells with the concept of overlaying is discussed.\[^{[22,26–29]}\] In this system a hierarchy of cells is assumed to exist. A macrocell system is assumed at the top of the hierarchy with smaller cells systems at its bottom. A mobile with high mobility is assigned to a macrocell system, whereas the one with a
low mobility, to smaller cell systems. A design incorporating various combinations of different multiple access schemes reflects the ease of handoff and other traffic management strategies. A SDMA scheme has an important role to play in this concept, with various beams placed at the bottom of the hierarchy.

**CHANNEL ALLOCATION AND ASSIGNMENT**

Various multiple access schemes discussed in a previous section are used to divide a given spectrum into a set of disjoint channels. These channels are then allocated to various cells for their use. Channel allocation may be carried out using one of the three basic schemes, namely, fixed channel allocation, dynamic channel allocation, and hybrid channel allocation.[30]

**Fixed Channel Allocation Schemes**

In fixed channel allocation schemes, a number of channels are allocated to a cell permanently for its use such that these channels satisfy certain channel reuse constraints as discussed in the previous section. In its simplest form, the same number of channels are allocated to each cell. For a system with uniform traffic distribution across all cells, this uniform channel allocation scheme is efficient in the sense that the average call blocking probability in each cell is the same as that of the overall system. For systems where the distribution is not uniform, the call blocking probability differs from cell to cell, resulting in the call being blocked in some cells when there are spare channels available in other cells.

This situation could be improved by allocating channels nonuniformly as per the expected traffic in each cell or employing one of many prevailing channel borrowing schemes. One of these is referred to as a static borrowing scheme where some channels are borrowed from cells with light traffic and allocated to those with heavy traffic. Rearrangements of channels between cells are performed periodically to meet the variation in traffic load. In this scheme, the borrowed channels stay with the new cell until reallocated. There are other temporary borrowing schemes where a cell that has used all its channels is allowed to borrow a free channel from a neighbor provided it does not interfere with existing calls. The borrowed channel is returned to the original cell once the call is complete. Some temporary borrowing schemes allow any channel from a cell to be borrowed, whereas in others only nominated channels are allowed to be borrowed. Many borrowing strategies are available for selecting a channel, ranging from a simple scheme to pick the first available channel that satisfies the cochannel distance constraints to the one that performs an exhaustive search to select a channel that yields maximum SIR and minimizes the future probability of call blocking.

**Dynamic Channel Allocation Schemes**

Fixed channel allocation schemes discussed thus far are simple to implement and are generally useful for relatively stable traffic conditions. These schemes are not very efficient for fast changing user distribution because they are not designed to adapt to short-term variations in traffic conditions. Dynamic channel allocation schemes are most suited for such situations. In these schemes, channels are not allocated to various cells but are kept in a central pool, and are assigned to calls as they arrive. At the completion of a call, the assigned channel is released and goes back to the pool. The process of channel assignment involves working out a cost of assigning a channel to a call and a channel with the minimum cost is chosen for the purpose. The various channel assignment schemes differ in the way the cost function is selected using various parameters of interest such as reuse distance, SIR ratio, and probability of call blocking. Some schemes base their assignment only on the existing traffic conditions in the service area whereas the others take the past and the present conditions into account.

Dynamic channel assignment schemes may be implemented centrally where a central controller assigns the channels to calls from the pool. The central controller is able to achieve very efficient channel assignment but requires high overhead. The channel assignment may also be implemented in a distributed manner by base stations where calls are originated. The channel implementation by base stations requires less overhead than that required by a central controller and is more suitable for microcell systems. The distributed channel assignment schemes can be divided into two categories. In one case each base station keeps detailed status information about existing available channels in its neighborhood by exchanging status information with other base stations. The schemes in this category may provide near optimum allocation but pay a heavy price in terms of increased communication with other base stations, particularly in heavy traffic. The other category of distributed channel assignment schemes uses simple algorithms that rely on mobiles to measure signal strength to decide the suitability of a channel.

**Hybrid Channel Allocation Schemes**

The fixed channel allocation schemes are efficient under uniformly distributed heavy traffic. On the other hand, the dynamic channel allocation schemes perform better under low traffic conditions with varying and
nonuniformly distributed loads. The hybrid channel allocation schemes maximize advantages of both these schemes by dividing channels into fixed and dynamic sets. The channels in fixed sets are allocated as per fixed channel allocation strategies and those in the other set are free to be assigned to calls in a cell that has used all its allocated channels. The channels in this set are assigned as per the dynamic channel allocation procedures. Apparently, no optimum ratio of channels is assigned to two sets and the design parameter is dependent on local traffic conditions. More details on these and related issues may be found in Katzela and Naghsheh [30] and references therein.

**HANDOFF**

It is common for a mobile to move away from its servicing base station while a call is in progress. As the mobile approaches the cell boundary, the strength and quality of the signal it receives starts to deteriorate. At some stage, near the cell boundary, it receives a stronger signal from a neighboring base station than it does from its serving base station. At this point the control of the mobile is handed over to the new base station by assigning a channel belonging to the new cell. This process where a radio channel used by a mobile is changed, is referred to as handoff or handover. [22,25,29-32] When handoff is between two base stations as described earlier, it is referred to as intercell handoff. On the other hand, when handoff is between two channels belonging to the same base stations, it is referred to as intracell handoff. The situation arises when the network, while monitoring its channels, finds a free channel of better quality than that used by a mobile and decides to move the mobile to this new channel to improve the quality of channels in use. Sometimes, the network rearranges channels to avoid congestion and initiates intracell handoff.

Handoff is also necessary between different layers of overlaid systems consisting of microcells and macrocells. In these systems, the channels are divided into microcell channels and macrocell channels. When a mobile moves from one microcell to another and there is no available channel for handoff, a macrocell channel is used to meet the handoff request. This avoids the forced termination of a call. Later if a channel becomes available at an underlayered microcell, then the macrocell channel may be released and a microcell channel is assigned to the call by initiating a new handoff.

Forced termination of a call in progress is undesirable and to minimize it, a number of strategies are employed. These include reserving channels for handoff, using channel assignment schemes that give priority to a handoff request over new calls, and queuing the handoff request. The channel reservation and handoff priority scheme reduce the probability of forced termination by increasing the probability of blocking new calls. The queuing schemes are effective when handoff requests arrive in groups and there is a reasonable likelihood of channel availability in the near future.

The handoff is initiated when the quality of existing channels deteriorates below an acceptable threshold or a better channel is available. The channel quality is measured in terms of BER, received signal strength, or some other signal quality such as eye opening of radio signal that indicates signal to interference plus noise ratio.

For handoff initiation, the signal strength is used as an indication of the distance between the base and the mobile. For this reason, a drop in signal strength resulting from Rayleigh fading is normally not used to initiate handoff and some kind of averaging is used to avoid the problem. In some systems, the round trip delay between mobile and base is also used as an indication of the distance.

The measurement of various parameters may be carried out either at the mobile or at the base. Depending on where the measurements are made and who initiates the handoff, various handoff implementation schemes are possible including network-controlled handoff, mobile-controlled handoff, and mobile-assisted handoff.

**Network-Controlled Handoff**

In network-controlled handoff, each base station monitors the signal strength received from mobiles in their cells and makes periodic measurements of the received signal from mobiles in their neighboring cells. The MSC then initiates and completes the handoff of a mobile as and when it decides. The decision is based on the received signal strength at the base station serving the mobiles and base stations in neighboring cells. Because of its centralized nature, the collection of these measurements generates a large network traffic. This could be reduced to an extent by making measurements less frequently and by not requiring the neighboring base station to send the measurements continually. However, this reduces the accuracy. The execution of handoff by this method takes a few seconds and for this reason the method is not preferred by microcellular systems where a quick handoff is desirable.

**Mobile-Controlled Handoff**

Mobile-controlled handoff is a highly decentralized method and does not need any assistance from the MSC. In this scheme a mobile monitors the signal strength on its existing channel and measures signals...
received from the neighboring base stations. It receives BER and signal strength information from its serving base stations about uplink channels. Based on all this information, it initiates the handoff process by requesting the neighboring base for allocation of a low interference channel. The method has a handoff execution time on the order of 100 ms and is suitable for microcell systems.

Mobile-Assisted Handoff

In mobile-assisted handoff methods, as the name suggests, a mobile helps the network in the handoff decision making by monitoring the signal strength of its neighboring base stations and passing them to MSC via its serving base station. The handoff is initiated and completed by the network. The execution time is on the order of 1 second.

Hard Handoff and Soft Handoff

Handoff may be classified into hard handoff and soft handoff. During hard handoff the mobile can communicate only with one base station. The communication link gets broken with the existing base station before the new one is established and there is normally a small gap in communication during the transition. In the process of soft handoff, the mobile is able to communicate with more than one base station. It receives signals from more than one base station and the received signals are combined after appropriate delay adjustment. Similarly, more than one station receives signals from mobiles and the network combines different signals. This scheme is also known as macroscopic diversity and is mostly employed by CDMA systems.

Hard handoff, on the other hand, is more appropriate for TDMA and FDMA systems. It is also simple to implement compared with soft handoff. However, it may lead to unnecessary handoff back and forth between two base stations when the signals from two base stations fluctuate. The situation may arise when a mobile, presently being served, e.g., by Base 1, receives a stronger signal, from say Base 2 and is handed over to Base 2. Immediately after that it receives a stronger signal from Base 1 compared what it receives from Base 2, causing a handoff. This phenomenon, known as the ping-pong effect, may continue for some time and is undesirable because every handoff has a cost associated with it requiring network signaling of varying amount for authentication, database updates, circuit switching, and so on. This is avoided by using a hysteresis margin such that the handoff is not initiated until the difference between the signals received from the two base stations is more than the margin. For example, if the margin is $\Delta$ dB then the handoff is initiated when the signal received by the mobile from Base 2 is $\Delta$ dB more than that from Base 1. More details on various handoff implementation issues may be found in Tripathi et al.,[22] Noerpel and Lin,[31] and Tekinay and Jabbari[33] and references therein.

CELL SPLITTING AND CELL SECTORIZATION

Each cell has a limited channel capacity and thus could only serve so many mobiles at a given time. Once the demand in that cell exceeds this limit, the cell is further subdivided into smaller cells, each new cell with its own base station and its frequency allocation. The power of the base station transmitters is adjusted to reflect the new boundaries. The power transmitted by new base stations is less than that of the old one.

The consequence of the cell splitting is that the frequency assignment has to be done again, which affects the neighboring cells. It also increases the handoff rate because the cells are now smaller and a mobile is likely to cross cell boundaries more often compared with the case when the cells are big. Because of altered signaling conditions, this also affects the traffic in control channels.

Cell sectorization is referred to the case when a given cell is subdivided into several sectors and all sectors are served by the same base station. This is normally done by employing directional antennas such that the energy in each sector is directed by separate antennas. This has the effect of increased channel capacity similar to cell splitting. However, it uses the same base station and thus does not incur the cost of establishing new base stations associated with the cell splitting. This helps in reducing the cochannel interference because the energy is directed in the direction of the sector that does not cause interference in the cochannel cells, particularly in cochannel cells in the opposite direction to the sector. As in the case of cell splitting, this also affects the handoff rate.

POWER CONTROL

It is important that a radio receiver receives a power level that is enough for its proper function but not high enough for this level to disturb other receivers. This is achieved with maintaining constant power level at the receiver by transmitter power control. The receiver controls the power of the transmitter at the other end. For example, a base would control the power transmitted by mobile phones and vice versa.

It is done by a receiver monitoring its received power and sending a control signal to the transmitter to control
its power transmission as required. Sometimes a separate pilot signal is used for this purpose.

Power control reduces the near–far problem in CDMA systems and helps to minimize the interference near the cell boundaries when used in forwardlink.\textsuperscript{12,13}

REFERENCES


