UNIT I

SERVICES AND TECHNICAL CHALLENGES

Types of Services

TYPES AND SERVICES

- Analog Cellular Services
- GSM and Cellular Services
- Broadcast services
- Paging
- Cellular Telephony
- Trunking Radio
- Cordless Telephony
- Wireless Local Area Networks
- Personal Area Networks
- Fixed Wireless Access
- Ad hoc Networks and Sensor Networks
- Satellite Cellular Communications

Analog Cellular Services

The first generation of cellular systems used analog radio technology.Analog cellular systems consist of three basic elements:

- Mobile telephone (mobile radio),
- Cell sites,
- Mobile switching center (MSC).

- Figure 1.1 shows a basic cellular system in which a geographic service area such as a city is divided into smaller radio coverage area cells.
- A mobile telephone communicates by radio signals to the cell site within a radio coverage area.
- The cell sites base station (BS) converts these radio signals for transfer to the MSC via wired (landline) or wireless (microwave) communications links.
- The MSC routes the call to another mobile telephone in the system or the appropriate landline facility.
- These three elements are integrated to form a ubiquitous coverage radio system that can connect to the public switched telephone network (PSTN).

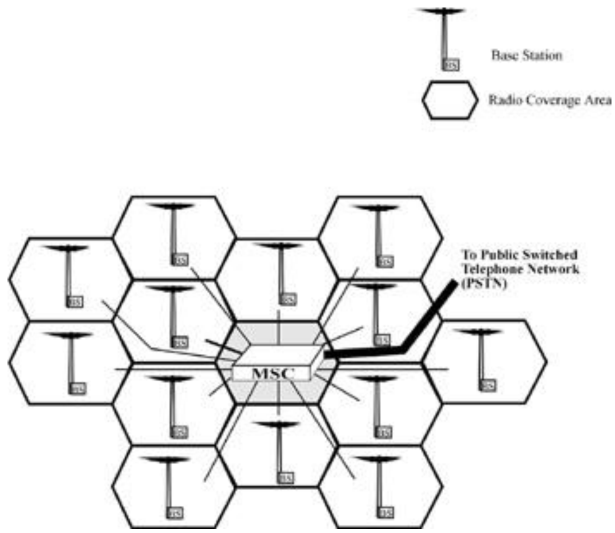


Figure 1.1: Basic cellular System

GSM and Cellular Services

GSM (Global System for Mobile Communications, originally *GroupeSpécial Mobile*), is a standard developed by the European Telecommunications Standards Institute (ETSI) to describe the protocols for second-generation (2G) digital cellular networks.

The network is structured into a number of discrete sections as shown in Figure 1.2

- Base Station Subsystem – the base stations and their controllers explained
- Network and Switching Subsystem – the part of the network most similar to a fixed network, sometimes just called the "core network"
- GPRS Core Network – the optional part which allows packet-based Internet connections
- Operations support system (OSS) – network maintenance

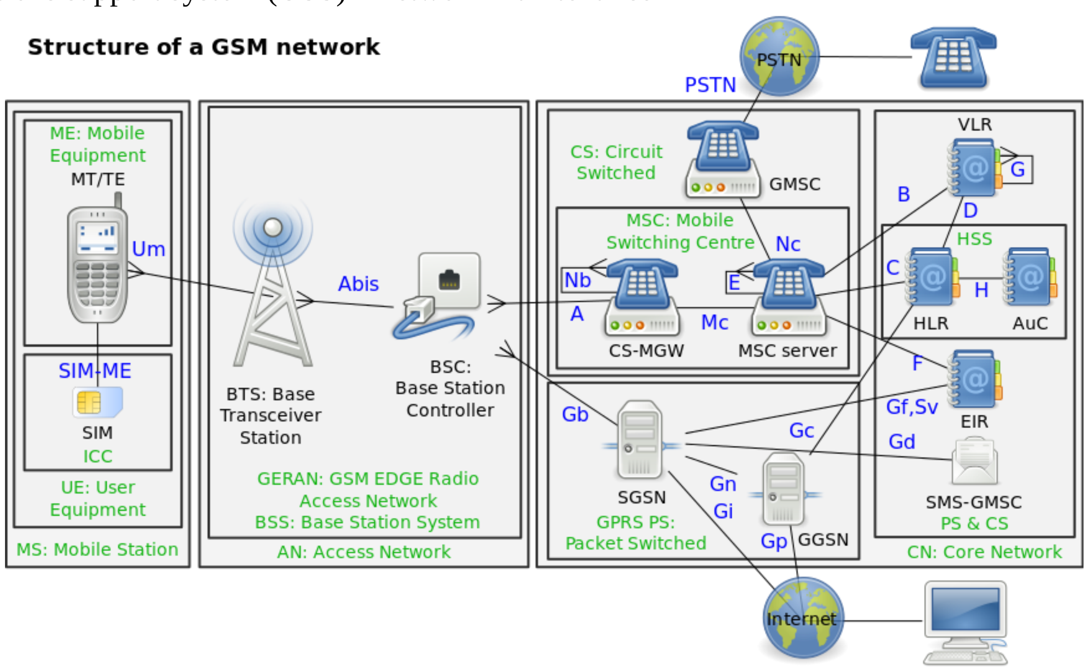


Figure 1.2 : Structure of GSM Network

Broadcast services

The first wireless service was broadcast radio. In this application, information is transmitted to infinite users.

The propertiesof broadcast radio are

1. The information is only sent in one direction.
2. It is only the broadcast station that sends information to the radio or TV receivers;
3. The listeners (or viewers) do not transmit any information back to the broadcast station.
4. The transmitted information is the same for all users.
5. The information is transmitted continuously.

The below Figure 1.3 represent the broadcast transmitter model where the receivers (users) are distributed randomly.

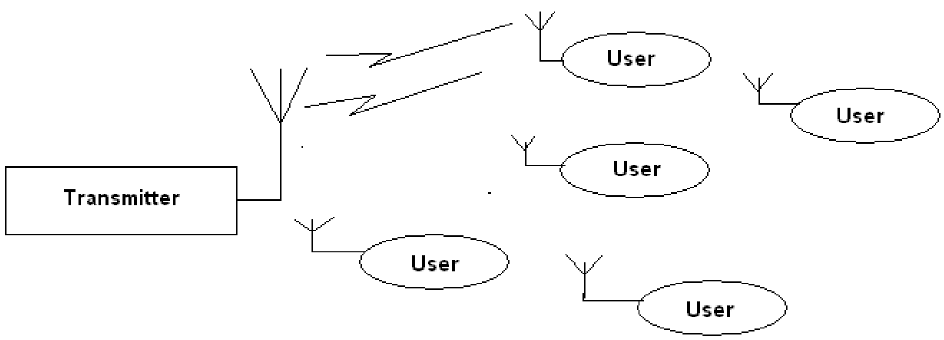


Figure 1.3: Broadcast Transmitter

Paging

Similar to broadcast, paging systems are unidirectional wireless communications systems. They are Characterized by the following properties.

- The user can onlyreceive information, but cannot transmit. Consequently, a “call” (message) can only be initiated by the call center, not by the user.

- The information is intended for, and received by, only a single user.

The amount of transmitted information is very small.

Originally, the received information Consisted of a single bit of information, which indicated to the user that “somebody has sent you a message.”

The user then had to make a phone call (usually from a payphone) to the call center, where a human operator repeated the content of the waiting message.

Now paging systems became more sophisticated, allowing the transmission of short messages (e.g., a different phone number that should be called, or the nature of an emergency).

Due to the unidirectional nature of the communications, and the small amount of information, the bandwidth required for this service is small.

This in turn allows the service to operate at lower carrier frequencies – e.g., 150MHz – where only small amounts of spectrum are available.

The below figure 1.4 represents the paging principle

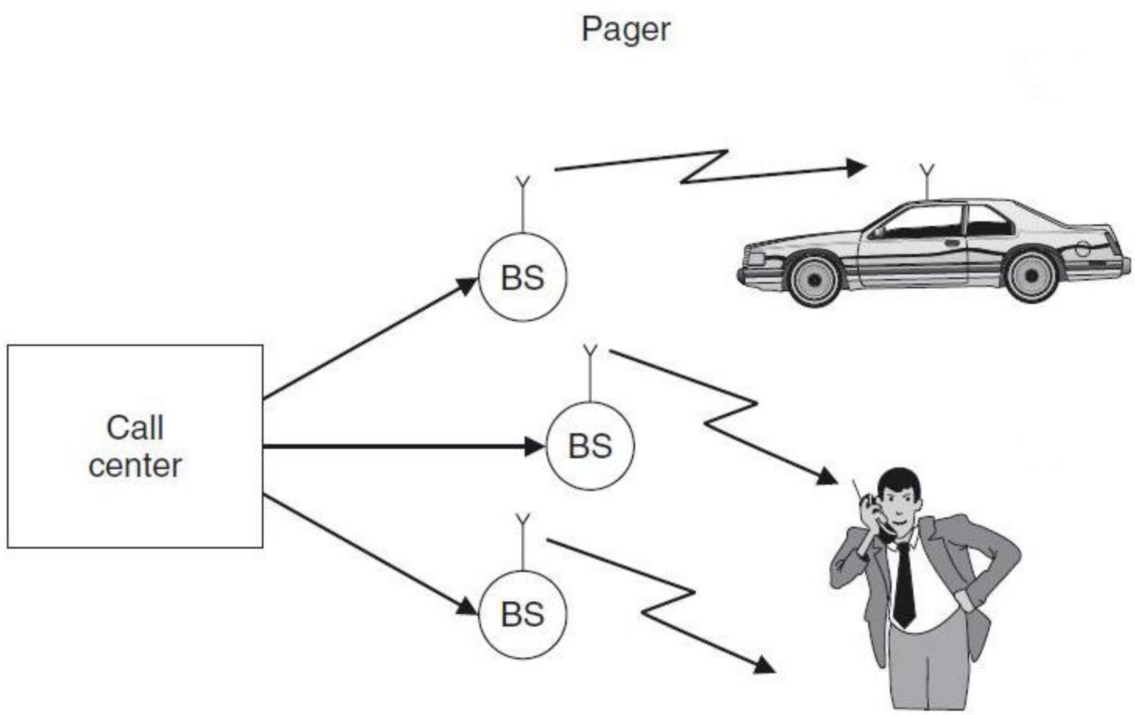


Figure 1.4 :Principle of Pager

Cellular Telephony

- In a cellular radio system, a land area to be covered with radio service is divided into regular shaped cells, which can be hexagonal, square, circular or some other regular shapes, although hexagonal cells are conventional.

- Each of these cells is assigned with multiple frequencies for example (f_1 to f_4) which have corresponding radio base stations. The group of frequencies can be reused in other cells, provided that the same frequencies are not reused in adjacent neighboring cells as that would cause co-channel interference.

- The key characteristic of a cellular network is the ability to re-use frequencies to increase both coverage and capacity. As described above, adjacent cells must use different frequencies; however there is no problem with two cells sufficiently far apart operating on the same frequency. Refer Figure 1.5.

- The working principle follows the analog cellular principle. The radio link will be like from Mobile switching center to base station and from base station to mobile phone and vise versa.

- In cities, each cell site may have a range of up to approximately ½ mile (0.80 km), while in rural areas, the range could be as much as 5 miles (8.0 km).

  Small cells, which have a smaller coverage area than base stations, are categorised as follows:

i) Microcell, (less than 2 Kilometres) ii) Picocell, (less than 200 Metres) iii) Femtocell, (around 10 Mts)

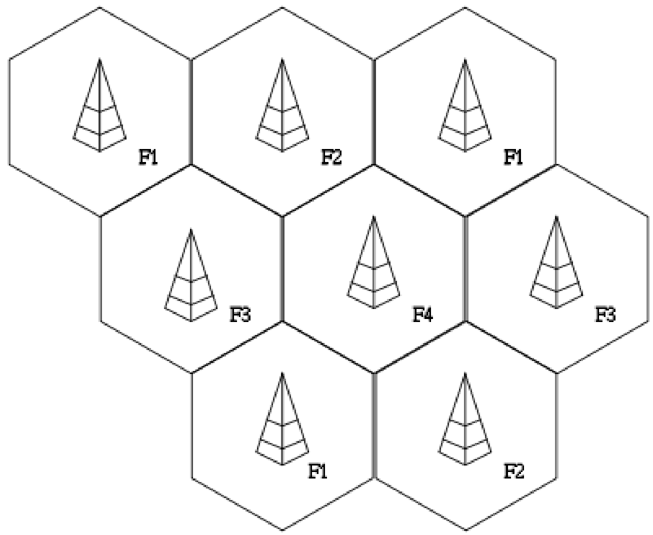


Figure 1.5: Frequency reuse concept and cell splitting

Hand over concept:

- As the phone user moves from one cell area to another cell while a call is in progress, the mobile station will search for a new channel to attach to in order not to drop the call.

- Once a new channel is found, the network will command the mobile unit to switch to the new channel and at the same time switch the call onto the new channel.

Trunking Radio

- Trunking is a cost effective radio communication technique that makes communication significantly easier and more effective than traditional radio communication methods systems such as ComReps (Community Repeaters),
- Here different users operate on separate radio frequencies or radio channels.
- Radio trunking smartly controls and guides the users toward a free channel, thereby reducing waiting times.
- The principal of trunking is based on automatic and dynamic allocation of a small number of radio channels among a large number of radio users in the most efficient and transparent way.
- Finding a free channel is therefore not the user’s obligation as the system does it automatically and in the shortest possible time.
- Trunking permits a large number of users to share a relatively small number of communication channels.Refer Figure 1.6 for Queuing operation and working model.

- This process is managed through tried and tested control equipment and the entire allocation process is transparent to the individual user.
- Trunking optimizes resources to reduce queues and speed up channel allocation.

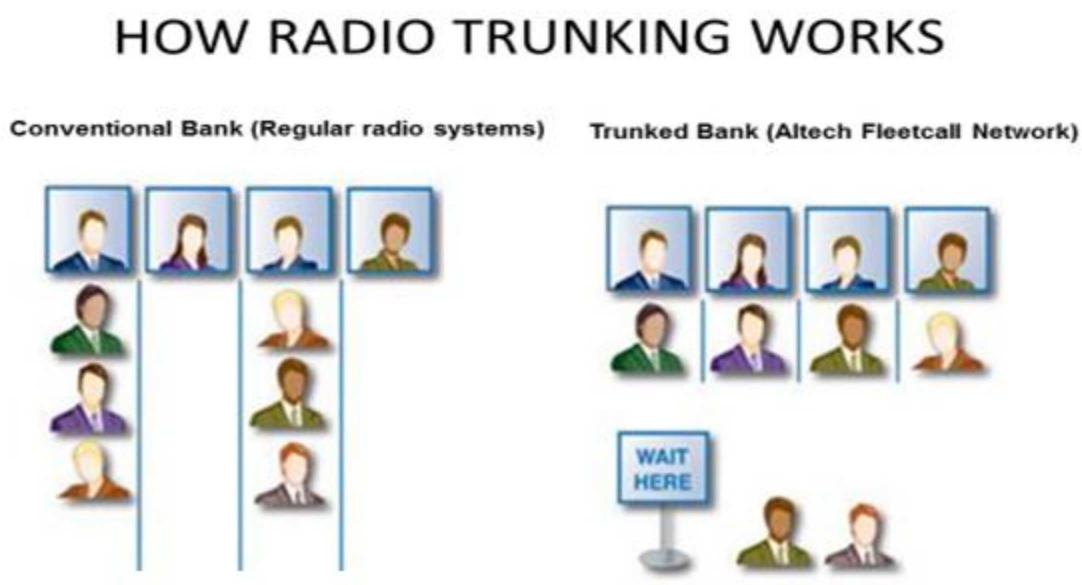


Figure 1.6: Queuing operation and working model.

Cordless Telephony

Cordless telephony describes a wireless link between a handset and a Base Set (BS) that is directly connected to the public telephone system. The block diagram representing cordless telephony is shown in figure 1.7. The main difference from a cellphone is that the cordless telephone is associated with, and can communicate only with a single BS.

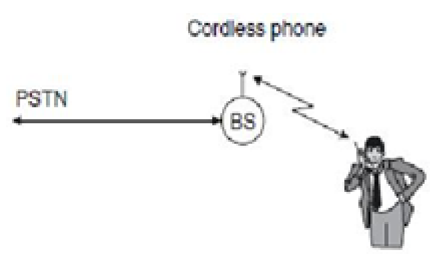


Figure 1.7: Cordless Telephony

Wireless Local Area Networks

- A wireless local area network (WLAN) is a local area network (LAN) that doesn't rely on wired Ethernet connections. A WLAN can be either an extension to a current wired network or an alternative to it.

- WLANs have data transfer speeds ranging from 1 to 54Mbps, with some manufacturers offering proprietary 108Mbps solutions. The 802.11n standard can reach 300 to 600Mbps.
- The wireless signal is broadcast hence everybody nearby can share it.The general set up is shown in Figure 1.8.
- Several security precautions are necessary to ensure only authorized users can access your WLAN.
- A WLAN signal can be broadcast to cover an area ranging in size from a small office to a large campus.
- Most commonly, a WLAN access point provides access within a radius of 65 to 300 feet.

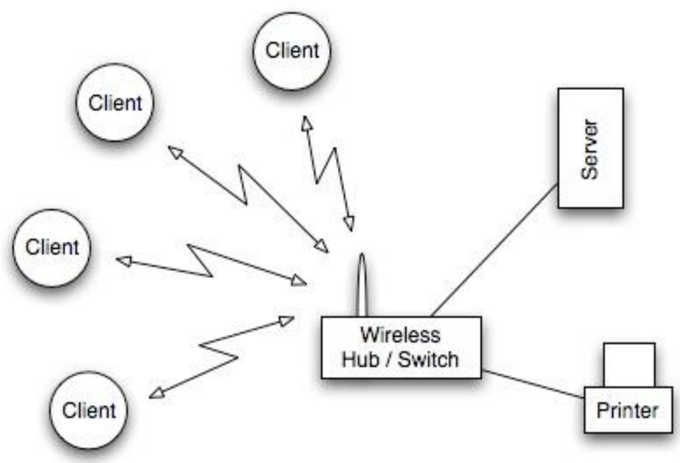


Figure 1.7 : set up of simple Wireless Local Area Networks

WLAN STANDARDS

WLAN standard	characteristics
802.11a	Faster data transfer rates (up to 54Mbps) Supports more simultaneous connections Less susceptible to interference
802.11b	Better at penetrating physical barriers Longest range (70-150 feet) Hardware is usually less expensive
802.11g	Faster data transfer rates (up to 54Mbps) Better range than 802.11b (65-120 feet)
802.11n	The 802.11n standard was recently ratified by the Institute of Electrical and Electronics Engineers (IEEE), as compared to the previous three standards. Though specifications may change, it is expected to allow data transfer rates up to 600Mbps, and may offer larger ranges.

Personal Area Networks

- Personal Area networks are mostly intended for simple “cable replacement” duties.
- Devices following the Bluetooth standard allows to connect a hands-free headset to a phone without requiring a cable.
- The distance between the two devices is less than a meter.
- Data rates are fairly low (<1Mbit/s).
- Recently, wireless communications between components in an entertainment system (DVD player to TV), between computer and peripheral devices (printer, mouse), and similar applications have gained importance, and a number of standards for PANs have been developed by the IEEE 802.15 group.
- For the above applications, data rates in excess of 100 Mbit/s are used.
- Networks for even smaller distances are called Body Area Networks (BANs), which enable communications between devices located on various parts of a user’s body.
- BANs play an increasingly important role in the monitoring of patients’ health and of medical devices (e.g., pacemakers).

Fixed Wireless Access

Fixed wireless access systems can also be considered as a derivative of cordless phones or WLANs,essentially replacing a dedicated cable connection between the user and the public landline system.

The main difference from a cordless system is that

(i) There is no mobility of the user devices and
(ii) The BS almost always serves multiple users.
(iii) The distances bridged by fixed wireless access devices are much larger (between 100m and several tens of kilometers) than those bridged by cordless telephones.
(iv) Fixed wireless access has its main market for covering rural areas, and for establishing connections in developing countries that do not have any wired infrastructure in place.
(v) The IEEE 802.16 (WiMAX) standard tries to ease that problem by allowing some limited mobility in the system, and thus blurs the distinction from cellular telephony.

Ad hoc Networks and Sensor Networks

A **wireless ad hoc sensor network** consists of a number of sensors spread across a geographical area. Each sensor has wireless communication capability and some level of intelligence for signal processing and networking of the data. Some examples of wireless ad hoc sensor networks are the following:

1. Military sensor networks to detect and gain as much information as possible about enemy movements, explosions, and other phenomena of interest.

2. Sensor networks to detect and characterize Chemical, Biological, Radiological, Nuclear, and Explosive (CBRNE) attacks and material.

3. Sensor networks to detect and monitor environmental changes in plains, forests, oceans, etc.

4. Wireless traffic sensor networks to monitor vehicle traffic on highways or in congested parts of a city.

5. Wireless surveillance sensor networks for providing security in shopping malls, parking garages, and other facilities.

6. Wireless parking lot sensor networks to determine which spots are occupied and which are free.

The above list suggests that wireless ad hoc sensor networks offer certain capabilities and enhancements in operational efficiency in civilian applications as well as assist in the national effort to increase alertness to potential terrorist threats.

Two ways to classify wireless ad hoc sensor networks are whether or not the nodes are individually addressable, and whether the data in the network is aggregated.

The sensor nodes in a parking lot network should be individually addressable, so that one can determine the locations of all the free spaces. This application shows that it may be necessary to broadcast a message to all the nodes in the network.

If one wants to determine the temperature in a corner of a room, then addressability may not be so important. Any node in the given region can respond. The ability of the sensor network to aggregate the data collected can greatly reduce the number of messages that need to be transmitted across the network. This function of data fusion is discussed more below.

The basic goals of a wireless ad hoc sensor network generally depend upon the application, but the following tasks are common to many networks:

1. Determine the value of some parameter at a given location: In an environmental network, one might one to know the temperature, atmospheric pressure, amount of sunlight, and the relative humidity at a number of locations. This example shows that a given sensor node may be connected to different types of sensors, each with a different sampling rate and range of allowed values.

2. Detect the occurrence of events of interest and estimate parameters of the detected event or events: In the traffic sensor network, one would like to detect a vehicle moving through an intersection and estimate the speed and direction of the vehicle.

3. Classify a detected object: Is a vehicle in a traffic sensor network a car, a mini-van, a light truck, a bus, etc.

4. Track an object: In a military sensor network, track an enemy tank as it moves through the geographic area covered by the network.

In these four tasks, an important requirement of the sensor network is that the required data be disseminated to the proper end users. In some cases, there are fairly strict time requirements on this communication. For example, the detection of an intruder in a surveillance network should be immediately communicated to the police so that action can be taken.
Wireless ad hoc sensor network requirements include the following:

1. Large number of (mostly stationary) sensors: Aside from the deployment of sensors on the ocean surface or the use of mobile, unmanned, robotic sensors in military operations, most nodes in a smart sensor network are stationary. Networks of 10,000 or even 100,000 nodes are envisioned, so scalability is a major issue.

2. Low energy use: Since in many applications the sensor nodes will be placed in a remote area, service of a node may not be possible. In this case, the lifetime of a node may be determined by the battery life, thereby requiring the minimization of energy expenditure.

3. Network self-organization: Given the large number of nodes and their potential placement in hostile locations, it is essential that the network be able to self-organize; manual configuration is not feasible. Moreover, nodes may fail (either from lack of energy or from physical destruction), and new nodes may join thenetwork.Therefore, the network must be able to periodically reconfigure itself so that it can continue to function. Individual nodes may become disconnected from the rest of the network, but a high degree of connectivity must be maintained.

4. Collaborative signal processing: Yet another factor that distinguishes these networks from MANETs is that the end goal is detection/estimation of some events of interest, and not just communications. To improve the detection/estimation performance, it is often quite useful to fuse data from multiple sensors. This data fusion requires the transmission of data and control messages, and so it may put constraints on the network architecture.

5. Querying ability: A user may want to query an individual node or a group of nodes for information collected in the region. Depending on the amount of data fusion performed, it may not be feasible to transmit a large amount of the data across the network. Instead, various local sink nodes will collect the data from a given area and create summary messages. A query may be directed to the sink node nearest to the desired location.

Satellite Cellular Communications

Satellite phones: The concept of using a mobile phone from anywhere on the globe is one that has many applications. Although the terrestrial cellular network is widely available, there are still very many areas where coverage is not available. In these situations satellite phones are of great use.

As an example satellite phones are widely used by the emergency services for situations when they are in remote areas, even of countries that might have a good cellular network, but not in remote areas. They may also be for communications in rural areas where no cellular coverage may be available. They also find uses at sea, and in developing countries, or in uninhabited areas of the globe.

The below Figure 1.9 represents Satellite Cellular Communications

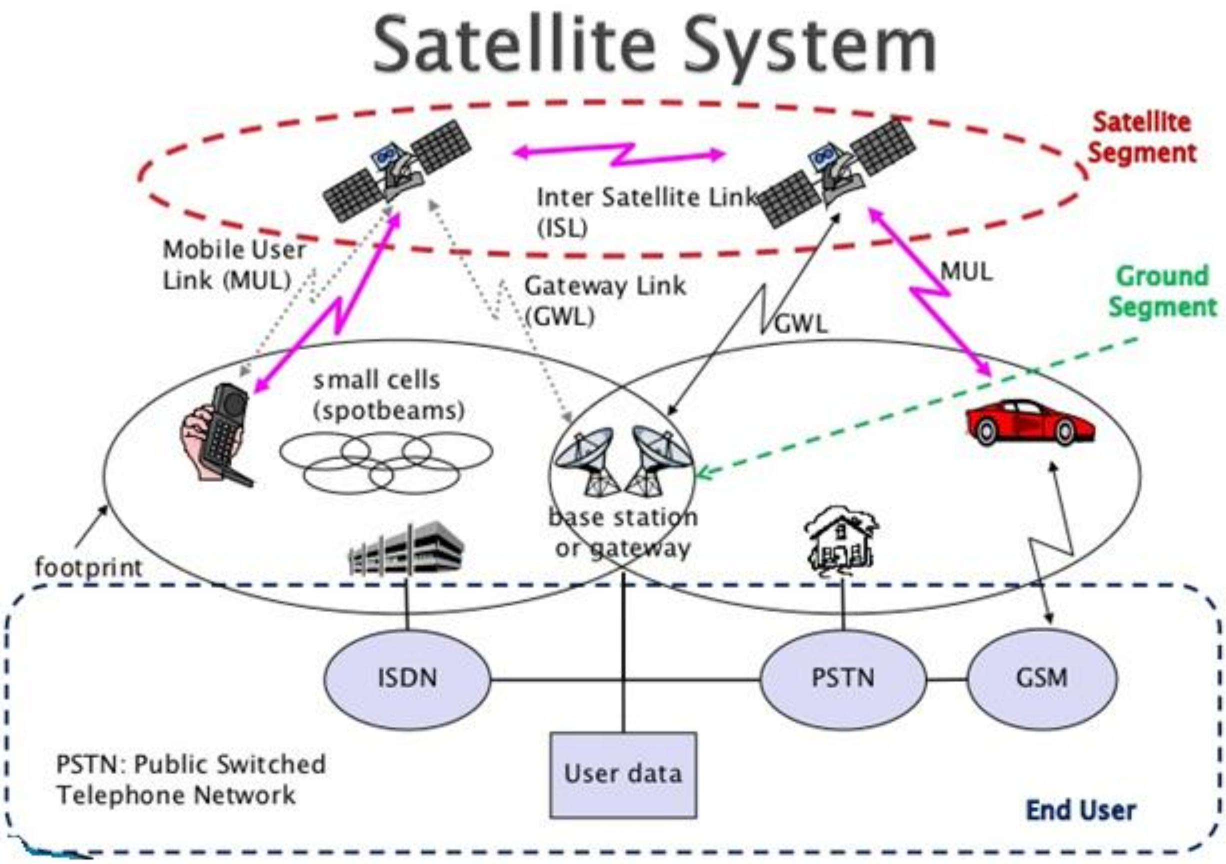


Figure 1.9 : Satellite Cellular Communications

The satellite mobile is connected to the satellite through mobile user link (MUL)

Satellite in turn is connected to the base station or gateway through gate way link (GWL) and then to the end user through GSM.

A call established between user A to user B is represented in below figure 1.10

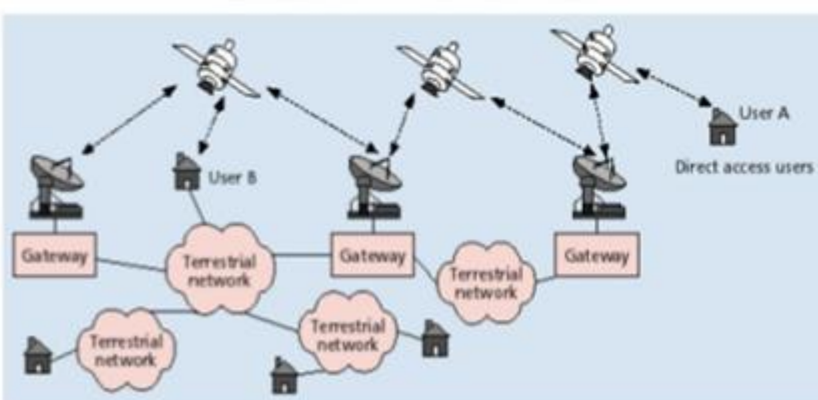


Figure 1.10: Call establishment between user A and user B

Benefits:

- Less complex- switching circuits are on the ground and satellites are just reflectors.
- Most of the call is transferred over public telephone network.

Issues:

- Gateway must be in line of sight of the satellite
- Significant number of ground geteways to provide direct satellite links.

Requirements for the Services

The key Requirements for the wireless communication offered Services are

- Data Rate
- Range and Number of Users
- Mobility
- Energy Consumption
- Use of Spectrum
- Direction of Transmission
- Service Quality

Data Rate

- **Speech communications:** It usually require between 5 and 64 kbit/s depending on the required quality and the amount of compression.
- **Elementary data services:** Require between 10 and 100 kbit/s.
- **Cellular systems:** Require data rates of about 10 kbit/s are standard.
- **Communications between computer peripherals and similar devices:** Wireless links with data rates around 1Mbit/s are used.
- **High-speed data services:** WLANs and 3G cellular systems are range from 0.5 to 100 Mbit/s.
- **Personal Area Networks(PANs):** The range of data rates (over100 Mbit/s).
- **Sensor networks:** Require data rates from a few bits per second to about 1 kbit/s.

Range and number of users:

- Peculiarity among the different networks is the range and the number of users that they serve.
- By “range,” we mean here the distance between one transmitter and receiver.
- The coverage area of a wireless communication system can be made almost independent of the range, by just combining a larger number of BSs into one big network.

SERVICES	RANGE	USERS
BANs	~1m	1
PANs	~10m	1
WLANs	~100m	~10
Cellular systems	Microcells 500m, Macrocells 10-30km	More
Fixed wireless access	100m-several km	More
Satellite systems	Large (even the size of a country)	Limited

Mobility:

The ability to move around while communicating through wireless equipment (mobile phone) is mobility

Mobility exists in different grades .

- *Fixed devices* are placed only once, and after that time they communicate with their Base Set (BS), or with each other, always from the same location. The main motivation is to avoiding the laying of cables. Example cordless phones.
- *Nomadic devices* are placed at a certain location for a limited duration of time and then moved to a different location (e.g.) Laptop, wireless Gadgets.
- *Low mobility devices* are operated at pedestrian speeds.(e.g.) cellphones operated by walking human.
- *High mobility devices*- Cellphones operated by people in moving vehicles (about 30 to 150km/h).
- *Extremely high mobility* - high-speed trains and planes (300 to 1000 km/h). These speeds pose unique challenges both for the design of the physical layer (Doppler shift) and for the handover between cells.

Energy Consumption

- Energy consumption is a critical aspect for wireless devices.
- Most wireless devices use rechargeable batteries, as they should be free of *any* wires

Rechargeable batteries**:** nomadic and mobile devices, like laptops, cellphones, and cordless phones, are usually operated with rechargeable batteries. Standby times as well as operating times are one of the determining factors for customer satisfaction.

Energy consumption is determined on one hand by the distance over which the data have to be transmitted (remember that a minimum SNR has to be maintained).

One-way batteries**:** sensor network nodes often use one-way batteries, which offer higher energydensity at lower prices. Furthermore, changing the battery is often not an option; rather, the sensor including the battery and the wireless transceiver is often discarded after the battery has run out.

Power mains: BSs and other fixed devices can be connected to the power mains. Therefore, energy efficiency is not a major concern for them.

Use of Spectrum

- The term radio spectrum typically refers to the full frequency range from 3 kHz to 300 GHz that may be used for wireless communication
- Sections of spectrum are called "bands".
- The portions of spectrum set aside for wireless mobile phone service are split into two bands.
- The first is Cellular, which is centered roughly around 800 MHz.
- The second is Personal Communication System (PCS), which is centered roughly around 1900 MHz.
- Examples of frequency spectrum and its uses are summarized in Table 1.1

Table 1.1: Frequency spectrum and its uses

Band name	Abbreviation	ITU band	Frequency and wavelength / spectrum	Example uses
N/A			< 3 Hz > 100,000 km	Natural and artificial electromagnetic noise
Extremely low frequency	ELF	1	3–30 Hz,100,000 km – 10,000 km	Communication with submarines
Super low frequency	SLF	2	30–300 Hz,10,000 km – 1000 km	Communication with submarines
Ultra low frequency	ULF	3	300–3000 Hz,1000 km – 100 km	Submarine communication, communication within mines
Very low frequency	VLF	4	3–30 kHz,100 km – 10 km	Navigation, time signals, submarine communication, wirelessheart rate monitors, geophysics
Low frequency	LF	5	30–300 kHz,10 km – 1 km	Navigation, clock time signals, AM longwave broadcasting (Europe and parts of Asia), RFID, amateur radio
Medium frequency	MF	6	300–3000 kHz,1 km – 100 m	AM (medium-wave) broadcasts, amateur radio, avalanche beacons
High frequency	HF	7	3–30 MHz,100 m – 10 m	Shortwave broadcasts, citizens' band radio, amateur radio andover-the-horizon aviation communications, RFID, over-the-horizon radar, automatic link establishment (ALE) / near-vertical incidence skywave (NVIS) radio communications, marine and mobile radio telephony
Very high frequency	VHF	8	30–300 MHz,10 m – 1 m	FM, television broadcasts and line-of-sight ground-to-aircraft and aircraft-to-aircraft communications, land mobile and maritime mobile communications, amateur radio, weather radio
Ultra high frequency	UHF	9	300–3000 MHz,1 m – 100 mm	Television broadcasts, microwave oven, microwavedevices/communications, radio astronomy, mobile phones,wireless LAN, Bluetooth, ZigBee, GPS and two-way radios such as land mobile, FRS and GMRS radios, amateur radio
Super high frequency	SHF	10	3–30 GHz,100 mm – 10 mm	Radio astronomy, microwave devices/communications, wireless LAN, most modern radars, communications satellites, satellite television broadcasting, DBS, amateur radio
Extremely high frequency	EHF	11	30–300 GHz,10 mm – 1 mm	Radio astronomy, high-frequency microwave radio relay, microwave remote sensing, amateur radio, directed-energy weapon, millimeter wave scanner
Terahertz or Tremendously high frequency	THz or THF	12	300–3,000 GHz,1 mm – 100 μm	Terahertz imaging – a potential replacement for X-rays in some medical applications, ultrafast molecular dynamics, condensed-matter physics, terahertz time-domain spectroscopy, terahertz computing/communications, sub-mm remote sensing, amateur

Direction of Transmission:

There are 3 different transmission modes characterized according to the direction of the Transmission.

They are i) simplex, ii) Half duplex iii) Full duplex

A simplex connection is a connection in which the data flows in only one direction, from the transmitter to the receiver. This type of connection is useful if the data do not need to flow in both directions (for example, from your computer to the printer or from the mouse to your computer...).The below Figure 1.11 represents simplex connection.

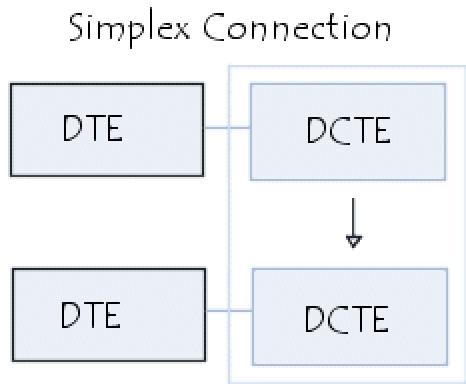


Figure 1.11 : simplex connection.

A half-duplex connection (sometimes called an *alternating connection* or *semi-duplex*) is a connection in which the data flows in one direction or the other, but not both at the same time. The below Figure 1.12 represents half duplex connection.

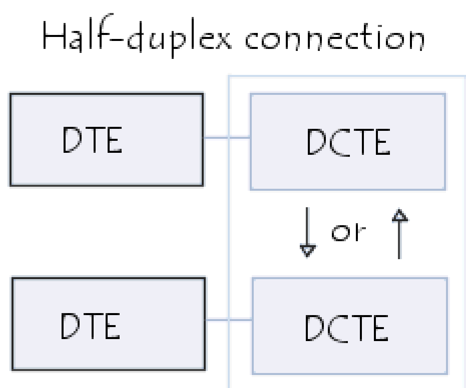


Figure 1.12 :Half-duplex connection.

A full-duplex connection is a connection in which the data flow in both directions simultaneously. Each end of the line can thus transmit and receive at the same time, which means that the bandwidth is divided in two for each direction of data transmission if the same transmission medium is used for both directions of transmission.The below Figure 1.13 represents full duplex connection.

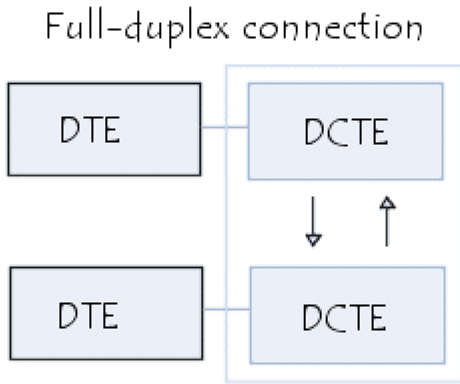


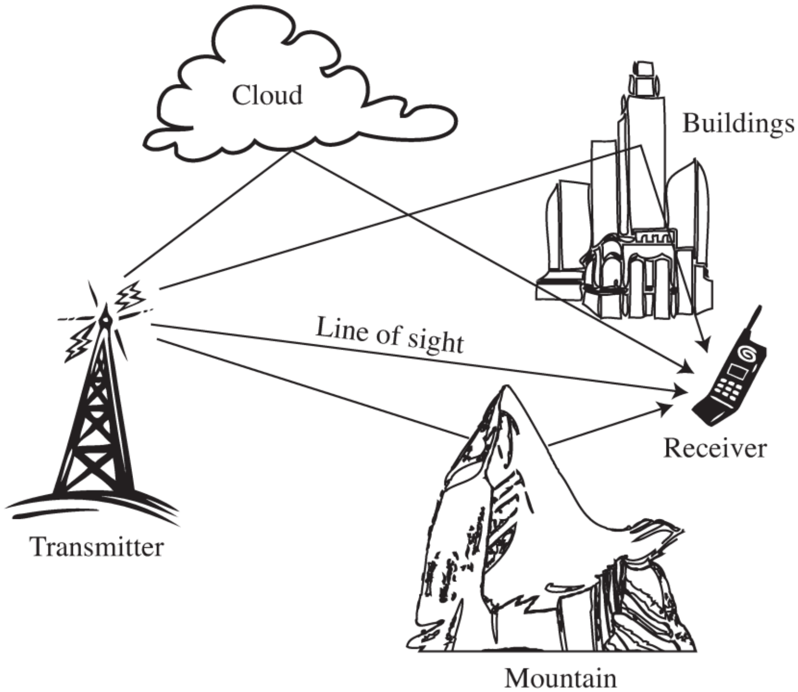
Figure 1.13: Full-duplex connection.

Service Quality

- The first main indicator for service quality is *speech quality* for speech services. Speech quality is usually measured by the *Mean Opinion Score* (MOS).
- The service quality for data services are defined by *file transfer speed*.
- The speed of data transmission is simply measured in bit/s
- For cellphones and other speech services, the *service quality* is often computed as the complement of "fraction of blocked calls Plus 10 times the fraction of dropped calls.

MULTIPATH PROPAGATION

- In wireless telecommunications, multipath is the propagation phenomenon that results in radio signals reaching the receiving antenna by two or more paths.
- Causes of multipath include atmospheric ducting, ionospheric reflection and refraction, and reflection from water bodies and terrestrial objects such as mountains and buildings.
- The effects of multipath include constructive and destructive interference, and phase shifting of the signal.
- Destructive interference causes fading. Where the magnitudes of the signals arriving by the various paths have a distribution known as the Rayleigh distribution, this is known as Rayleigh fading.



Fading:

- In wireless communications, fading is deviation of the attenuation affecting a signal over certain propagation media. The fading may vary with time, geographical position or radio frequency, and is often modeled as a random process.
- A fading channel is a communication channel that experiences fading. In wireless systems, fading may either be due to multipath propagation, referred to as multipath induced fading, or due to shadowing from obstacles affecting the wave propagation, sometimes referred to as shadow fading.
- The figure 1.15 given below represents the fading impact on signal amplitudes.

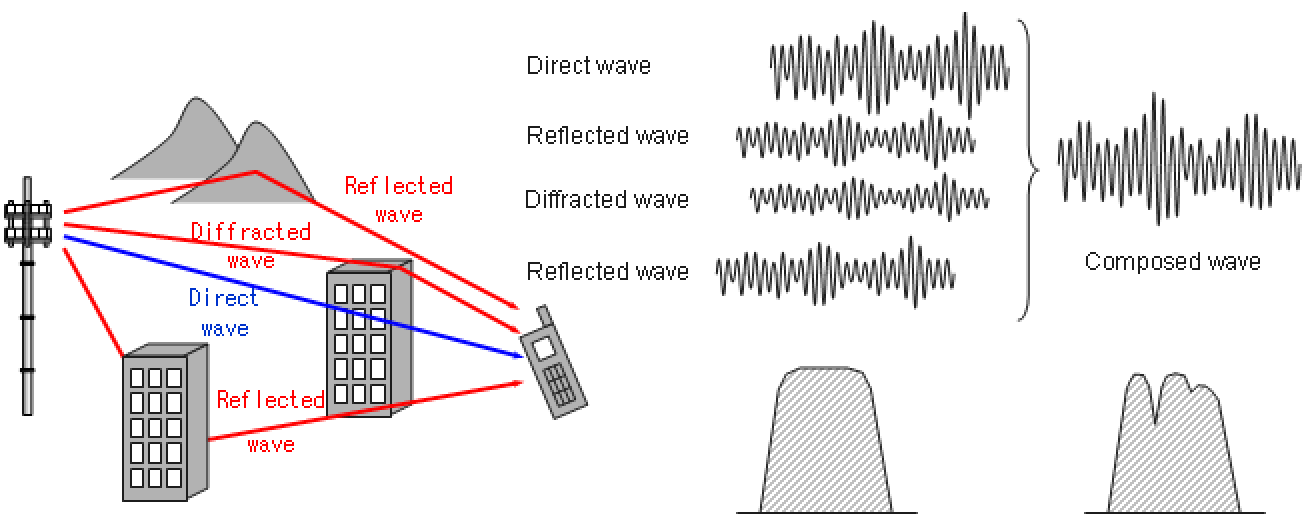
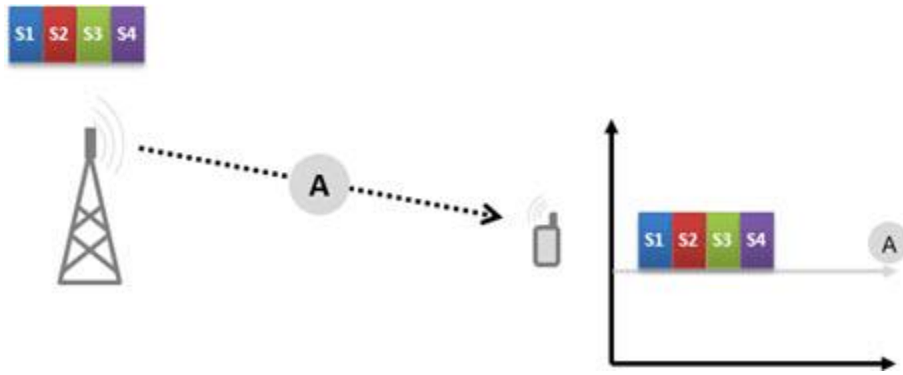


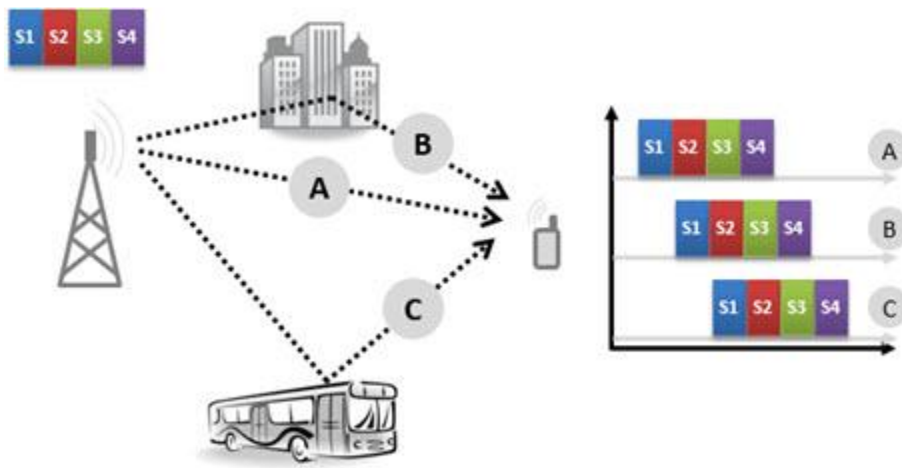
Figure 1.15: Fading impact on signal amplitudes.

Inter symbol interference:

In an ideal system (theoretical), the transmitted symbols arrive at the receiver without any loss or interference, as shown in the following figure.



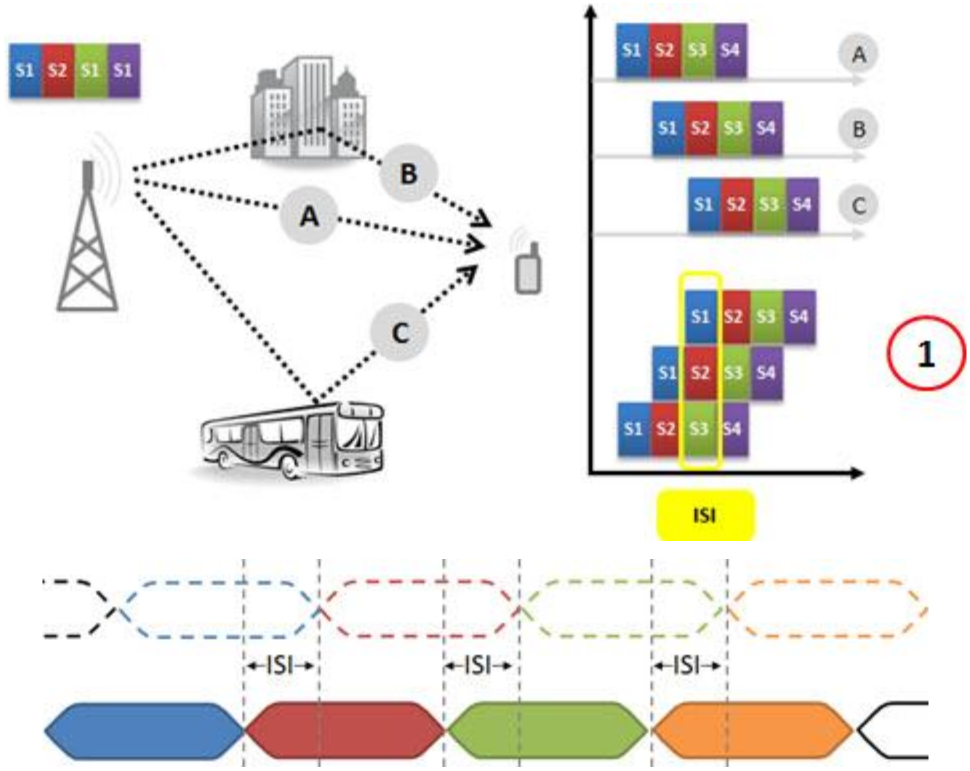
But in a real scenario the transmitted signals are affected in different ways, for example, according to the propagation environment.



Although the "Multipath" bring positive benefits, "Multipath" and "Delay Spread" also end up causing interference inter symbols.

At the receiver, all these "multipath" components are summed as A,B and C have multiple symbols being received "simultaneously" (Symbols "Overlap") - this is the intersymbol interference (ISI)

The below figure represents the ISI in the real-time scenario.



Spectrum Limitations

The spectrum available for wireless communications services is limited, and regulated by international agreements.

The spectrum has to be used in a highly efficient manner.

Two approaches are used for spectrumallocation.

i) Regulated spectrum usage:-Where a single network operator has control over the usage of the spectrum.

ii) Unregulated spectrum usage:-Each user can transmit without additional control, as long as he complies with certain restrictions on the emission power and bandwidth.

Assigned Frequencies

- The frequency assignment for different wireless services is regulated by the International TelecommunicationsUnion (ITU)

The other regulatory bodies are

- Federal Communications Commission (FCC)
- Association of Radio Industries and Businesses (ARIB)
- European Conference of Postal and Telecommunications Administrations (CEPT)

Few examples of frequency assignment and their uses are listed below.

- Below 100 MHz: At these frequencies, we find Citizens' Band (CB) radio, pagers, and analog cordless phones.
- 100–800 MHz: these frequencies are mainly used for broadcast (radio and TV) applications.

- 400–500 MHz: a number of cellular and trunking radio systems make use of this band. It is mostly systems that need good coverage, but show low user density.

- 800–1000 MHz: several cellular systems use this band (analog systems as well as 2G cellular). Also some emergency communications systems (trunking radio) make use of this band.

- 1.8–2.1 GHz: this is the main frequency band for cellular communications. The current (2G) cellular systems operate in this band, as do most of the third-generation systems.

- 2.4–2.5 GHz: the Industrial, Scientific, and Medical (ISM) band. Cordless phones, Wireless Local

- Area Networks (WLANs) and wireless Personal Area Networks (PANs) operate in this band; they share it with many other devices, including microwave ovens.

- 3.3–3.8 GHz: is envisioned for fixed wireless access systems.
- 4.8–5.8 GHz: in this range, most WLANs can be found. Also, the frequency range between

- 5.7 and 5.8 GHz can be used for fixed wireless access, complementing the 3-GHz band. Also car-to-car communications are working in this band.

- 11–15 GHz: in this range we can find the most popular satellite TV services, which use 14.0–14.5 GHz for the uplink, and 11.7–12.2 GHz for the downlink.

Noise and Interference Limited Systems.

Noise-Limited Systems

Noise is unwanted electrical or electromagnetic energy that degrades the quality of signals and data.

Noise occurs in digital and analog systems, and can affect communications of all types, including text, programs, images, audio, and telemetry.

In general, noise originating from outside the system is inversely proportional to the frequency, and directly proportional to the wavelength.

Fixing upof noise threshold for the wireless system is called as noise limited. we can call it *signal power limited.*

Depending on the interpretation, it is too much noise or too little signalpower that leads to bad link quality.

$$P_{RX} = P_{TX} G_{RX} G_{TX} (\lambda / 4\pi d)^2 \quad \text{---} \qquad (1)$$

where G_{RX} and G_{TX} are the gains of the receive and transmit antennas, respectively. λ is the wavelength, P_{TX} is the transmit power, P_{RX} is the Receiver power and d is the distance between the transmitter and receiver.

The noise that disturbs the signal can consist of several components. They are

Thermal noise: The power spectral density of thermal noise depends on the environmentaltemperature T_e that the antenna "visualizes."

The temperature of the Earth is around 300 K, while the temperature of the (cold) sky is approximately $T_e \approx 4$K (the temperature in the direction of the Sun is of course much higher). As a first approximation, it is usually assumed that the environmental temperature is isotropic ally 300 K. Noise power spectral density is then

$$N_0 = k_B T_e \text{--} \quad (2)$$

wherek_B is Boltzmann's constant and is $k_B = 1.38 * 10^{-23}$ J/K,

$$\text{Noise power is } P_n = N_0 B \text{---} \quad (3)$$

Where B is RX bandwidth . It is common to write Eq. (2) using logarithmic units

Power P expressed in units of dBm is 10 log10 (P/1 mW)

$N_0 = -174$ dBm/Hz --- (4)

This means that the noise power contained in a 1-Hz bandwidth is −174 dBm.

The noise power contained in bandwidth B is −174 + 10 log10(B)dB.------- (5)

The logarithm of bandwidth B, specifically 10 log10(B), has the units dBHz.

The other types of noises are Man-made noise and Other intentional emission sources

The **noise factor** F of a system is defined as F= SNR_{in}/ $SNR_{out,}$

whereSNR_{in} is input signal to noise ratio and SNR_{out} is output signal to noise ratio.

Receiver noise:

Mathematically, the total noise figure F_{eq} of a cascade of components is

$$F_{eq} = F_1 + (F_2 - 1/G_1) + (F_3 - 1/G_1 G_2) + \cdot \cdot \cdot$$

where F_i and G_i are noise figures and noise gains of the individual stages in absolute units Note that for this equation, passive components, like attenuators with gain $m < 1$, can be interpreted as *either* having a noise figure of $F = 1/m$ and unit gain of $G = 1$, *or* unit noise figure $F = 1$, and gain $G = m$.

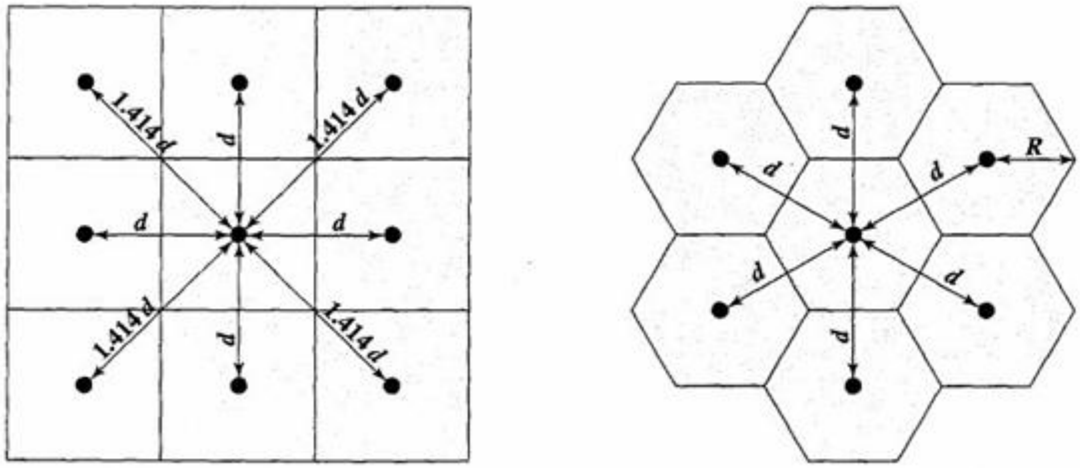
Cellular Wireless Networks

Principles of Cellular Networks & its Organization

- The essence of a cellular network is the use of multiple low-power transmitters, on the order of 100W or less.
- Due to small range transmitters an area can be divided into cells, each one served by its own antenna.
- Each cell is allocated a band of frequencies and is served by a **base station,** consisting of transmitter, receiver, and control unit.
- Adjacent cells are assigned different frequencies to avoid interference or crosstalk.

- If the width of a square cell is d, then a cell has .four neighbors at a distance d and four neighbors at a distance $\sqrt{2}d$.
- For a cell radius R, the distance between the cell center and each adjacent cell center is

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



(a) Square pattern (b) Hexagonal pattern

Frequency Reusability

- To use the same frequency band in multiple cells at some distance from one another.
- This allows the same frequency band to be used for multiple simultaneous conversations in different cells.
- In characterizing frequency reuse, the following parameters are commonly used:

D = minimum distance between centers of cells that use the same frequency band (called co-channels)
R = radius of a cell
d = distance between centers of adjacent cells (d =$\sqrt{3}R$)
N = number of cells in a repetitious pattern (each cell in the pattern uses a unique set of frequency bands), termed the **reuse factor**

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

Increasing Capacity

In time, as more customers use the system, traffic may build up so that there are not enough frequency bands assigned to a cell to handle its calls.

Following approaches have been used to cope with this situation:

- **Frequency borrowing**
  - Frequencies are taken from adjacent cells by congested cells.

- **Cell splitting**
  - Cells in areas of high usage can be split into smaller cells.

- **Cell sectoring**

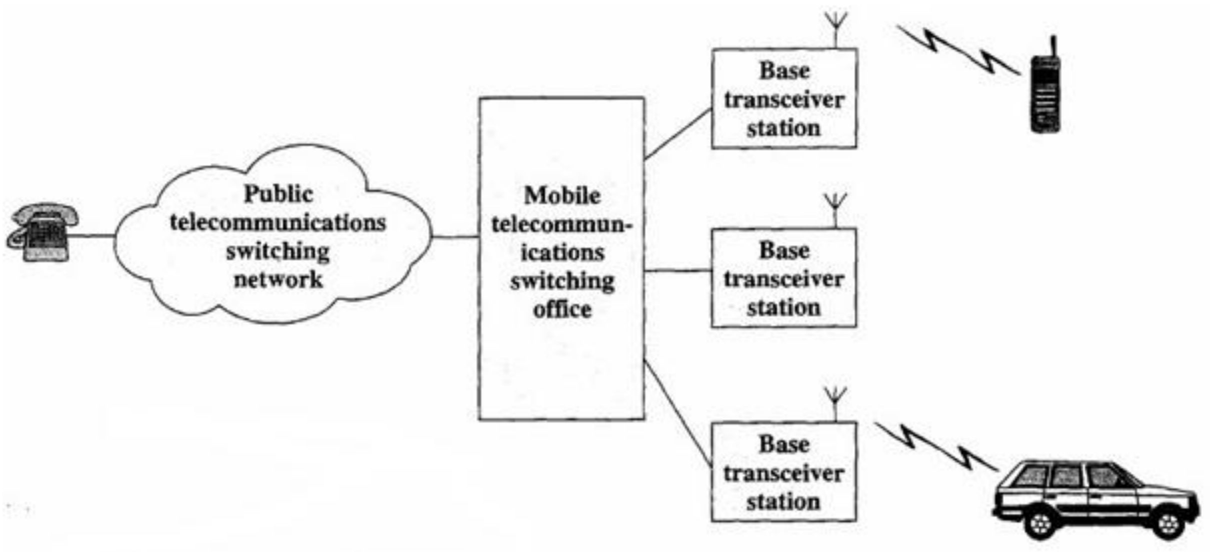
- A cell is divided into a number of wedge-shaped sectors, each with its own set of channels, typically 3 or 6 sectors per cell.

- **Microcells**
    - As cells become smaller, antennas move from the tops of tall buildings or hills, and finally to lamp posts, where they form microcells.
    - Microcells are useful in city streets in congested areas, along highways, and inside large public buildings.

Typical parameters for Macro cells and Microcells

	Macrocell	**Microcell**
Cell radius	**1 to 20 Km**	**.1 to 1 Km**
Transmission power	**1 to 10W**	**.1 to 1W**
Average delay spread	**.1 to 10 µs**	**10 to 100 ns**
Maximum bit rate	**,3 Mbps**	**1 Mbps**

<u>**Cellular Systems**</u>



- Center of each cell is a base station (BS).
- The BS includes an antenna, a controller, and a number of transceivers, for communicating on the channels assigned to that cell.
- The controller is used to handle the call process between the mobile unit and the rest of the network.
- Each BS is connected to a mobile telecommunications switching office (**MTSO**).
- The link between an MTSO and a BS is by a wire line.
- Two types of channels are available between the mobile unit and the base station (BS):
    1. **Control channels**
        - Used to exchange information having to do with setting up and maintaining calls and with establishing a relationship between a mobile unit and the nearest BS.
    2. **Traffic channels**
        - Carry a voice or data connection between users.

Mobile Telecommunications Switching Office (MTSO)

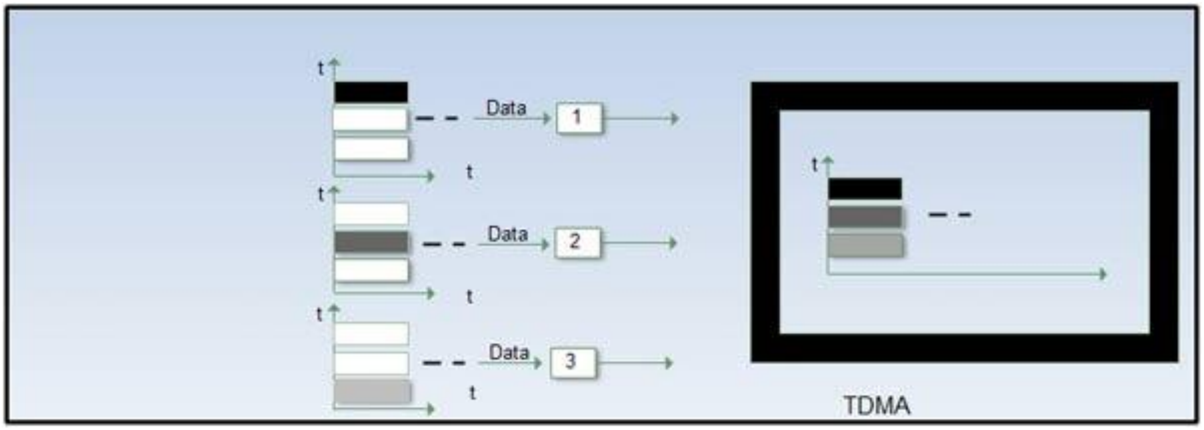
- The MTSO connects calls between mobile units.
- The MTSO is also connected to the telecommunications network and can make a connection between a fixed subscriber to the public network and a mobile subscriber to the cellular network.
- The following are the steps in a typical call between two mobile users within an area controlled by a single MTSO:
    - Mobile unit initialization
    - Mobile-originated call
    - Paging
    - Call accepted
    - Ongoing call
    - Handoff

- Other functions
    - Call blocking
    - Can termination
    - Call drop
    - Calls to/from fixed and remote mobile subscriber

Multiple Access Schemes

Time Division Multiple Access (TDMA).

- In TDMA, the bandwidth of channel is dividend amongst various stations on the basis of time.
- Each station is allocated a time slot during which it can sent its data *i.e.* each station can transmit its data in its allocated time slot only.
- Each station must know the beginning of its slot and the location of its slot.
- TDMA requires synchronization between different stations.
- Synchronization is achieved by using some synchronization bits (preamble bits) at the beginning of each slot.
- TDMA is different from TDM, although they are conceptually same.
- TDM is a physical layer technique that combines the data from slower channels and transmits then by using a faster channel. This process uses physical multiplexer.
- TDMA, on other hand, is an access method in the data link layer. The data link layer in each station tells its physical layer to use the allocated time slot. There is no physical multiplexer at the physical layer.



General specification of TDMA

- Rx: 869-894MHz.
- Tx: 824-849MHz.
- 832 Channels spaced 30kHz apart (3 users/channel).
- DQPSK modulation scheme.
- 48.6kbps bit rate.
- Uses Time Division Duplexing (TDD) usually.
- Interim Standard (IS) – 54

Advantages of TDMA

- Flexible bit rate
- Extended Battery life
- TDMA installations offer savings in base station equipment, space and maintenance
- No frequency guard band required.
- Easy for mobile or base stations to initiate and execute handoffs.

Disadvantages of TDMA

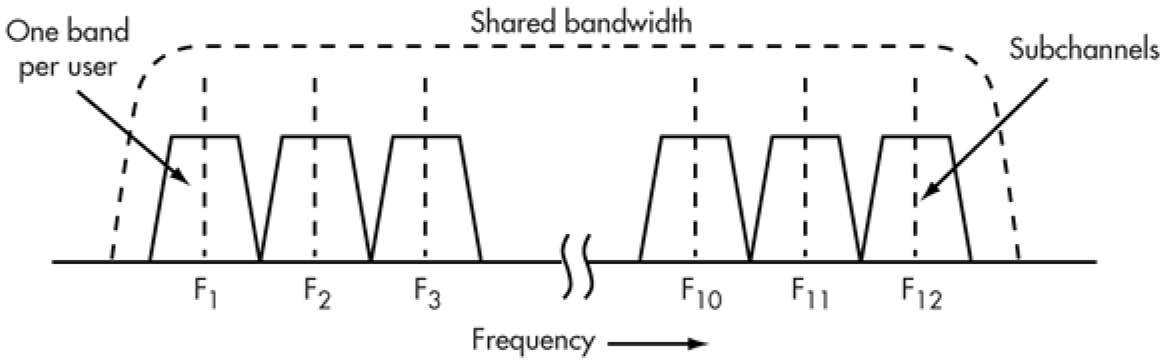
- Multipath distortion
- Requires signal processing for matched filtering and correlation detection.
- Demands high peak power on uplink in transient mode.
- Requires network wide timing synchronization.

<u>Frequency Division Multiple Access (FDMA).</u>

FDMA is the process of dividing one channel or bandwidth into multiple individual bands, each for use by a single user .

Each individual band or channel is wide enough to accommodate the signal spectra of the transmissions to be propagated.

The data to be transmitted is modulated on to each subcarrier, and all of them are linearly mixed together. The frequency spectrum of FDMA is shown in below Figure.



- FDMA divides the shared medium bandwidth into individual channels. Subcarriers modulated by the information to be transmitted occupy each sub-channel.

- The best example of this is the cable television system. The medium is a single coax cable that is used to broadcast hundreds of channels of video/audio programming to homes.

- The coaxial cable has a useful bandwidth from about 4 MHz to 1 GHz. This bandwidth is divided up into 6-MHz wide channels.

- Initially, one TV station or channel used a single 6-MHz band. But with digital techniques, multiple TV channels may share a single band today thanks to compression and multiplexing techniques used in each channel.

1General specification of FDMA

- Rx: 869-894MHz.
- Tx: 824-849MHz.
- 832 Channels spaced 30kHz apart (3 users/channel).
- DQPSK modulation scheme.
- 48.6kbps bit rate.
- Uses Frequency Division Duplexing (FDD) ISI (Inter symbol Interference) is low.

Advantages of FDMA:

- Fairly efficient when the number of stations is small and the traffic is uniformly constant.
- No need for network timing
- Capacity increase can be obtained by reducing the information bit rate and using efficient digital code

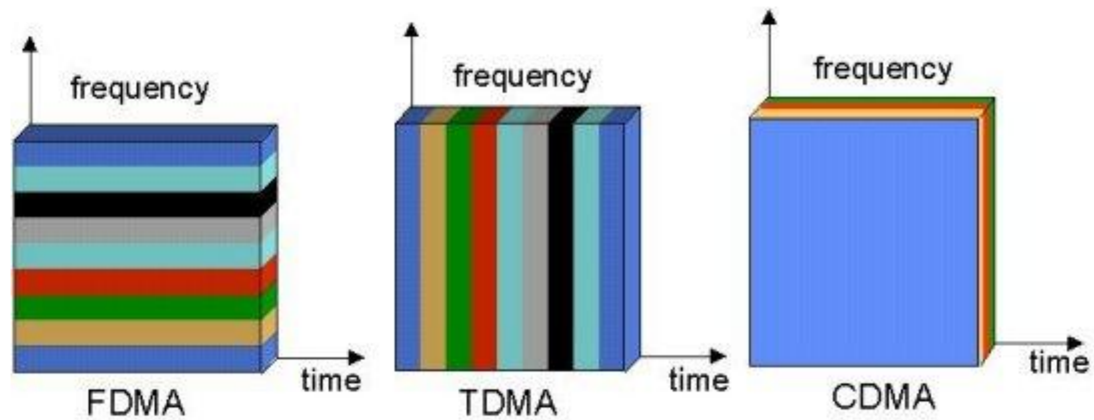
No restriction regarding the type of baseband or type of modulation.

Disadvantages of FDMA:

- Presence of guard bands.
- Does not differ significantly from analog system.
- Maximum bit rate per channel is fixed.
- Small inhibiting flexibility in bit rate CAPABILITY. Allotted bandwidth of FDMA is 12.5 MHz.

Code division multiple Access (CDMA)

- In CDMA, the User's signal is spread by unique sequence. All users share the same bandwidth with different sequences.
- Communication between the transmitter and receiver is identified by the sequence. The transmitter and receiver know the sequence in advance.
- CDMA is based on spread spectrum technology.
- Spread spectrum is a transmission technology wherein data occupy a larger bandwidth than necessary. Bandwidth spreading is accomplished before transmission using a code that is independent of the transmitted data
- The same code is used to demodulate the data at receiving end. CDMA is originally designed for military to avoid jamming.
- Random signature sequences can be implemented in CDMA. With the help of direct sequence spread spectrum, it is possible for the user in the CDMA system to occupy the entire spectrum .
- With the help of this sequence, we can able to receive multiple packets simultaneously without losing any packets due to collision.
- The capability were the multiple packets can be received by the physical layer is known as multi packet reception (MPR).
- Spreading in CDMA is achieved with the help of direct spreading sequence (DSSS) technique. In this DSSS, each code symbol is spread by a unit energy chip waveform. The spreading sequence is independent from code symbol to code symbol and across users.
- The below figure depicts the CDMA spectrum with respect toTDMA and FDMA.



General specification of CDMA

- Rx: 869-894MHz.
- Tx: 824-849MHz.
- 20 Channels spaced 1250kHz apart (798 users/channel).
- QPSK/(Offset) OQPSK modulation scheme
- 1.2288Mbps bit rate.
- Operates at both 800 and 1900 MHz frequency bands.
- IS-95 standard

Advantages of CDMA

- No absolute limit on the number of users
- Easy addition of more users
- Better signal quality
- Impossible for the hackers to decipher the code sent.
- Multipath fading may be substantially reduced because of large signal bandwidth.

Disadvantages of CDMA

- Near-far problem arises
- Self-jamming.
- As the number of users increases, the overall quality of service decreases.

UNIT II

WIRELESS PROPAGATION CHANNELS (Qualitative Treatment only): Free space propagation model- basic propagation mechanisms –reflection- ground reflection model-diffraction-scattering- -outdoor and indoor propagation models. Small scale fading and multipath: Small scale multipath propagation-Impulse response model of multipath channel – small scale multipath measurements –parameters of mobile multipath channels -types of small scale fading.

INTRODUCTION

A wireless propagation channel is the medium linking the transmitter and the receiver. Its properties determine the information-theoretic capacity, i.e., the ultimate performance limit of wireless communications, and also determine how specific wireless systems behave. Wireless channels differ from wired channels by *multipath propagation,* i.e., the existence of a multitude of propagation paths from transmitter and receiver, where the signal can be reflected, diffracted, or scattered along its way.

Free Space model – the basic propagation mechanisms:

We start with the simplest possible scenario: a transmit and a receive antenna in free space, and derive the received power as a function of distance. Energy conservation dictates that the integral of the power density over any closed surface surrounding the transmit antenna must be equal to the transmitted power. If the closed surface is a sphere of radius d, centered at the transmitter (TX) antenna, and if the TX antenna radiates isotropically, then the power density on the surface is $P_{TX}/(4\pi d^2)$. The receiver (RX) antenna has an "effective area" A_{RX}. We can envision that all power impinging on that area is collected by the RX antenna [Stutzman and Thiele 1997]. Then the received power is given by:

$$P_{RX}(d) = P_{TX}\frac{1}{4\pi d^2}A_{RX}$$

If the transmit antenna is not isotropic, then the energy density has to be multiplied with the antenna gain G_{TX} in the direction of the receive antenna.

$$P_{RX}(d) = P_{TX}G_{TX}\frac{1}{4\pi d^2}A_{RX} \tag{4.1}$$

The product of transmit power and gain in the considered direction is also known as *Equivalent Isotropically Radiated Power* (EIRP).

The effective antenna area is proportional to the power that can be extracted from the antenna connectors for a given energy density. For a parabolic antenna, e.g., the effective antenna area is roughly the geometrical area of the surface. However, antennas with very small geometrical area – e.g., dipole antennas – can also have a considerable effective area.

It can be shown that there is a simple relationship between effective area and antenna gain

$$G_{RX} = \frac{4\pi}{\lambda^2} A_{RX} \quad (4.2)$$

Most noteworthy in this equation is the fact that – for a fixed antenna area – the antenna gain increases with frequency. This is also intuitive, as the directivity of an antenna is determined by its size in terms of wavelengths.

Substituting Eq. (4.2) into (4.1) gives the received power PRX as a function of the distance d in free space, also known as *Friis' law*:

$$P_{RX}(d) = P_{TX} G_{TX} G_{RX} \left(\frac{\lambda}{4\pi d} \right)^2 \quad (4.3)$$

The factor $(\lambda/4\pi d)$2 is also known as the *free space loss factor*.

Friis' law seems to indicate that the "attenuation" in free space increases with frequency. This is counterintuitive, as the energy is not lost, but rather redistributed over a sphere surface of area $4\pi d$2. This mechanism has to be independent of the wavelength. This seeming contradiction is caused by the fact that we assume the *antenna gain* to be independent of the wavelength. If we assume, on the other hand, that the effective *antenna area* of the RX antenna is independent of

frequency, then the received power becomes independent of the frequency (see Eq. (4.1)). For wireless systems, it is often useful to assume constant gain, as different systems (e.g., operating at 900 and 1800 MHz) use the same antenna *type* (e.g., $\lambda/2$ dipole or monopole), and not the same antenna *size*.

The validity of Friis' law is restricted to the far field of the antenna – i.e., the TX and RX antennas have to be at least one *Rayleigh distance* apart. The Rayleigh distance (also known as the *Fraunhofer distance*) is defined as:

$$d_R = \frac{2L_a^2}{\lambda} \quad (4.4)$$

where$L$a is the largest dimension of the antenna; furthermore, the far field requires $d \gg \lambda$ and $d \gg L$a. For setting up link budgets, it is advantageous to write Friis' law on a logarithmic scale. Equation (4.3) then reads

$$P_{RX}|_{dBm} = P_{TX}|_{dBm} + G_{TX}|_{dB} + G_{RX}|_{dB} + 20\log\left(\frac{\lambda}{4\pi d}\right)$$

where$L$a is the largest dimension of the antenna; furthermore, the far field requires $d \gg \lambda$ and $d \gg L$a. For setting up link budgets, it is advantageous to write Friis' law on a logarithmic scale. Equation (4.3) then reads

$$P_{RX}|_{dBm} = P_{TX}|_{dBm} + G_{TX}|_{dB} + G_{RX}|_{dB} + 20\log\left(\frac{\lambda}{4\pi d}\right)$$

where |dB means “in units of dB.” In order to better point out the distance dependence, it is

advantageous to first compute the received power at 1-m distance:

$$P_{\mathrm{RX}}(1\mathrm{m}) = P_{\mathrm{TX}}|_{\mathrm{dBm}} + G_{\mathrm{TX}}|_{\mathrm{dB}} + G_{\mathrm{RX}}|_{\mathrm{dB}} + 20\log\left(\frac{\lambda|_{\mathrm{m}}}{4\pi \cdot 1}\right)$$

The last term on the r.h.s. of Eq. (4.8) is about −32 dB at 900MHz and −38 dB at 1800 MHz. The

actual received power at a distance *d* (in meters) is then:

$$P_{\mathrm{RX}}(d) = P_{\mathrm{RX}}(1\mathrm{m}) - 20\log(d|_{\mathrm{m}})$$

Basic Methods of Propagation

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

Reflection

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

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

$$sin(\theta_B) = \sqrt{\frac{\epsilon_1}{\epsilon_1 + \epsilon_2}}.$$

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

$$\theta_i = \theta_r$$

and for vertical polarization,

$$E_i = E_r$$

and for horizontal polarization.

$$E_i = -E_r$$

Diffraction

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

Scattering

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

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

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

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

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

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

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

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

Scattering by Rough Surfaces

Scattering on rough surfaces (Figure 4.13) is a process that is very important for wireless communications. Scattering theory usually assumes roughness to be *random*. However, in wireless communications it is common to also describe *deterministic*, possibly periodic, structures (e.g., bookshelves or windowsills) as rough. For ray-tracing predictions (see Section 7.5), “roughness” thus describes all (physically present) objects that are not included in the used maps and building plans.

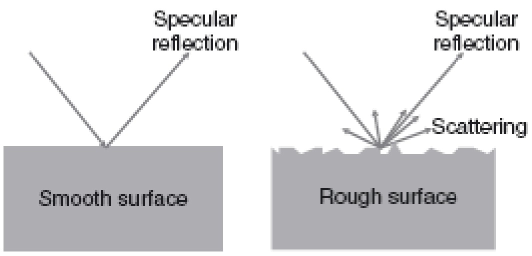


Figure 4.13 Scattering by a rough surface.

The justifications for this approach are rather heuristic: (i) the errors made are smaller than some other error sources in ray-tracing predictions and (ii) there is no better alternative. That being said, the remainder of this section will consider the mathematical treatment of genuinely

rough surfaces. This area has been investigated extensively in the last 30 years, mostly due to its great importance in radar technology. Two main theories have evolved: the Kirchhoff theory and the perturbation theory.

1. The Kirchhoff Theory

The Kirchhoff theory is conceptually very simple and requires only a small amount of information – namely, the probability density function of surface amplitude (height). The theory assumes that height variations are so small that different *scattering points* on the surface do not influence each other – in other words, that one point of the surface does not “cast a shadow” onto other points of the surface. This assumption is actually not fulfilled very well in wireless communications.

Assuming that the above condition is actually fulfilled, surface roughness leads to a reduction in power of the specularly reflected ray, as radiation is also scattered in other directions (see r.h.s. of Figure 4.13). This power reduction can be described by an *effective* reflection coefficient ρrough. In the case of Gaussian height distribution, this reflection factor

$$\rho_{\text{rough}} = \rho_{\text{smooth}} \exp\left[-2(k_0 \sigma_h \sin\psi)^2\right] \qquad (4.45)$$

becomes:

whereσh is the standard deviation of the height distribution, *k*0 is the wavenumber $2\pi/\lambda$, and ψ is the angle of incidence (defined as the angle between the wave vector and the surface). The term 2*k*0σh sin$\psi$is also known as Rayleigh roughness. Note that for grazing incidence *(*$\psi \approx 0$*)*, the effect of the roughness vanishes, and the reflection becomes specular again.

2 Perturbation Theory

The perturbation theory generalizes the Kirchhoff theory, using not only the probability density function of the surface height but also its spatial correlation function. In other words, it takes into account the question "how fast does the height vary if we move a certain distance along the surface?" (see Figure 4.14).

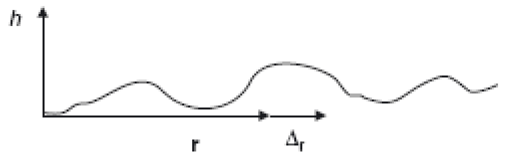


Figure 4.14 Geometry for perturbation theory of rough scattering.

Mathematically, the spatial correlation function is defined as:

$$\sigma_h^2 W(\Delta_r) = E_r\{h(r)h(r + \Delta_r)\} \quad (4.46)$$

where**r** and _r are (two-dimensional) location vectors, and *E*ris expectation with respect to **r**. We need this information to find whether one point on the surface can "cast a shadow" onto another point of the surface. If extremely fast amplitude variations are allowed, shadowing situations are much more common. The above definition enforces spatial statistical stationarity – i.e., the correlation is independent of the absolute location **r**. The correlation length *L*c is defined as the distance so that *W(L*c*)* = 0.5 · *W(*0*)*.

<u>Outdoor Propagation Models</u>

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.

<u>1. Okumura Model</u>

- 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 propogation 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, *G*(*ht*) is the base station antenna height gain factor, *G*(*hr*) is the mobile antenna height gain factor, *G AREA* 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.

2. Hata Model

The Hata model is an empirical formulation of the graphical path-loss data provided by the Okumura and is valid over roughly the same range of frequencies, 150-100. This empirical formula simplifies the calculation of path loss because it is closed form formula and it is not based on empirical curves for the different parameters. The standard formula for empirical path loss in urban areas under the Hata model is

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

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

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

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

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

else it is

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

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

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

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

Indoor Propagation Models

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

1. Partition Losses Inside a Floor (Intra-floor)

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

2. Partition Losses Between Floors (Inter-floor)

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

3. Log-distance Path Loss Model

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

$$PL(dB) = PL(d_0) + 10n \log10(d/d_0) + X\sigma$$

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

Multipath Propagation

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

Multipath & Small-Scale Fading

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

Fading

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

Multipath Fading Effects

In principle, the following are the main multipath effects:

1. Rapid changes in signal strength over a small travel distance or time interval.

2. Random frequency modulation due to varying Doppler shifts on different multipath signals.

3. Time dispersion or echoes caused by multipath propagation delays.

Factors Influencing Fading

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

(1) Multipath propagation – Multipath is the propagation phenomenon that 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 (quantified by *coherence bandwidth*), the received signal will be distorted.

Small Scale Fading

- Describes *rapid fluctuations* of the amplitude, phase of multipath delays of radio signal over short period of time or travel distance

- Caused by interference between two or more versions of the transmitted signal which arrive at the receiver at slightly different times.

- These waves are called multipath waves and combine at the receiver antenna to give a resultant signal which can vary widely in amplitude and phase.

Small Scale Multipath Propagation

Effects of multipath

- Rapid changes in the signal strength
- Over small travel distances,
- Over small time intervals
- Random frequency modulation due to varying Doppler shifts on different multiples signals
- Time dispersion (echoes) caused by multipath propagation delays

Multipath occurs because of

- Reflections
- Scattering

Multipath

At a receiver point Radio waves generated from the same transmitted signal may come from different directions

- with different propagation delays
- with (possibly) different amplitudes (random)
- with (possibly) different phases (random)
- with different angles of arrival (random).

These multipath components combine vectorially at the receiver antenna and cause the total signal

- to fade
- to distort

Impulse Response Model of a Multipath Channel

Mobile radio channel may be modeled as a linear filter with time varying impulse response in continuous time. To show this, consider time variation due to receiver motion and time varying impulse response $h(d, t)$ and $x(t)$, the transmitted signal.

The received signal $y(d, t)$ at any position d would be

$$y(d,t) = x(t) * h(d,t) = \int_{-\infty}^{\infty} x(\tau)\, h(d, t-\tau)\, d\tau$$

For a causal system: $h(d, t) = 0$, for $t<0$ and for a stable system ∞

Applying causality condition in the above equation, $h(d, t-\tau) = 0$ for $t-\tau < 0$ $\Rightarrow \tau > t$, i.e., the integral limits are changed to

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

Since the receiver moves along the ground at a constant velocity v, the position of the receiver is $d = vt$, i.e.,

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

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

$$y(t) = \int_{-\infty}^{t} x(\tau)\, h(vt, t-\tau)\, d\tau = x(t) * h(vt,t) = x(t) * h(d,t) \qquad (5.14)$$

It is useful to discretize the multipath delay axis τ of the impulse response into equal time delay segments called *excess delay bins*, each bin having a time delay width equal to

$(\tau_{i+1} - \tau_i) = \Delta\tau$ and $\tau_i = i\Delta\tau$ for $i \in \{0, 1, 2, ..N-1\}$,

where N represents the total number of possible equally-spaced multipath components, including the first arriving component. The useful frequency span of the model is $2/\Delta\tau$. The model may be used to analyze transmitted RF signals having bandwidth less than $2/\Delta\tau$. If there are N multipaths, maximum excess delay is given by $N\Delta\tau$.

If there are N multipaths, maximum excess delay is given by $N\Delta\tau$.

$$\{y(t) = x(t) * h(t, \tau_i) | i = 0, 1, ...N-1\} \tag{5.15}$$

Bandpass channel impulse response model is

$$x(t) \rightarrow h(t, \tau) = Re\{h_b(t, \tau)e^{j\omega_c t} \rightarrow y(t) = Re\{r(t)e^{j\omega_c t}\} \tag{5.16}$$

Baseband equivalent channel impulse response model is given by

$$c(t) \rightarrow \frac{1}{2}h_b(t, \tau) \rightarrow r(t) = c(t) * \frac{1}{2}h_b(t, \tau) \tag{5.17}$$

Average power is

$$\overline{x^2(t)} = \frac{1}{2}|c(t)|^2 \tag{5.18}$$

Small-Scale Multipath Measurements

Direct RF Pulse System

A wideband pulsed bistatic radar usually transmits a repetitive pulse of width *Tbb*s, and uses a receiver with a wide bandpass filter (*BW* = 2 *Tbb*Hz). The signal is then amplified, envelope detected, and displayed and stored on a high speed oscilloscope. Immediate measurements of the square of the channel impulse response convolved with the probing pulse can be taken. If the oscilloscope is set on averaging mode, then this system provides a local average power delay profile.

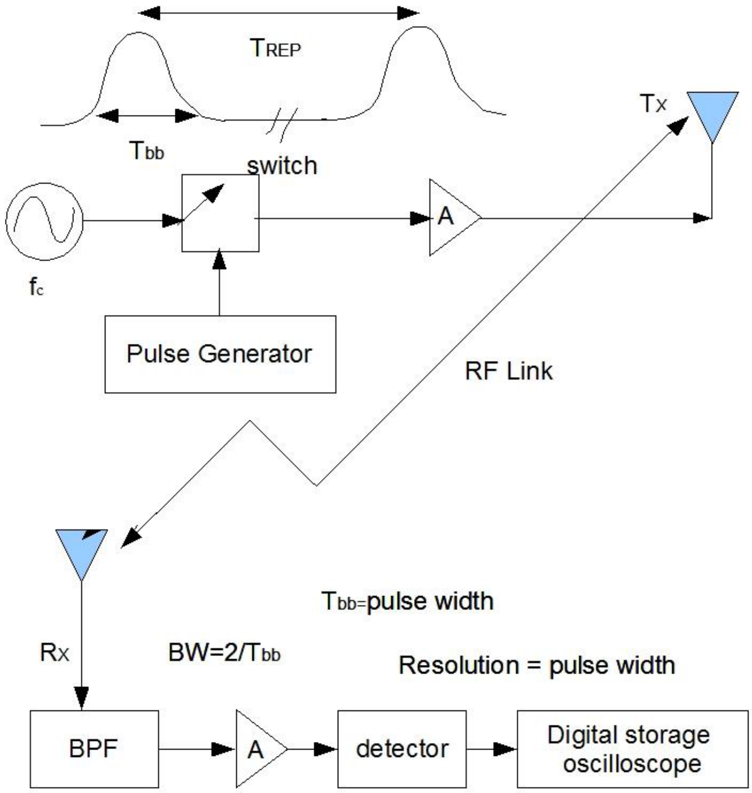
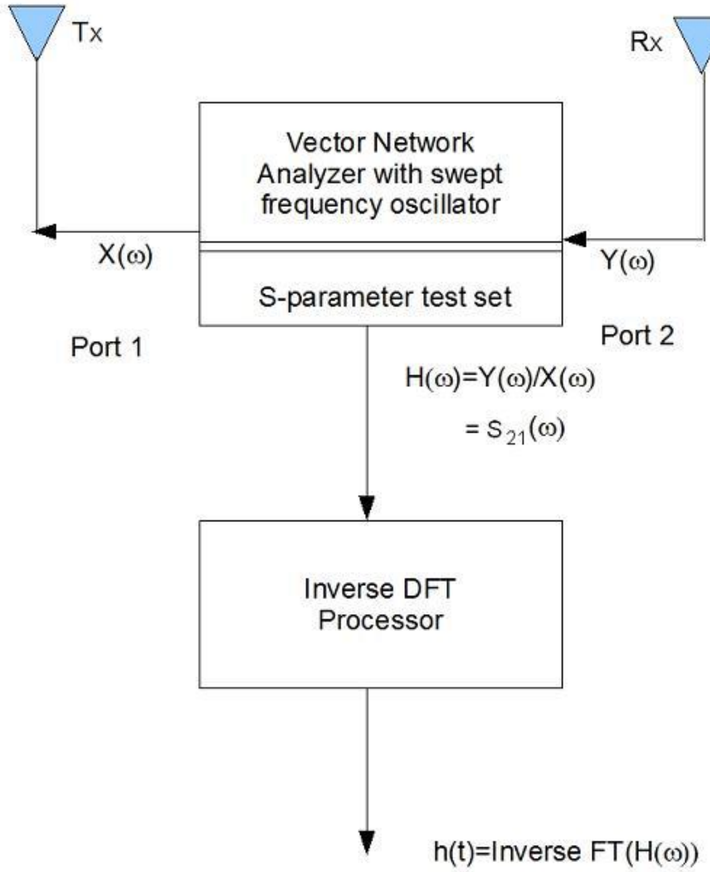


Figure 5.4: Direct RF pulsed channel IR measurement.

This system is subject to interference noise. If the first arriving signal is blocked or fades, severe fading occurs, and it is possible the system may not trigger properly.

Frequency Domain Channel Sounding

In this case we measure the channel in the frequency domain and then convert it into time domain impulse response by taking its inverse discrete Fourier transform (IDFT). A vector network analyzer controls a swept frequency synthesizer. An S parameter test set is used to monitor the frequency response of the channel. The sweeper scans a particular frequency band, centered on the carrier, by stepping through discrete frequencies. The number and spacing of the frequency step impacts the time resolution of the impulse response measurement. For each frequency step, the S-parameter test set transmits a known signal level at port 1 and monitors the received signal at port 2. These signals allow the analyzer to measure the complex response, $S21(\omega)$, of the channel over the measured frequency range. The $S21(\omega)$ measure is the measure of the signal flow from transmitter antenna to receiver antenna.



Frequency domain channel IR measurement

This system is suitable only for indoor channel measurements. This system is also non real-time. Hence, it is not suitable for time-varying channels unless the sweep times are fast enough.

Multipath Channel Parameters

To compare the different multipath channels and to quantify them, we define some parameters. They all can be determined from the power delay profile. These parameters can be broadly divided in to two types.

1. **Time Dispersion Parameters**

These parameters include the mean excess delay,rms delay spread and excess delay spread. The mean excess delay is the first moment of the power delay profile and is defined as where $a_k$is the amplitude, $\tau_k$is the excess delay and $P(\tau_k)$ is the power of the individual multipath signals.

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

The mean square excess delay spread is defined as

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

Since the rms delay spread is the square root of the second central moment of the power delay profile, it can be written as

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

As a rule of thumb, for a channel to be flat fading the following condition must be satisfied

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

where*TS* is the symbol duration. For this case, no equalizer is required at the receiver.

2. **Frequency Dispersion Parameters**

To characterize the channel in the frequency domain, we have the following parameters.

(1) Coherence bandwidth: it is a statistical measure of the range of frequencies over which the channel can be considered to pass all the frequency components with almost equal gain and linear phase. When this condition is satisfied then we say the channel to be flat.

Practically, coherence bandwidth is the minimum separation over which the two frequency components are affected differently. If the coherence bandwidth is considered to be the bandwidth over which the frequency correlation function is above

0.9, then it is approximated as

$$B_C \approx \frac{1}{50\sigma_\tau}.$$

However, if the coherence bandwidth is considered to be the bandwidth over which the frequency correlation function is above 0.5, then it is defined as

$$B_C \approx \frac{1}{5\sigma_\tau}.$$

The coherence bandwidth describes the time dispersive nature of the channel in the local area. A more convenient parameter to study the time variation of the channel is the coherence time. This variation may be due to the relative motion between the mobile and the base station or the motion of the objects in the channel.

(2) Coherence time: this is a statistical measure of the time duration over which the channel impulse response is almost invariant. When channel behaves like this, it is said to be slow faded. Essentially it is the minimum time duration over which two received signals are affected differently. For an example, if the coherence time is considered to be the bandwidth over which the time correlation is above 0.5, then it can be approximated as

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

where*fm*is the maximum doppler spread given be *fm*= *ν/λ*.

Another parameter is the Doppler spread (*BD*) which is the range of frequencies over which the receved Doppler spectrum is non zero.

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.

A. Fading Effects due to Multipath Time Delay Spread

1. Flat Fading

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

$$B_S \ll B_C$$

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

$$T_S \gg \sigma_\tau$$

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

2. Frequency Selective Fading

Frequency selective fading occurs when the signal bandwidth is more than the co-herence bandwidth of the mobile radio channel or equivalently the symbols duration of the signal is less than the rms delay spread.At the receiver, we obtain multiple copies of the transmitted signal, all attenuated and delayed in time. The channel introduces inter symbol interference.

B.Fading Effects due to Doppler Spread

1.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.

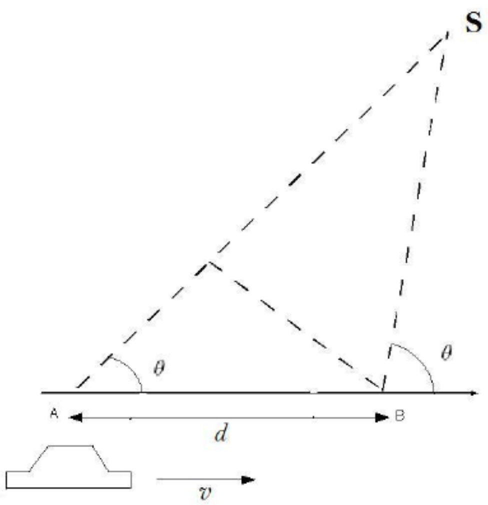
2.Slow Fading

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

3.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.

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 $\pm v_r c/ fo$or, $\pm v_r \lambda$.As the BS is located at an elevated place, a $\cos\varphi$ factor would also be multiplied with this. The exact scenario, as given in Figure is illustrated below.



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

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

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

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

Unit III

SIGNAL PROCESSING IN WIRELESS SYSTEMS

Principle of Diversity, Macro diversity,Micro diversity,Signal Combining Techniques, Transmit diversity, Equalizers- Linear and decision Feedback Equalizers, Review of Channel Coding and Speech coding techniques.

PRINCIPLE OF DIVERSITY

- The basic principle of diversity is that the receiver has multiple copies of the transmit signal, where each of the copies goes through a statistically independent channel.
- A first important step in designing diversity antennas is thus to establish a relationship between antenna spacing and the correlation coefficient.
- This relationship is different for BS antennas and MS antennas, and thus will be treated separately.

MACRODIVERSITY

- A separate base station (BS2) that is placed in such a way that the hill is not in the connection line between the MS and BS2. This in turn implies a large distance between BS1 and BS2, which gives rise to the word macrodiversity.
- The simplest method for macrodiversity is the use of on-frequency repeaters that receive the signal and retransmit an amplified version of it similar to **Simulcast** - the same signal is transmitted simultaneously from different BSs.
- Note that synchronization can only be obtained if the runtimes from the two BSs to the MS are known.
- Simulcast is also widely used for broadcast applications, especially digital TV.
- A disadvantage of simulcast is the large amount of signaling information that has to be carried on landlines. Synchronization information as well as transmit data has to be transported on landlines (or microwave links) to the BSs.
- The use of on-frequency repeaters is simpler than that of simulcast, as no synchronization is required. On the other hand, delay dispersion is larger, because

  (i) The runtime from BS to repeater, and repeater to MS is larger (compared with the runtime from a second BS),

  (ii) The repeater itself introduces additional delays due to the group delays ofelectronic components, filters, etc.

MICRODIVERSITY

The methods that can be used to combat small-scale fading, which are therefore called “microdiversity.”

The five most common methods are as follows:

1. Spatial diversity: several antenna elements separated in space.
2. Temporal diversity: transmission of the transmit signal at different times.
3. Frequency diversity: transmission of the signal on different frequencies.

4. Angular diversity: multiple antennas (with or without spatial separation) with different antenna patterns.
5. Polarization diversity: multiple antennas with different polarizations (e.g., vertical and horizontal).

Spatial Diversity - Spatial diversity is the oldest and simplest form of diversity and is most widely used. The transmit signal is received at several antenna elements.

Irrespective of the processing method, the performance is influenced by correlation of the signals between the antenna elements. A large correlation between signals at antenna elements is undesirable, as it decreases the effectiveness of diversity.

Temporal Diversity

Temporal diversity can be realized in different ways:

1. **Repetition coding:**
   - **T**his is the simplest form. The signal is repeated several times,where the repetition intervals are long enough to achieve decorrelation. This obviously achieves diversity, but is also highly bandwidth inefficient.
   - Spectral efficiency decreases by a factor that is equal to the number of repetitions.

2. **Automatic Repeat request (ARQ)**:
   - The receiver sends a message to the TX toindicate whether it received the data with sufficient quality. If this is not the case, then the transmission is repeated (after a wait period that achieves decorrelation).
   - The spectral efficiency of ARQ is better than that of repetition coding, since it requires multiple transmissions only when the first transmission occurs in a bad fading state, while for repetition coding, retransmissions occur always.
   - On the downside, ARQ requires a feedback channel.

3. **Combination of interleaving and coding**:

   - **This is an** advanced version of repetitioncoding is forward error correction coding with interleaving.
   - The different symbols of a codeword are transmitted at different a time, which increases the probability that at least some of them arrive with a good SNR.
   - The transmitted codeword can then be reconstructed.

Frequency Diversity

- In frequency diversity, the same signal is transmitted at two (or more) different frequencies.
- If the frequencies are spaced apart by more than the coherence bandwidth of the channel, then their fading is approximately independent, and the probability is low that the signal is in a deep fade at both frequencies simultaneously.

Angle Diversity

- A fading dip is created when MPCs, which usually come from different directions, interfere destructively.
- Angular diversity is usually used in conjunction with spatial diversity; it enhances the decorrelation of signals at closely spaced antennas. Different antenna patterns can be achieved very easily.
- The different types of antennas have different patterns. But even identical antennas can have different patterns when mounted close to each other.

Polarization Diversity

- Horizontally and vertically polarized MPCs propagate differently in a wireless channel,6 as the reflection and diffraction processes depend on polarization.
- The fading of signals with different polarizations is statistically independent. Receiving both polarizations using a dual-polarized antenna, and processing the signals separately, offers diversity.
- This diversity can be obtained without any requirement for a minimum distance between antenna elements.
- Let us now consider more closely the situation where the transmit signal is vertically polarized, while the signal is received in both vertical and horizontal polarization. In that case, fading of the two received signals is independent, but the average received signal strength in the two diversity branches is not identical.
- Various antenna arrangements have been proposed in order to mitigate this problem.

SIGNAL COMBINING TECHNIQUE

• Phase correction causes the signal amplitudes to add up, while, on the other hand, noise is added incoherently, so that noise powers add up.

• For amplitude weighting, two methods are widely used: Maximum Ratio Combining (MRC) weighs all signal copies by their amplitude. It can be shown that (using some assumptions) this is an optimum combination strategy. An alternative is Equal Gain Combining (EGC), where all amplitude weights are the same (in other words, there is no weighting, but just a phase correction).

Maximum Ratio Combining

- Signals from all of the m branches are weighted according to their individual signal voltage to noise power ratios and then summed. Refer figure 3.1.
- Individual signals must be cophased before being summed, which generally requires an individual receiver and phasing circuit for each antenna element.
- This produces an output SNR equal to the sum of all individual SNR. The advantage of producing an output with an acceptable SNR is even when none of the individual signals are themselves acceptable.
- Modern DSP techniques and digital receivers are now making this optimal form, as it gives the best statistical reduction of fading of any known linear diversity combiner.

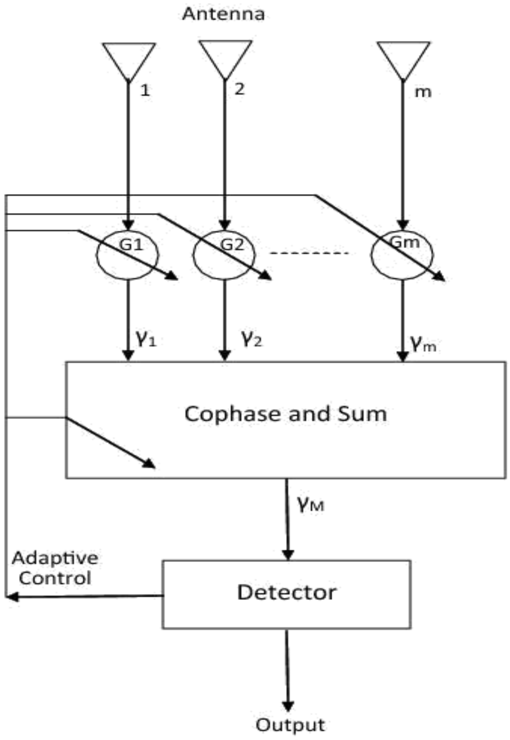


Figure 3.1: Maximum ratio combiner

Equal Gain Combining:

- In some cases it is not convenient to provide for the variable weighting capability required for true maximal ratio combining.Figure 3.2 depicts Equal Gain Combining
- In such cases, the branch weights are all set unity, but the signals from each branch are co-phased to provide equal gain combining diversity.
- It allows the receiver to exploit signals that are simultaneously received on each branch. Performance of this method is marginally inferior to max-imal ratio combining and superior to Selection diversity.

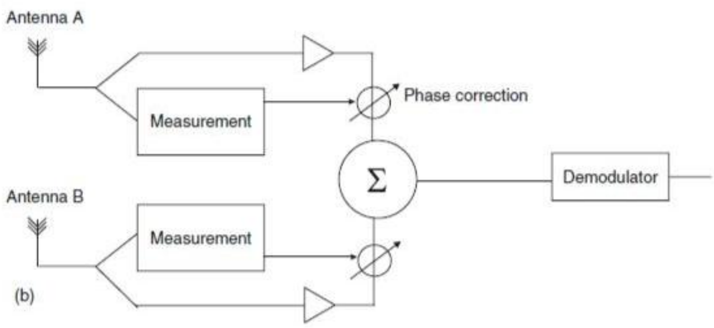


Figure 3.2 :Equal Gain Combining

- If the branches are statistically independent, then the moment-generating function of the total SNR can be computed as the product of the characteristic functions of the branch SNRs.

Optimum Combining

- In order to maximize the Signal-to-Interference-and-Noise Ratio (SINR), the weights should then be determined according to a strategy called optimum combining. The first step is the determination of the correlation matrix of noise and interference at the different antenna elements:

$$\mathbf{R} = \sigma_n^2 \mathbf{I} + \sum_{k=1}^{K} E\{\mathbf{r}_k \mathbf{r}_k^{\dagger}\}$$

  Where,

  $\mathbf{r}_k$ is the receive signal vector of the k^{th} interferer.

- The vector containing the optimum receive weights is then $\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{h}_d$
- A noise-limited system the correlation matrix becomes a (scaled) identity matrix, and optimum combining reduces to MRC.
- Optimum combining of signals from Nr diversity branches gives N_r degrees of freedom. This allows interference from $N_r - 1$ interferers to be eliminated.
- Alternatively, $N_s \leq N_r - 1$ interferers can be eliminated, while the remaining $N_r - N_s$ antennas behave like "normal" diversity antennas that can be used for noise reduction.

TRANSMIT DIVERSITY

- Multiple antennas can be installed at just one link end (usually the BS).
- For the uplink transmission from the MS to BS, multiple antennas can act as receive diversity branches.
- For the downlink, any possible diversity originates at the transmitter.
- Transmit diversity is the way of transmitting signals from several transmitting antennas and achieve a diversity effect with it.

Case 1: Transmitter Diversity with Channel State Information - TX knows the channel perfectly

- Here, we find that (at least for the noise-limited case) there is a complete equivalence between transmit diversity and receive diversity.

Case 2:**Transmitter Diversity without Channel State Information -** Channel State Information (CSI) is not available at the TX

- It is not possible to simply transmit weighted copies of the same signal from different transmit antennas, because we cannot know how they would add up at the RX.
- In order to give benefits, transmission of the signals from different antenna elements has to be done is such a way that it allows the RX to distinguish different transmitted signal components.

Method1- delay diversity:

- The signals transmitted from different antenna elements are delayed copies of the same signal. This makes sure that the effective impulse response is delay dispersive, even if the channel itself is flat fading.
- The effective impulse response of the channel then becomes

$$h(\tau) = \frac{1}{\sqrt{N_t}} \sum_{n=1}^{N_t} h_n \delta(\tau - nT_s)$$

Where, h_n- gains from the nth transmit antenna to the receive antenna, and

- The impulseresponse has been normalized so that total transmit power is independent of the number of antenna elements.
- The signals from different transmit antennas to the RX act effectively as delayed MPCs.
- The delay between signals transmitted from different antenna elements should be at least as large as the maximum excess delay of the channel.

Method 2 - phase-sweeping diversity:

- In this method, especially useful, if there are only two antenna elements, the same signal is transmitted from both antenna elements.
- However, one of the antenna signals undergoes a time-varying phase shift. This means that at the RX the received signals add up in a time-varying way;
- The reason for this is that – even if the TX, RX, and the IOs are stationary – the signal does not remain stuck in a fading dip.

EQUALIZER

- An equalizer within a receiver compensates for the average range of expected channel amplitude and delay characteristics.
- Equalizers must be adaptive since the channel is generally unknown and time varying.
- Equalization compensates for inter symbol interference (ISI) created by multipath within time dispersive channels.
- If the modulation bandwidth exceeds the coherence bandwidth of the radio channel, ISI occurs and modulation pulses are spread in time..

DECISION FEEDBACK EQUALIZER

- The basic idea behind decision feedback equalization is that once an information symbol has been detected and decided upon, the 1ST that it induces on future symbols can be estimated and subtracted out before detection of subsequent symbols.
- The DFE can be realized in either the direct transversal form or as a lattice filter. It consists of a feed forward filter (FFF) and a feedback filter (FBF).
- The FBF is driven by decisions on the output of the detector, and its coefficients can be adjusted to cancel the ISI on the current symbol from past detected symbols.
- The equalizer has $N^1 + N^2 + I$ taps in the feed forward filter and N3 taps in the feedback filter, and its output can be expressed as:

$$\hat{d}_k = \sum_{n=-N_1}^{N_2} c_n^* y_{k-n} + \sum_{i=1}^{N_3} F_i d_{k-i}$$

Where,

c^*_n, and y_n, are tap gains and the inputs, respectively .

F_1 are tap gains for the feedback filter, and

d_1 (ic_k) is the previous decision made on the detected signal.

The minimum mean squared error a DFE can achieve is,

$$E\left[|e(n)|^2\right]_{min} = \exp\left\{\frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \ln\left[\frac{N_0}{\left|F(e^{j\omega T})\right|^2 + N_0}\right] d\omega\right\}$$

The equation can be simplified as

$$E\left[|e(n)|^2\right] = \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} \frac{N_0}{\left|F(e^{j\omega T})\right|^2 + N_0} d\omega$$

- The lattice implementation of the DFE is equivalent to a transversal DFE having a feed forward filter of length N1 and a feedback filter of length N2, where N1 > N2. Another form of DFE proposed by Belfiore and Park is called a predictive DFE, and is shown in Figure below 3.4:

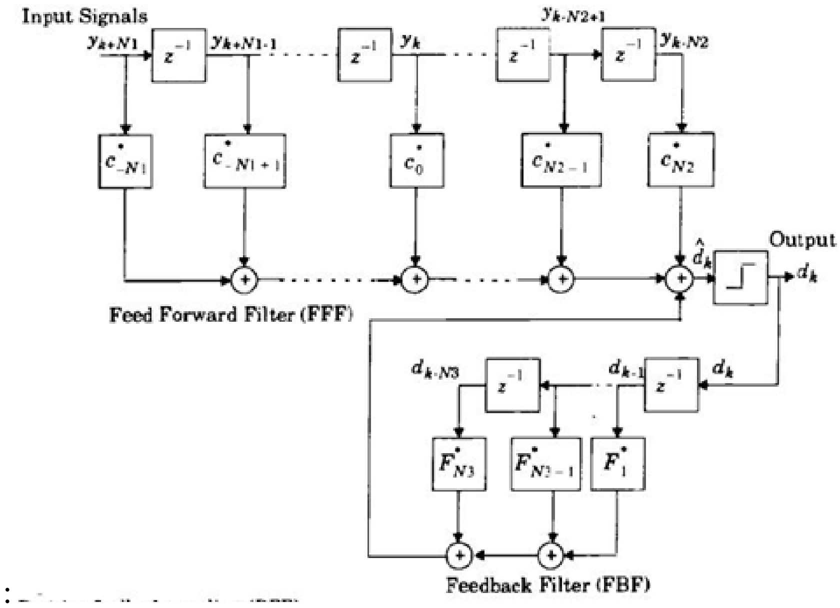


Figure 3.4 : Predictive DFE

- It also consists of a feed forward filter (FFF) as in the conventional DFE. However, the feedback filter (FBF) is driven by an input sequence formed by the difference of the output of the detector and the output of the feed forward filter.
- Hence, the FBF here is called a noise predictor because it predicts the noise and the residual 151 contained in the signal at the FFF output and subtracts from it the detector output after some feedback delay.
- The predictive DFE performs as well as the conventional DFE as the limit in the number of taps in the FFF and the FBF approach infinity.
- The FEF in the predictive DFE can also be realized as a lattice structure The RLS lattice algorithm can be used in this case to yield fast convergence.

CHANNEL CODING

- In channel coding, redundant data bits are added in the transmitted message so that if an instantaneous fade occurs in the channel, the data may still be recovered at the receiver without the request of retransmission.
- A channel coder maps the transmitted message into another specific code sequence containing more bits. Coded message is then modulated for transmission in the wireless channel.
- Channel Coding is used by the receiver to detect or correct errors introduced by the channel.
- Codes that used to detect errors, are error detection codes. Error correction codes can detect and correct errors.
- The error correction and detection codes are classified into three groups based on their structure.

  1. Block Code
  2. Convolution Code
  3. Concatenated Code.

Block Code

- Block codes are forward error correction (FEC) codes that enable a limited number of errors to be detected and corrected without retransmission.
- Block codes can be used to improve the performance of a communications system when other means of improvement (such as increasing transmitter power or using a more sophisticated demodulator) are impractical.
- In block codes, parity bits are added to blocks of message bits to make codewords or code blocks.
- In a block encoder, k information bits are encoded into n code bits. A total of n k redundant bits are added to the k information bits for the purpose of detecting and correcting errors.
- The block code is referred to as an (n, k) code, and the rate of the code is de ned as R_c = k=n and is equal to the rate of information divided by the raw channel rate.

Parameters

Code Rate (R_c): R_c = k=n.

Code Distance (d): Distance between two code words is the number of elements in which two code words C_i and C_j di ers denoted by d (C_i; C_j). If the code used is binary, the distance is known as **'Hamming distance'**.

For example d(10110, 11011) is 3. If the code 'C' consists of the set of code words, then the minimum distance of the code is given by d_{min} = min fd (C_i; C_j)g.

Code Weight (w): Weight of a code word is given by the number of nonzero elements in the code word. For a binary code, the weight is basically the number of 1s in the code word. For example weight of a code 101101 is 4.

Ex 1: The block code C = 00000, 10100, 11110, 11001 can be used to represent two bit binary numbers as:

00 { 00000
01 { 10100
10 { 11110
11 { 11001

Here number of codewords is 4, k = 2, and n = 5.

To encode a bit stream 1001010011

1. First step is to break the sequence in groups of two bits, i.e., 10 01 01 00 11
2. Next step is to replace each block by its corresponding codeword, i.e., 11110 10100 10100 00000 11001

Here, d_{min} = min fd (C_i; C_j)g = 2.

Properties of Block Codes

(a**) Linearity:** Suppose C_i and C_j are two code words in an (n, k) block code. Let $_1$ and $_2$ be any two elements selected from the alphabet. Then the code is said to be linear if and only if $_1C_1$ + $_2C_2$ is also a code word. A linear code must contain the all-zero code word.

(b) Systematic: A systematic code is one in which the parity bits are appended to the end of the information bits. For an (n, k) code, the rst k bits are identical to the information bits, and the remaining n k bits of each code word are linear combinations of the k information bits.

(c) Cyclic: Cyclic codes are a subset of the class of linear codes which satisfy the following cyclic shift property: If C = [C_{n1}; C_{n2}; :::; C_0] is a code word of a cyclic code, then [C_{n2}; C_{n3}; :::; C_0; C_{n1}], obtained by a cyclic shift of the elements of C, is also a code word. That is, all cyclic shifts of C are code words.

BCH Codes:

- BCH code is one of the most powerful known class of linear cyclic block codes, known for their multiple error correcting ability, and the ease of encoding and decoding.
- It's block length is n = 2^m 1 for m 3 and number of errors that they can correct is bounded by t

< (2^m 1)/2. Binary BCH codes can be generalized to create classes of non binary codes which use m bits per code symbol.

Reed Solomon (RS) Codes:

- Reed-Solomon code is an important subset of the BCH codes with a wide range of applications in digital communication and data storage.
- Typical application areas are storage devices (CD, DVD etc.), wireless communications, digital TV, high speed modems.
- It's coding system is based on groups of bits, such as bytes, rather than individual 0 and 1. This feature makes it particularly good at dealing with burst of errors: six consecutive bit errors.
- Block length of these codes is n = 2^m 1, and can be extended to 2^m or $2^m + 1$.
- Number of parity symbols that must be used to correct e errors is n k = 2e. Minimum distance d_{min} = 2e + 1, and it achieves the largest possible d_{min} of any linear code.

Convolutional Codes

- A continuous sequence of information bits is mapped into a continuous sequence of encoder output bits.
- A convolutional code is generated by passing the information sequence through a nine stage shift register. A convolutional encoder with n=2 and k=1 is shown in figure 4.5.
- Shift register contains 'N' k-bit stages and m linear algebraic function generators based on the generator polynomials.

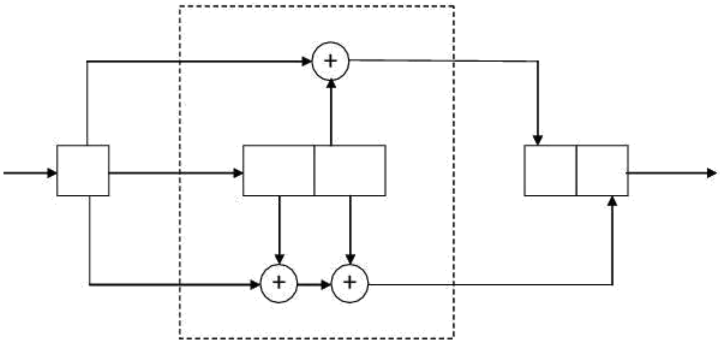


Figure 4.5:A convolutional encoder with n=2 and k=1.

- Input data is shifted into and along the shift register, k-bits at a time.
- Number of output bits for each k-bit user input data sequence is n bits, so the code rate R_c = k=n.
- The shift register of the encoder is initialized to all-zero-state before encoding operation starts.

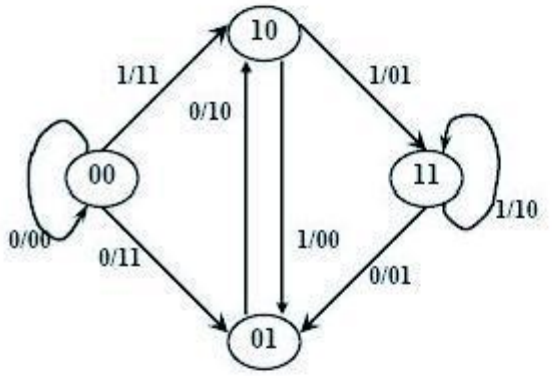


Figure 4.6: State diagram representation of a convolutional encoder

Generator Matrix The generator matrix for a convolutional code is semi-infinite since the input is semi-infinite in length. Hence, this may not be a convenient way of representing a convolutional code.

Generator Polynomials —Here, we specify a set of n vectors, one for each of the n modulo-2adders used. Each vector of dimension 2k indicates the connection of the encoder to that modulo-2 adder. A 1 in the ith position of the vector indicates that the corresponding shift register stage is connected and a 0 indicates no connection.

Logic Table —A logic table can be built showing the outputs of the convolutional encoder andthe state of the encoder for the input sequence present in the shift register.

State Diagram -—Since the output of the encoder is determined by the input and the currentstate of the encoder, a state diagram can be used to represent the encoding process. The state diagram is simply a graph of the possible states of the encoder and the possible transitions from one state to another.Figure 4.6 shows the State diagram representation of a convolutional encoder

Thee Diagram —The tree diagram shows the structure of the encoder in the form of a tree withthe branches representing the various states and the outputs of the coder.

Trellis Diagram —All branches emanating from two nodes having the same state are identical in the sense that they generate identical output sequences. This means that the two nodes having the same label can be merged. By doing this throughout the tree diagram, we can obtain another diagram called a Trellis diagram which is a more compact representation.

Concatenated Codes

- Concatenated codes are basically concatenation of block and convolutional codes. It can be of two types: serial and parallel codes. Below, we discuss a popular parallel concatenated code, namely, turbo code.
- A turbo encoder is built using two identical convolutional codes of special type with parallel concatenation.
- An individual encoder is termed a com-ponent encoder. An interleaver separates the two component encoders.
- The inter-leaver is a device that permutes the data sequence in some predetermined manner.

- Figure 4.7 below shows the block diagram of a turbo encoder using two identical encoders.
- The first encoder outputs the systematic V_0 and recursive convolutional V_1 sequences while the second encoder discards its systematic sequence and only outputs the recursive convolutional V_2 sequence.
- Encoders are also categorized as systematic or non-systematic. If the component encoders are not identical then it is called an asymmetric turbo code.

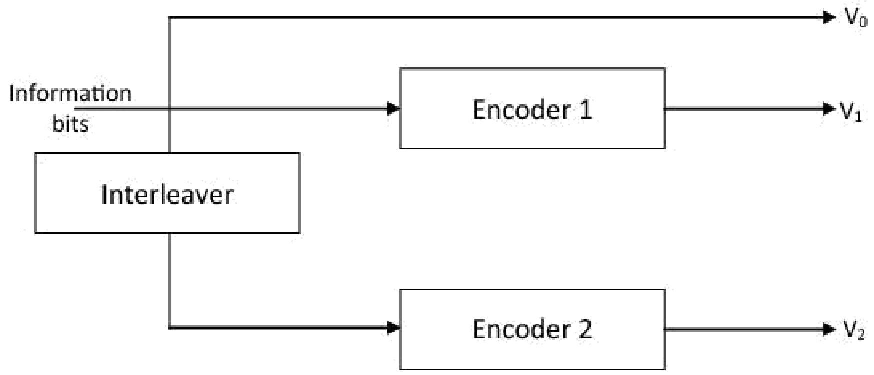


Figure 4.7: Block diagram of a turbo encoder

SPEECH CODING

- Determines the quality of the recovered speech and the capacity of the system.
- In mobile communication systems, bandwidth is a precious commodity, and service providers are continuously met with the challenge of accommodating more users within a limited allocated bandwidth which is very difficult.
- The lower the bit rate at which the coder can deliver toll quality speech, the more speech channels can be compressed within a given bandwidth. For this reason, manufacturers and service providers are continuously in search of speech codes that will provide toll quality speech at lower bit rates.
- To make speech coding practical, implementations must consume little power and provide tolerable, if not excellent speech quality.
- The goal of all speech coding systems is to transmit speech with the highest possible quality using the least possible channel capacity. The more complex an algorithm is, the more its processing delay and cost of implementation.
- Types of speech coders are given below in figure 4.8

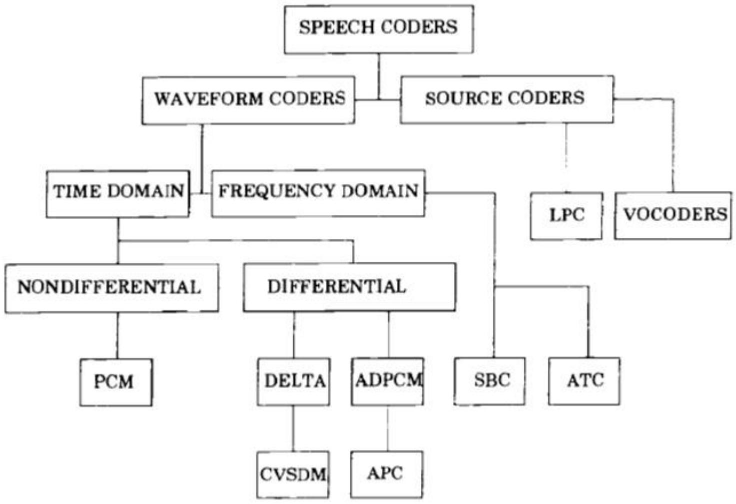


Figure 4.8: Types of speech coders

- Based on the means by which they achieve compression, speech coders are broadly classified into two categories:
  - ➢ Waveform Coders and
  - ➢ Vocoders.
- Waveform coders essentially strive to reproduce the time waveform of the speech signal as closely as possible.
- Examples of waveform coders include pulse code modulation (PCM), differential pulse code modulation (DPCM), and adaptive differential pulse code modulation (ADPCM), delta modulation (DM), and continuously variable slope delta modulation (CVSDM), and adaptive predictive coding.

Advantages:

- ➢ They are robust for a wide range of speech characteristics
- ➢ Suitable for noisy environments.
- ➢ Preserved with minimal complexity
- ➢ Achieves only moderate economy in transmission bit rate.

Advantages of Vocoders:

- ❖ Achieve very high economy in transmission bit rate
- ❖ Generally more complex.
- ❖ They are based on using a priori knowledge about the signal to be coded,

Characteristics of Speech Signals

- Some of the properties that are most often utilized in coder design include:
  - ❖ non-uniform probability distribution of speech amplitude,
  - ❖ the nonzero autocorrelation between successive speech samples,
  - ❖ the nonflat nature of the speech spectra, the existence of voiced and unvoiced segments in speech, and

- ❖ Thequasiperiodicity of voiced speech signals.
- ❖ The most basic property of speech waveforms that is exploited by all speech coders is that they are bandlimited. A finite bandwidth means that it can be time discretized (sampled) at a finite rate and reconstructed completely from its samples, provided that the sampling frequency is greater than twice the highest frequency component in the low pass signal.

- **Probability Density Function**
  - ➢ The pdf of a speech signal is in general characterized by a very high probability of near-zero amplitudes, a significant probability of very high amplitudes, and a monotonically decreasing function of amplitudes between these extremes.
  - ➢ The exact distribution, however, depends on the input bandwidth and recording conditions. The two-sided exponential (Laplacian) provides a good approximation to the long-term pdf of telephone quality speech signals

$$p(x) = \frac{1}{\sqrt{2}\sigma_x} \exp(-\sqrt{2}|x|/\sigma_x)$$

- **Autocorrelation Function (ACF)**
  - ➢ In every sample of speech, there is a large component that is easily predicted from the value of the previous samples with a small random error.
  - ➢ All differential and predictive coding schemes are based on exploiting this property.
  - ➢ The autocorrelation function (ACF) gives a quantitative measure of the closeness or similarity between samples of a speech signal as a function of their time separation. This function is mathematically defined,

$$C(k) = \frac{1}{N} \sum_{n=0}^{N-|k|-1} x(n)x(n+|k|)$$

where x (k) represents the k th speech sample. The autocorrelation function is often normalized to the variance of the speech signal and hence is constrained to have values in the range (-1,1) with C (0) = I . signals have an adjacent sample correlation, C (I) , as high as 0.85 to 0.9.

- **Power Spectral Density Function (PSI))**
  - ➢ The nonflat characteristic of the power spectraldensity of speech makes it possible to obtain significant compression by coding speech in the frequency domain.
  - ➢ The nonflat nature of the PSD is basically a frequency domain manifestation of the nonzero autocorrelation property.
  - ➢ It should be noted that the high frequency components, though insignificant in energy are very important carriers of speech information and hence need to be adequately represented in the coding system.

- A qualitative measure of the theoretical maximum coding gain that can be obtained by exploiting the nonilat characteristics of the speech spectra is given by the spectral flatness measure (SFM). The SFM is defined as the ratio of the arithmetic to geometric mean of the samples of the PSD taken at uniform intervals in frequency. Mathematically,

$$SFM = \frac{\left[\frac{1}{N}\sum_{k=1}^{N} S_k^2\right]}{\left[\prod_{k=1}^{N} S_k^2\right]^{\frac{1}{N}}}$$

where, Sk is the k th frequency sample of the PSD of the speech signal. Typically, speech signals have a long-term SFM value of 8 and a short-time SFM value varying widely between 2 and 500.

Unit IV

SIGNAL PROCESSING IN WIRELESS SYSTEMS

Spread Spectrum Systems- Cellular Code Division Multiple Access Systems- Principle, Power control, Effects of multipath propagation on Code Division Multiple Access,Orthogonal Frequency Division Multiplexing – Principle, Cyclic Prefix, Transceiver implementation, Second Generation(GSM, IS–95) and Third Generation Wireless Networks and Standards.

Spread Spectrum Systems

A spread spectrum modulation scheme is a digital modulation technique that utilizes a transmission bandwidth much greater than the modulating signal bandwidth, independently of the bandwidth of modulating signal. Spread-spectrum techniques are methods by which a signal (e.g. an electrical, electromagnetic, or acoustic signal) generated in a particular bandwidth is deliberately spread in the frequency domain, resulting in a signal with a wider bandwidth. These techniques are used for a variety of reasons, including the establishment of secure communications, increasing resistance to natural interference and jamming, to prevent detection, and to limit power flux density

Spread Spectrum System Criteria

For a system to be described as a spread spectrum system the following two criteria must be satisfied:

i. The transmitted signal must occupy a bandwidth much greater than the bandwidth of the modulating signal (i.e. the input signal to the system).

ii. The bandwidth occupied by the transmitted signal must be determined by a prescribed waveform and not by the modulating frequency (i.e. carrier frequency)

Reasons for use of Spread Spectrum Systems

There are three major reasons for the use of spread spectrum techniques in communication systems today.

i. They aid privacy of the transmission, since the spectral density of the spread spectrum may be less than the noise spectral density of the receiver.

ii. The de-spreading process in the receiver will spread the spectra of unwanted narrowband signals, thus improving interference rejection.

iii. The effect on a spread spectrum receiver, that receives a spread spectrum from a different spread spectrum system using the same frequency bands but implementing a different spreading pattern, approximates to noise in the receiver.

The principal advantages of spread-spectrum transmission are as follows:

1. Spread-spectrum signals can be overlaid onto bands where other systems are already operating, with minimal performance impact to or from the other systems.

2. The anti-multipath characteristics of spread-spectrum signaling and reception techniques are attractive in applications where multipath is likely to be prevalent. (Achieving good performance in frequency-selective fading may require the use of a RAKE receiver, which is in effect a matched filter for a multipath channel.)

3. The anti-interference characteristics of spread spectrum are important in some applications, such as networks operating on manufacturing floors, where the signal interference environment can be harsh.

4. Cellular systems designed with CDMA spread-spectrum technology offer greater operational flexibility and possibly a greater overall system capacity than those of systems built on TDMA access methods.

5. The convenience of unlicensed spread-spectrum operation in ISM bands is attractive to manufacturers and users alike.

6. High-resolution ranging achievable with spread-spectrum technology is attractive for emerging location-aware broadband wireless ad hoc networks.

   Spread spectrum is attractive to military users for the following fundamental reasons.

1. It can provide Low Probability of Interception (LPI).

2. It has Anti Jamming (AJ) properties.

3. It can allow for simultaneous multiple signal access.

4. It has inherent properties of secure communications, making it difficult for adversaries to recover the signal intelligence even if they are capable of detecting the presence of such signals.

Spread spectrum techniques to find their way into commercial applications. The benefits that commercial users are deriving are parallel to those that motivated the military users. If attempt to define a parallel set of benefits that Spread spectrum systems offer to the commercial environment, they are:

5. Low transmit power density function per Hertz.

   This helps the spectral contamination issue making it easier to comply with regulatory requirements. In addition, the interference generated from each operating channel to others is reduced.

6. Interference rejection capability.

   This provides a performance advantage over classical narrowband transmissions due to the fact that spread spectrum systems are more immune to interference sources.

7. Simultaneous multiple system transmissions over the same frequency bandwidth.

   This allows for code division multiplexing schemes that allow for simultaneous transmission of more than one signal. Communication security or less expensive privacy protection aspects are also important to commercial networks. The primary spread spectrum waveforms are Frequency Hoping (FH) and Direct Sequence (DS) spread spectrum. Both of these techniques capitalize on the utilization of a wide frequency bandwidth for their communications in order to extract the stated benefits

8. Use of the signal to measure propagation delay, which is important for radar systems to determine the position and direction of targets;

9. It is possible to hide the signal in background noise, so as to make it difficult to detect by those for which the spreading code is unknown.

10. Useful as a means of achieving diversity in channels with multipath.

TYPES OF SPREAD SPECTRUM

There are three ways to spread the bandwidth of the signal:

- **Direct sequence.** The digital data is directly coded at a much higher frequency. The code is generated pseudo-randomly, the receiver knows how to generate the same code, and correlates the received signal with that code to extract the data.
- **Frequency hopping**. The signal is rapidly switched between different frequencies within the hopping bandwidth pseudo-randomly, and the receiver knows beforehand where to find the signal at any given time.
- **Time hopping.** The signal is transmitted in short bursts pseudo-randomly, and the receiver knows beforehand when to expect the burst.

Direct sequence spread spectrum transmitter and receiver.

- A simple block diagram of a BPSK DS-SS system is shown in Figure 4.1.The BPSK modulated data, represented by d(t) is spread after multiplication by a pseudorandom (PN) sequence with a bandwidth much greater than the information signal. The transmitted signal x(t) can be expressed as

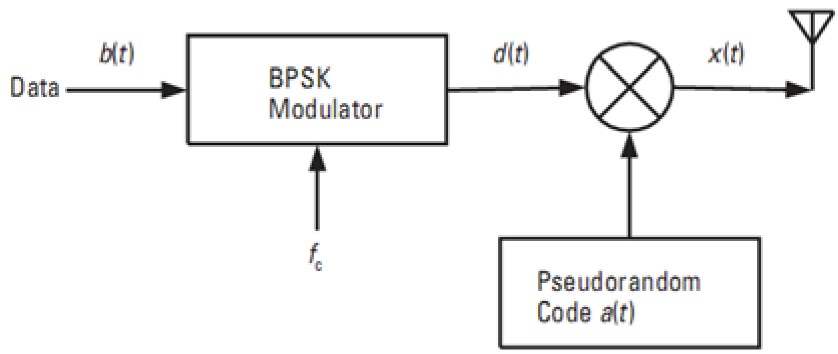


Figure 4.1 Block diagram of DS-SS Transmitter

$$x(t) = \sqrt{2P}b(t)a(t)\exp(J\omega t)..................(\ 1\ )$$

where a(t) is the spreading PN sequence with chips of ±1 of duration and code length of $N = T_b/T_c$. The signal spectrum at various stages of transmission is shown in Figure 4.2. As an information signal is multiplied by the PN sequence, its energy is spread over a wide bandwidth while the total signal energy remains constant. If the spreading ratio is large enough, the spread signal appears as very low power noise.

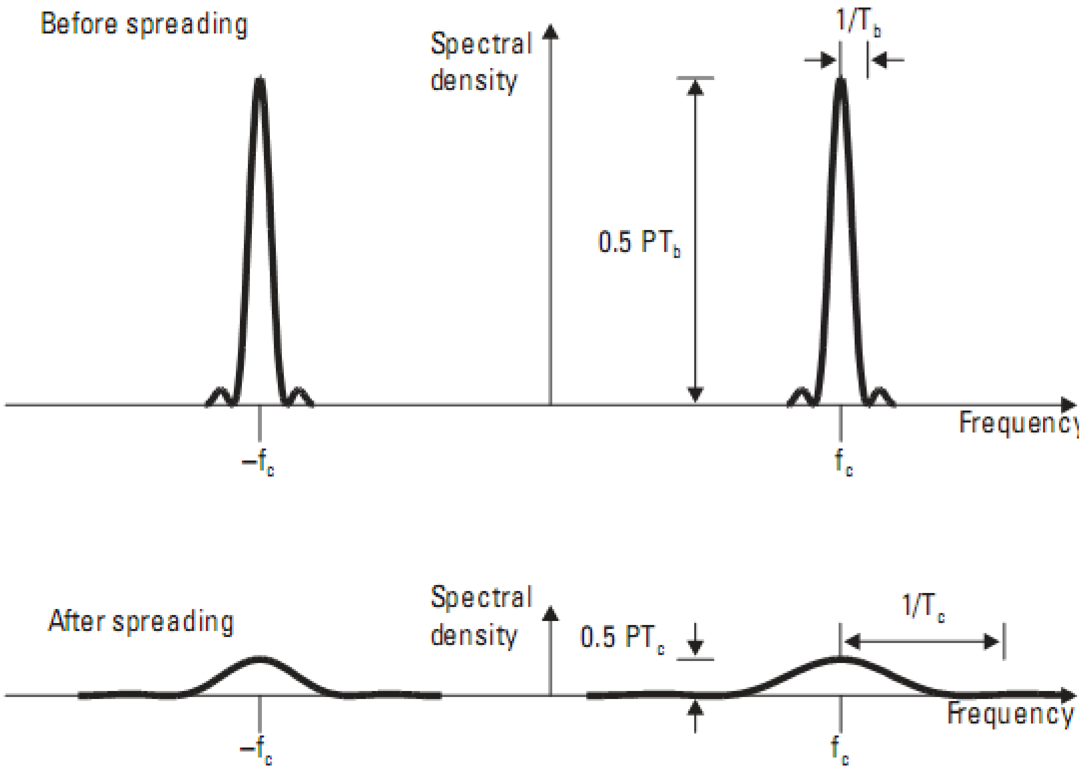


Figure 4.2 Signal spectrum of DS-SS modulation before and after spreading

A block diagram of a DS-SS receiver is shown in Figure 4.3.

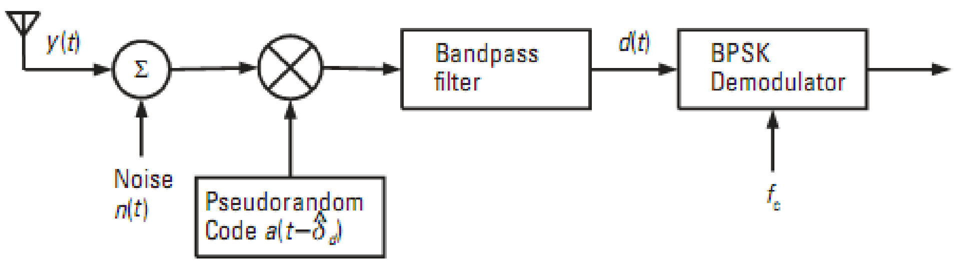


Figure 4.3 Block diagram of a DS-SS Receiver

FHSS is a spread spectrum communications technique whereby the carrier frequency of a transmitted message signal can be one of M possible carrier frequencies.

Each of the M possible carrier frequencies are determined by a PN sequence. Each bit of these sequences can be clustered so as to obtain a larger symbol size.

As a case in point, if each of the three obtained PN sequences are clustered, there exist 7 possible M carrier frequencies at the transmitter.

Generally, if N consecutive one bit symbols are clustered together, the highest possible number of M carrier frequencies is given by $M = 2^N - 1$

The sequence with which each of these frequency hops occur is known as the **frequency hopping pattern**

The entire set of all the available carrier frequencies is known as the **hopset**with the value M referred to as the **hopset size**

The hop pattern of a typical FHSS system is provided in Figure 4.4 below

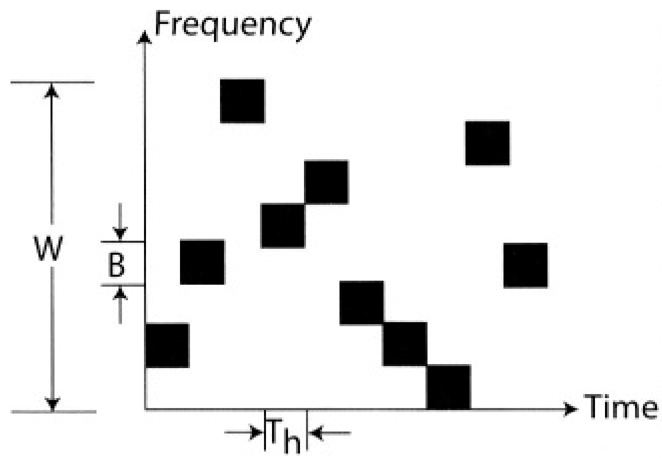


Figure 4.4: Frequency Hopping Pattern of a FHSS system

FHSSTransmitter
The FHSS-Block diagram is shown below in figure 4.5

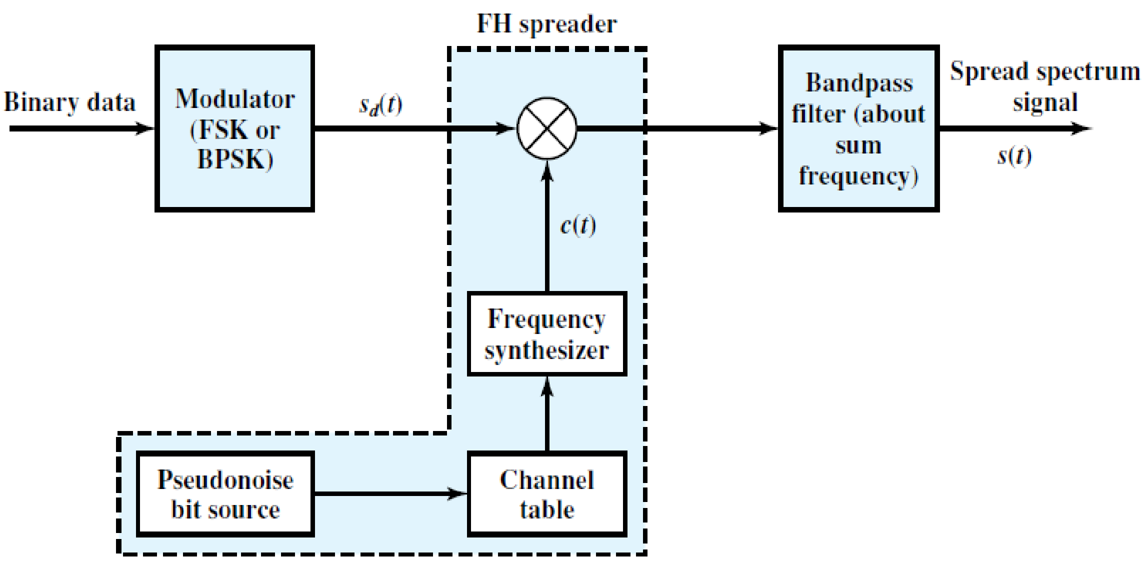


Figure 4.5 FHSS Transmitter

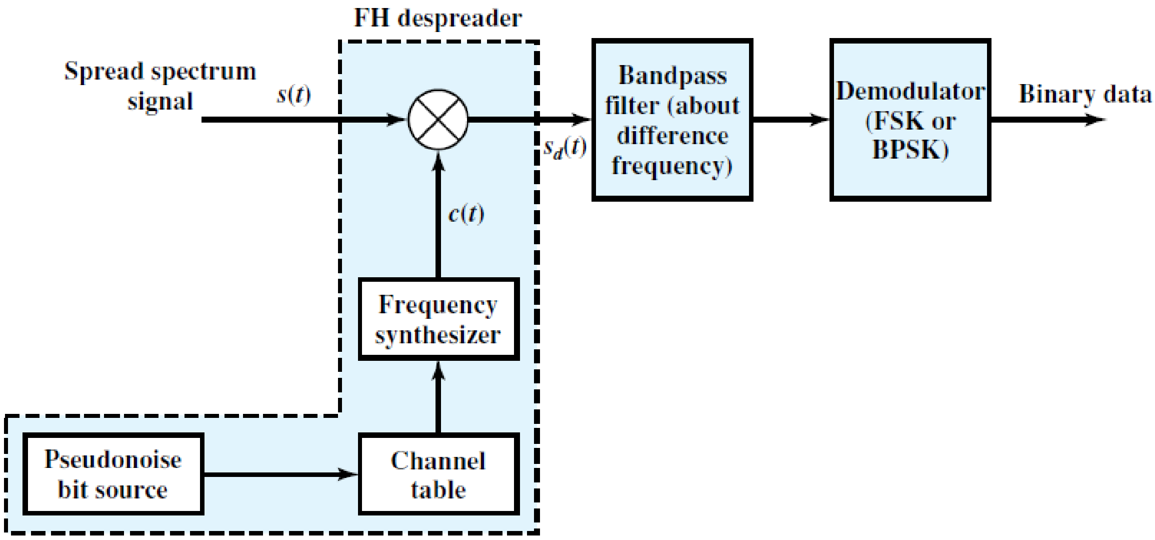


Figure 4.6 FHSS Receiver

At the transmitter, a binary sequence of bits that represents the data is digitally modulated using either MFSK or BPSK modulation

Based on the sequence that is received from the PN sequence generator, and a hopping algorithm w.r.t channel table afrequency is generated .

The signal c(t) that is generated by the frequency synthesizer is usually several times larger as compared to the modulated binary data

s(t) and c(t) are mixed together and passed through BPF for further transmission.

A similar process at the receiver generates the original binary data at the receiver.

Principle Behind Code Division Multiple Access

- In code division multiple access (CDMA) systems, the narrowband message signal is multiplied by a very large bandwidth signal called the spreading signal. The spreading signal is a pseudo-noise code sequence that has a chip rate which is orders of magnitudes greater than the data rate of the message.
- All users in a CDMA system, use the same carrier frequency and may transmit simultaneously. Each user has its own pseudorandom code word which is approximately orthogonal to all other codewords.
- The receiver performs a time correlation operation to detect only the specific desired codeword. All other code words appear as noise due to decorrelation.
- For detection of the message signal, the receiver needs to know the codeword used by the transmitter. Each user operates independently with no knowledge of the other users.
- In CDMA, the power of multiple users at a receiver determines the noise floor after decorrelation. If the power of each user within a cell is not controlled such that they do not appear equal at the base station receiver, then the near-far problem occurs.
- The near-far problem occurs when many mobile users share the same channel. In general, the strongest received mobile signal will capture the demodulator at a base station.

- In CDMA, stronger received signal levels raise the noise floor at the base station demodulators for the weaker signals, thereby decreasing the probability that weaker signals will be received.
- Power control is provided by each base station in a cellular system and assures that each mobile within the base station coverage area provides the same signal level to the base station receiver.
- This solves the problem of a nearby subscriber overpowering the base station receiver and drowning out the signals of far away subscribers.
- Power control is implemented at the base station by rapidly sampling the radio signal strength indicator (RSSI) levels of each mobile and then sending a power change command over the forward radio link.

Cellular Code-Division-Multiple-Access Systems

Many of the advantages of cellular CDMA are related to the fact that interference behaves almost like noise, especially in the uplink. This noise-like behavior is due to several reasons:

- The number of users (and therefore, of interferers) in each cell is large.
- Power control makes sure that all intracell signals arriving at the BS have approximately the same strength
- Interference from neighboring cells also comes from a large number of users.
- Spreading codes are designed in such a way that all signals in one cell have approximately the same cross-correlation with each signal in all neighboring cells.

Effects of multipath propagation on Code Division Multiple Access

Power control:

Power control is important to make sure that the desired user has a time-invariant signal strength, and that the interference from other users becomes noise-like.

For further considerations, we have to distinguish between power control for the uplink and that for the downlink:

Power control in the uplink**:** for the uplink, power control is vital for the proper operation ofCDMA. Power control is done by a closed control loop: the MS first sends with a certain power,the BS then tells the MS whether the power was too high or too low, and the MS adjusts its poweraccordingly.

The bandwidth of the control loop has to be chosen so that it can compensate forsmall-scale fading – i.e., has to be on the order of the Doppler frequency.

Due to time variations of the channel and noise in the channel estimate, there is a remaining variance in the powers arriving at the BS; this variance is typically on the order of 1.5–2.5 dB, while the dynamic range that has to be compensated is 60 dB or more. This variance leads to a reduction in the capacity of a CDMA cellular system of up to 20% compared with the case when there is ideal power control.
The open loop compensates for large-scale variations in the channel (path loss and shadowing), which are approximately the same at uplink and downlink frequencies. The closed loop is then used to compensate for small-scale variations.

***Power control in the downlink*:** for the downlink, power control is not necessary for CDMA to function: all signals from the BS arrive at one MS with the same power (the channel is the same for all signals).

The goal of downlink power control is to minimize the total transmit power while keeping the BER or SINR level above a given threshold.

The accuracy of downlink power control need not be as high as for the uplink; for many cases, open loop control is sufficient.

Soft Handover:

As all cells use the same frequencies, an MS can have contact with two BSs at the same time. If an MS is close to a cell boundary, it receives signals from two or more BSs and also transmits to all of these BSs.

Signals coming from different MSs have different delays, but this can be compensated by the Rake receiver, and signals from different cells can be added coherently.

This is in contrast to the hard handover in an FDMA-based system, where an MS can have contact with only one BS at a time, because it can communicate only on one frequency at a time. Consider now an MS that starts in cell A, but has already established a link to BS B as well.

At the outset, the MS gets the strongest signal from BS A. As it starts to move toward cell B, the signal from BS A becomes weaker, and the signal from BS B becomes stronger, until the system decides to drop the link to BS A.

Soft handover dramatically improves performance while the MS is near the borderline of the two cells, as it provides diversity (macrodiversity) that can combat large-scale as well as small-scale fading.

On the downside, soft handover decreases the available capacity in the downlink: one MS requires resources (Walsh–Hadamard codes) in two cells at the same time, while the user talks – and pays – only once. Furthermore, soft handover increases the amount of signaling that is required between BSs.

Methods for Capacity Increases:

***Quiet periods during speech transmission*:**

- For speech transmission, CDMA makes implicit useof the fact that a person does not talk continuously, but rather only about 50% of the time, theremainder of the time (s)he listens to the other participant.
- In addition, there are pauses between words and even syllables, so that the ratio of “talk time” to “total time of a call” is about 0.4. During quiet periods, no signal, or a signal with a very low data rate, has to be transmitted.

- In a CDMA system, not transmitting information leads to a decrease in total transmitted power, andthus interference in the system. But we have already seen above that decreasing the interferencepower allows additional users to place calls.

- There can be a worst case scenario where all users in a cell are talking simultaneously, but, statistically speaking, this is highly improbable, especially when the number of users is large.

- Thus, pauses in the conversation can be used very efficiently by CDMA in order to improve capacity (compare also discontinuous transmission in TDMA systems).

Flexible data rate:

- In an FDMA (TDMA) system, a user can occupy either one frequency (timeslot),or integer multiples thereof. In a CDMA system, arbitrary data rates can be transmitted by an appropriate choice of spreading sequences.
- This is not important for speech communications, which operate at a fixed data rate. For data transmission, however, the flexible data rate allows for better exploitation of the available spectrum.
-

***Soft capacity*:**

- The capacity of a CDMA system can vary from cell to cell. If a given cell adds more users, it increases interference to other cells.
- It is thus possible to have some cells with high capacity, and some with lower; furthermore, this can change dynamically, as traffic changes. This concept is known as *breathing cells*.

***Error correction coding*:**

- The drawback of error correction coding is that the data rate that is to be transmitted is increased, which decreases spectral efficiency.
- On the other hand, CDMA consciously increases the amount of data to be transmitted. It is thus possible to include error correction coding without decreasing spectral efficiency;
- in other words, different users are distinguished by different error correction codes (coding by spreading). Note, however, that commercial systems do not use this approach; they have separate error correction and spreading.

Orthogonal Frequency Division Multiplexing

Orthogonal Frequency Division Multiplexing or OFDM is a modulation format that is being used for many of the latest wireless and telecommunications standards.

OFDM has been adopted in the Wi-Fi arena where the standards like 802.11a, 802.11n, 802.11ac and more. It has also been chosen for the cellular telecommunications standard LTE / LTE-A, and in addition to this it has been adopted by other standards such as WiMAX and many more.

Orthogonal frequency division multiplexing has also been adopted for a number of broadcast standards from DAB Digital Radio to the Digital Video Broadcast standards, DVB. It has also been adopted for other broadcast systems as well including Digital Radio Mondiale used for the long medium and short wave bands.

Although OFDM, orthogonal frequency division multiplexing is more complicated than earlier forms of signal format, it provides some distinct advantages in terms of data transmission, especially where high data rates are needed along with relatively wide bandwidths.

OFDM has been adopted in the Wi-Fi arena where the standards like 802.11a, 802.11n, 802.11ac and more. It has also been chosen for the cellular telecommunications standard LTE / LTE-A, and in addition to this it has been adopted by other standards such as WiMAX and many more.

Orthogonal frequency division multiplexing has also been adopted for a number of broadcast standards from DAB Digital Radio to the Digital Video Broadcast standards, DVB. It has also been

adopted for other broadcast systems as well including Digital Radio Mondiale used for the long medium and short wave bands.

Although OFDM, orthogonal frequency division multiplexing is more complicated than earlier forms of signal format, it provides some distinct advantages in terms of data transmission, especially where high data rates are needed along with relatively wide bandwidths.Figure 4.7 and 4.8 illustrates the spectrum separation.

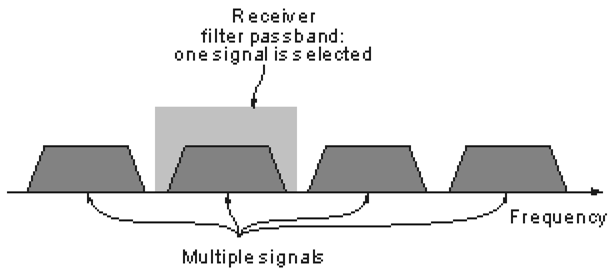


Figure 4.7 :Traditional view of receiving signals carrying modulation

To see how OFDM works, it is necessary to look at the receiver. This acts as a bank of demodulators, translating each carrier down to DC. The resulting signal is integrated over the symbol period to regenerate the data from that carrier.

The same demodulator also demodulates the other carriers. As the carrier spacing equal to the reciprocal of the symbol period means that they will have a whole number of cycles in the symbol period and their contribution will sum to zero - in other words there is no interference contribution.

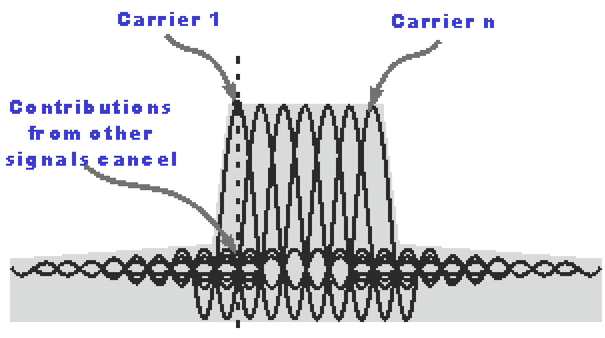


Figure 4.8:OFDM Spectrum

One requirement of the OFDM transmitting and receiving systems is that they must be linear. Any non-linearity will cause interference between the carriers as a result of inter-modulation distortion. This will introduce unwanted signals that would cause interference and impair the orthogonality of the transmission.

In terms of the equipment to be used the high peak to average ratio of multi-carrier systems such as OFDM requires the RF final amplifier on the output of the transmitter to be able to handle the peaks whilst the average power is much lower and this leads to inefficiency. In some systems the peaks are limited. Although this introduces distortion that results in a higher level of data errors, the system can rely on the error correction to remove them.

Data on OFDM

The data to be transmitted on an OFDM signal is spread across the carriers of the signal, each carrier taking part of the payload. This reduces the data rate taken by each carrier. The lower data rate has the advantage that interference from reflections is much less critical.

This is achieved by adding a guard band time or guard interval into the system. The below figure 4.9 illustrates the guard interval.This ensures that the data is only sampled when the signal is stable and no new delayed signals arrive that would alter the timing and phase of the signal.

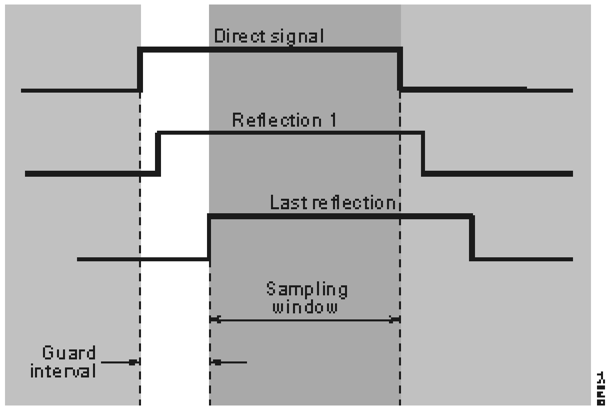


Figure 4.9:Guard Interval

The distribution of the data across a large number of carriers in the OFDM signal has some further advantages. Nulls caused by multi-path effects or interference on a given frequency only affect a small number of the carriers, the remaining ones being received correctly.

By using error-coding techniques, which does mean adding further data to the transmitted signal, it enables many or all of the corrupted data to be reconstructed within the receiver. This can be done because the error correction code is transmitted in a different part of the signal.

OFDM advantages & disadvantages

OFDM advantages

OFDM has been used in many high data rate wireless systems because of the many advantages it provides.

- ***Immunity to selective fading:*** One of the main advantages of OFDM is that is more resistant to frequency selective fading than single carrier systems because it divides the overall channel into multiple narrowband signals that are affected individually as flat fading sub-channels.
- ***Resilience to interference:*** Interference appearing on a channel may be bandwidth limited and in this way will not affect all the sub-channels. This means that not all the data is lost.
- ***Spectrum efficiency:*** Using close-spaced overlapping sub-carriers, a significant OFDM advantage is that it makes efficient use of the available spectrum.
- ***Resilient to ISI:*** Another advantage of OFDM is that it is very resilient to inter-symbol and inter-frame interference. This results from the low data rate on each of the sub-channels.
- ***Resilient to narrow-band effects:*** Using adequate channel coding and interleaving it is possible to recover symbols lost due to the frequency selectivity of the channel and narrow band interference. Not all the data is lost.
- ***Simpler channel equalisation:*** One of the issues with CDMA systems was the complexity of the channel equalisation which had to be applied across the whole channel. An advantage

of OFDM is that using multiple sub-channels, the channel equalization becomes much simpler.

OFDM Disadvantages

Whilst OFDM has been widely used, there are still a few disadvantages to its use which need to be addressed when considering its use.

- ***High peak to average power ratio:*** An OFDM signal has a noise like amplitude variation and has a relatively high large dynamic range, or peak to average power ratio. This impacts the RF amplifier efficiency as the amplifiers need to be linear and accommodate the large amplitude variations and these factors mean the amplifier cannot operate with a high efficiency level.
- ***Sensitive to carrier offset and drift:*** Another disadvantage of OFDM is that is sensitive to carrier frequency offset and drift. Single carrier systems are less sensitive.

OFDM Variants

There are several other variants of OFDM for which the initials are seen in the technical literature. These follow the basic format for OFDM, but have additional attributes or variations:

- ***COFDM:*** Coded Orthogonal frequency division multiplexing. A form of OFDM where error correction coding is incorporated into the signal.
- ***Flash OFDM:*** This is a variant of OFDM that was developed by Flarion and it is a fast hopped form of OFDM. It uses multiple tones and fast hopping to spread signals over a given spectrum band.
- ***OFDMA:*** Orthogonal frequency division multiple access. A scheme used to provide a multiple access capability for applications such as cellular telecommunications when using OFDM technologies.
- ***VOFDM:*** Vector OFDM. This form of OFDM uses the concept of MIMO technology. It is being developed by CISCO Systems. MIMO stands for Multiple Input Multiple output and it uses multiple antennas to transmit and receive the signals so that multi-path effects can be utilised to enhance the signal reception and improve the transmission speeds that can be supported.
- ***WOFDM:*** Wideband OFDM. The concept of this form of OFDM is that it uses a degree of spacing between the channels that is large enough that any frequency errors between transmitter and receiver do not affect the performance. It is particularly applicable to Wi-Fi systems.

Each of these forms of OFDM utilize the same basic concept of using close spaced orthogonal carriers each carrying low data rate signals. During the demodulation phase the data is then combined to provide the complete signal.

Second Generation (2G*):* 2G (or 2-G) is short for second-generation wireless telephone technology. Second generation 2G cellular telecom networks were commercially launched on the GSM standard in Finland by Radiolinja(now part of Elisa Oyj) in 1991.

- Three primary benefits of 2G networks over their predecessors were that phone conversations were digitally encrypted, 2G systems were significantly more efficient on the spectrum allowing for far greater mobile phone penetration levels; and 2G introduced data services for mobile, starting with SMS text messages.

- 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.

TDMA/FDD Standards

- 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.

- 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.

- 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.

Advanced Second Generation(2.5G)*:2.5G is a stepping stone between 2G and 3G cellular wireless technologies. The term “second and a half generation” is used to describe 2G-systems that have implemented a packet switched domain in addition to the circuit switched domain. It does not necessarily provide faster services because bundling of timeslots is used for circuit switched data services (HSCSD) as well.*

- The first major step in the evolution of GSM networks to 3G occurred with the introduction of General Packet Radio Service (GPRS). CDMA2000 networks similarly evolved through the introduction of 1xRTT. So the cellular services combined with enhanced data transmission capabilities became known as ’2.5G.’

  Main upgradation techniques are:
  - supporting higher data rate transmission for web browsing.
  - supporting e-mail traffic.
  - 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.

Third Generation ***(3G)***: International 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 telephone, video calls, and wireless data, all in a mobile environment.
- Compared to 2G and 2.5G services, 3G allows simultaneous use of speech and data services and higher data rates (at least 200 kbit/s peak bit rate to fulfill to IMT-2000 specification). Today's 3G systems can in practice offer up to 14.0 Mbit/s (1.75 MB/s) on the downlink and 5.8 Mbit/s (0.725 MB/s) on the up link
- 3G networks are wide area cellular telephone networks which evolved to incorporate high-speed internet access and video telephony

Some features of 3G networks are listed below

- 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.

3G STANDARDS

- W-CDMA
- CDMA2000
- TD-SCDMA

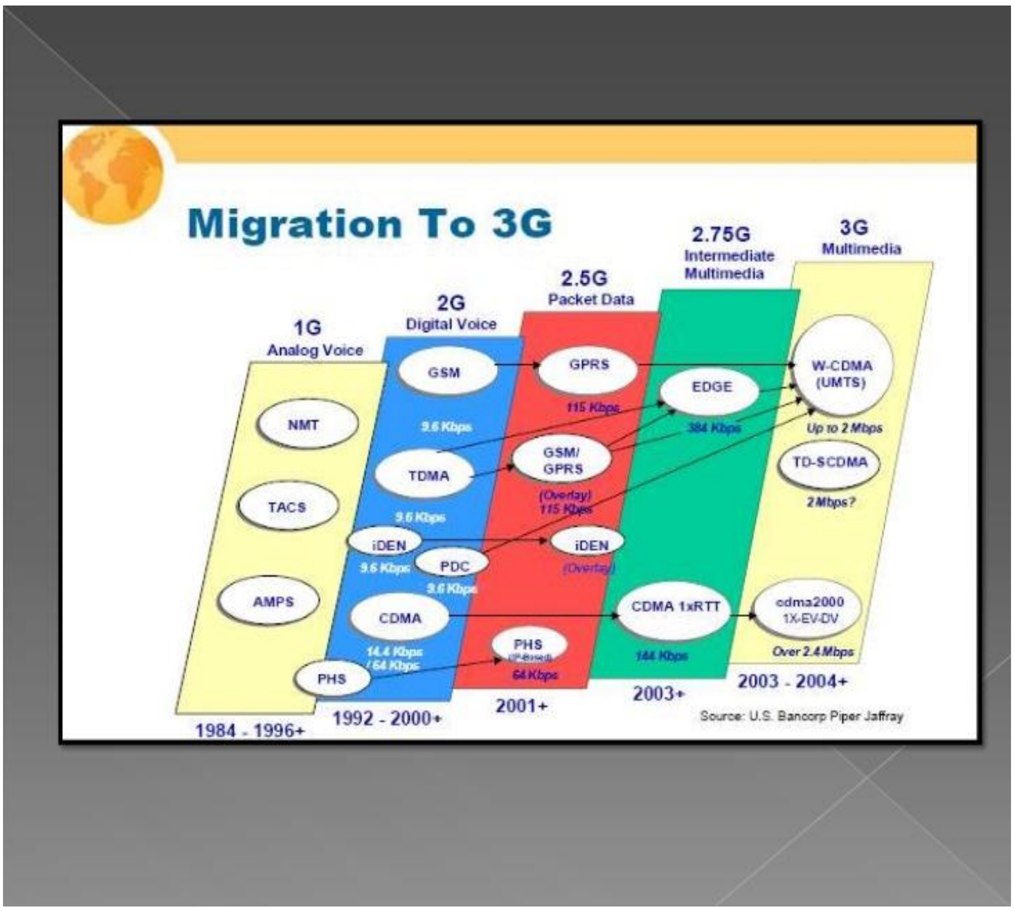
- **3G W-CDMA** (UMTS)
  - WCDMA is based on DS-CDMA (direct sequence code division multiple access) technology in which user-information bits are spread over a wide bandwidth (much larger than the information signal bandwidth) by multiplying the user data with the spreading code.

- **3G CDMA2000**
  - Code division multiple access 2000 is the natural evolution of IS-95 (cdmaOne). It includes additional functionality that increases its spectral e_ciency and data rate capability.(code division multiple access) is a mobile digital radio technology where channels are de_ned with codes (PN sequences).

  - CDMA permits many simultaneous transmitters on the same frequency channel.

- **3G TD-SCDMA**
  - TD-SCDMA uses TDD, in contrast to the FDD scheme used by W-CDMA. By dynamically adjusting the number of timeslots used for downlink and uplink, the system can more easily accommodate asymmetric traffic with different data rate requirements on downlink and uplink than FDD schemes.

  - The "S" in TD-SCDMA stands for "synchronous", which means that uplink signals are synchronized at the base station receiver, achieved by continuous timing adjustments. This reduces the interference between users of the same timeslot using different codes by improving the orthogonality between the codes, therefore increasing system capacity, at the cost of some hardware complexity in achieving uplink synchronization.



A GLIMPSE OF CELLULAR STRUCTURE

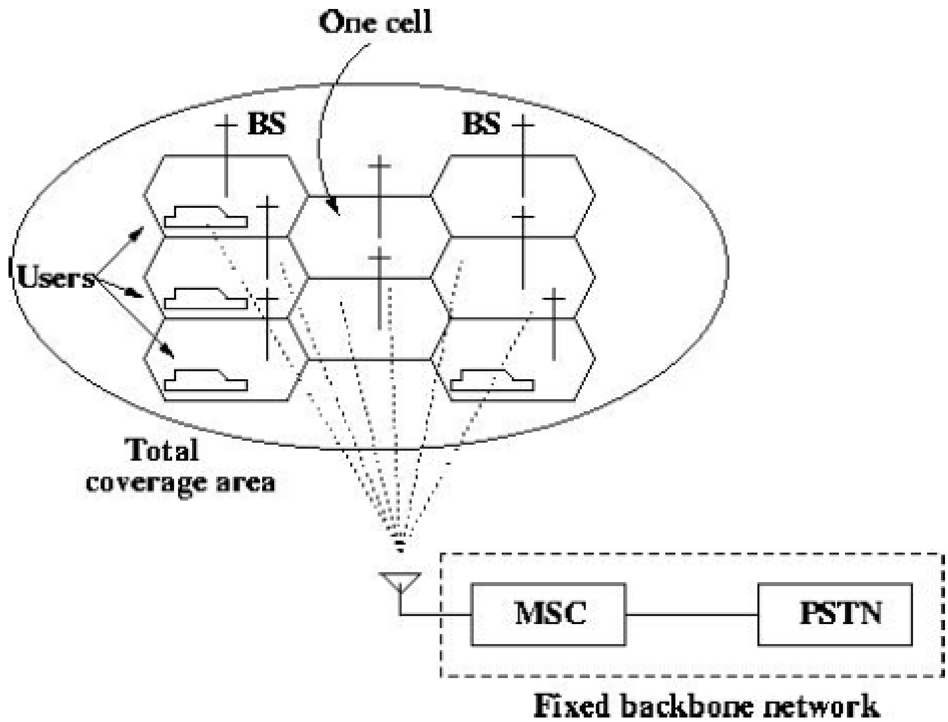


Figure 4.10 :Basic cellular structure

- Cellular telephone systems must accommodate a large number of users over a large geographic area with limited frequency spectrum, i.e., with limited number of channels. If a single transmitter/ receiver is used with only a single base station, then sufficient amount of power may not be present at a huge distance from the BS.
- Figure 4.10 illustrates the basicBasic cellular structure.
- For a large geographic coverage area, a high powered transmitter therefore has to be used. But a high power radio transmitter causes harm to environment.
- Mobile communication replaces the high power transmitters by low power transmitters by dividing the coverage area into small segments, called cells.
- Each cell uses a certain number of the available channels and a group of adjacent cells together use all the available channels. Such a group is called a cluster.
- This cluster can repeat itself and hence the same set of channels can be used again and again. Each cell has a low power transmitter with a coverage area equal to the area of thecell.
- This technique of substituting a single high powered transmitter by several low powered transmitters to support many users is the backbone of the cellular concept.

Operational Channels in cellular system

In each cell, there are four types of channels that take active part during a mobilecall. These are:

- **Forward Voice Channel (FVC):** This channel is used for the voice transmission from the BS to the MS.

- **Reverse Voice Channel (RVC**): This is used for the voice transmissionfrom the MS to the BS.

- **Forward Control Channel (FCC**): Control channels are generally usedfor controlling the activity of the call, i.e., they are used for setting up calls and to divert the call to unused voice channels. Hence these are also called setup channels. These channels transmit and receive call initiation and service request messages. The FCC is used for control signaling purpose from the BS to MS.

- **Reverse Control Channel (RCC):** This is used for the call control purpose from the MS to the BS. Control channels are usually monitored by mobiles.

Making a mobile call

- When a mobile is idle, i.e., it is not experiencing the process of a call, then it searches all the FCCs to determine the one with the highest signal strength.

- The mobile then monitors this particular FCC. However, when the signal strength falls below a particular threshold that is insufficient for a call to take place, the mobile again searches all the FCCs for the one with the highest signal strength.

- For a particular country or continent, the control channels will be the same. So all mobiles in that country or continent will search among the same set of control channels.

- However, when a mobile moves to a different country or continent, then the control channels for that particular location will be different and hence the mobile will not work.

- Each mobile has a mobile identification number (MIN). When a user wants to make a call, he sends a call request to the MSC on the reverse control channel. He also sends the MIN of the person to whom the call has to be made.

- The MSC then sends this MIN to all the base stations. The base station transmits this MIN and all the mobiles within the coverage area of that base station receive the MIN and match it with their own.

- If the MIN matches with a particular MS, that mobile sends an acknowledgment to the BS. The BS then informs the MSC that the mobile is within its coverage area. The MSC then instructs the base station to access specific unused voice channel pair.

- The base station then sends a message to the mobile to move to the particular channels and it also sends a signal to the mobile for ringing. In order to maintain the quality of the call, the MSC adjusts the transmitted power of the mobile which is usually expressed in dB or dBm.

- When a mobile moves from the coverage area of one base station to the coverage area of another base station i.e., from one cell to another cell, then the signal strength of the initial base station may not be sufficient to continue the call in progress.

- So the call has to be transferred to the other base station. This is called handoff In such cases, in order to maintain the call, the MSC transfers the call to one of the unused voice channels of the new base station or it transfers the control of the current voice channels to the new base station.

Block diagram of a IS-95 Transmitter.

The generalized block diagram for IS-95 Transmitter is shown in figure 4.7.we have considered the case when the output of the channel coder actually has a data rate of 28.8 kbit/s – i.e., a source rate of 14.4 or 9.6 kbit/s. However, depending on the source, data, a lower rate (14.4 kbit/s, 7.2 kbit/s, or 3.6 kbit/s) can also be the output of a convolutional encoder.

In this case, encoded symbols are repeated (several times, if necessary) until a data rate of 28.8 kbit/s is achieved. It is these repeated data that are sent to the block interleaver for further processing. However, it would waste resources to transmit all of these repeated data at full power.

For the uplink, this problem is solved by gating off the transmitter part of the time. If, e.g., the coded data rate is 14.4 kbit/s (source data rate 7.2 kbit/s), then the transmitter is turned on only 1/2 of the time.

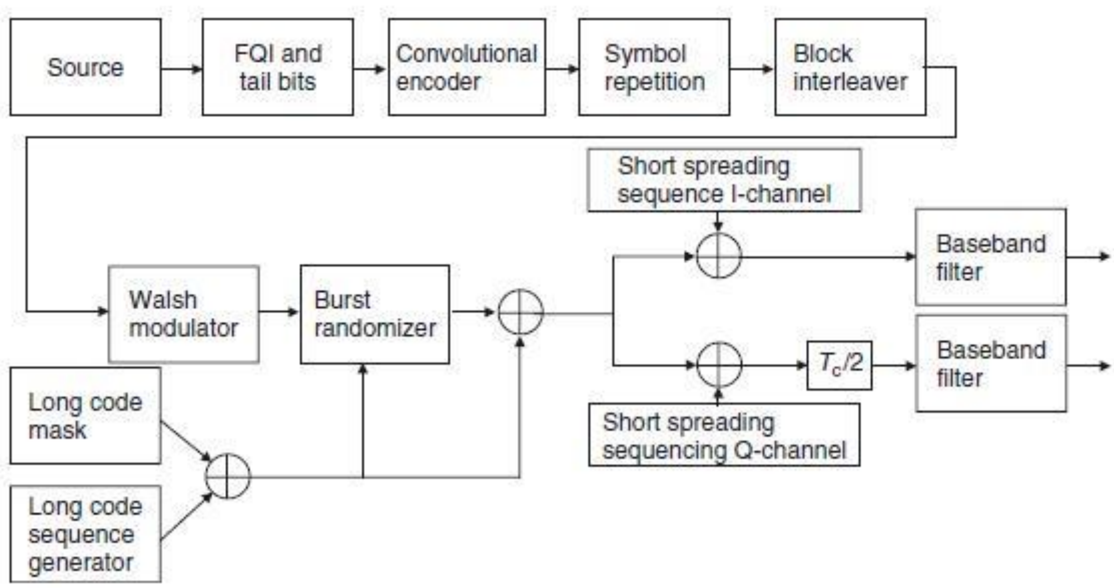


Figure 4.7: Block diagram of an IS-95 mobile station transmitter

As a consequence, average transmit power is only 1/2 of the full-data-rate case, and the interference seen by other users is only half as large. Determination of the time for gating off the transmitter is actually quite complicated. The first issue is that gating has to be coordinated with the interleaver. For example, for the 7.2-kbit source data rate mode, each 1.25-ms-long group of output symbols is repeated once. Gating thus eliminates one of these two symbol groups.

UNIT V

WIRELESS DATA SERVICES: First Wave of Mobile Data Services: Text-Based Instant Messaging. Second Wave of Mobile Data Services: Low-Speed Mobile Internet Services. Current Wave of Mobile Data Services: High-Speed and Multimedia Mobile Internet Services. IP-Based Wireless Networks - 3GPP, 3GPP2.

First Wave of Mobile Data Services: Text-Based Instant Messaging

The first globally successful mobile data service is SMS (Short Message Services), which was first introduced in Europe over GSM networks. SMS allows a mobile user to send and receive short text messages (up to 160 text characters) instantly.

Supporting SMS does not require a packet core network. Instead, SMS messages are delivered using the signaling protocol-Mobile Application Part (MAP)—that was originally designed to support mobility in GSM networks.

This allowed SMSservices to be provided over the completely circuit-switched 2G GSM networks long before packet core networks were introduced into wireless networks.

SMS services grew rapidly first in Europe. Today, SMS services are booming throughout the world.

In the United Kingdom, the number of transmitted SMS messages more than doubled in the two-year period from 2001 to 2002. Based on statistics from theMobile Data Association, an average of 52 million SMS messages weretransmitted every day in the United Kingdom in December 2002, whichtranslates into about 2.2 million messages per hour on the average.

Figure 5.1shows the SMS subscriber growth in the United Kingdom from 1998 to June2003.

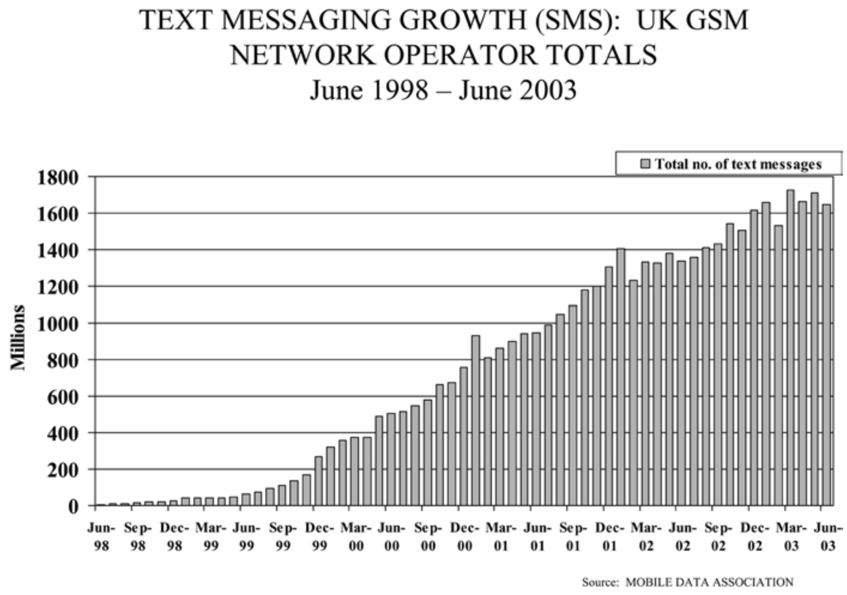


Figure 5.1 Growth of SMS message transmissions in the United Kingdom

In Europe, 186 billion SMS messages were transmitted in 2002.. In China, the revenues of SMS and value-added services (VAS) totaled 750 million in 2002. During an eight-day period around the Chinese New Yearin 2003 (early February), approximately 7 billion SMS messages were transmitted.

In the United States, SMS is experiencing an explosive growth, although its usage levels have not reached the levels in Western Europe and the AsiaPacific region. InphoMatch, a company that handles intercarrier SMS services for AT&T Wireless, Verizon, and T-Mobile, delivered 100 million SMS messages in August 2002 and saw the intercarrier SMS traffic doubling every two months.

In addition to providing a highly valuable data service to mobile users, SMS allowed mobile users to become familiar and comfortable with mobile data services and to appreciate the value of mobile data services. This helped pave the road for mobile users to adopt the more advanced data services that arrived next.

Second Wave of Mobile Data Services: Low-Speed Mobile Internet Services

Interactive and information-based mobile Internet services emerged as the secondwave of widespread mobile data services. An example of successful mobile Internetapplications is i-Mode, which was launched by NTT DoCoMo over its PDC radio
systems in Japan in February 1999.

The i-Mode services include:

Sending and receiving emails and instant messages.Commercial transactions, e.g., banking, ticket reservation, credit card billing inquiry, and stock tradingDirectory services, e.g., dictionary, restaurant guides, and phone directory.

Daily information, e.g., news, weather reports, road conditions, and trafficinformation.Entertainment, e.g., Karaoke, network games, and horoscope.

The i-Mode services has experienced a rapid and steady growth ever since its launch. Figure 1.7 shows the growth of i-Mode subscribers for the three-year period from 2001 to 2003 (Source: NTT DoCoMo). By the summer of 2003, i-Mode wassupporting over 38 million subscribers.

i-Mode represents a significant milestone in the evolution of mobile services. It was the first major success in bringing Internet-based services to a large population of mobile subscribers. It demonstrated the values and the potentials of the mobileInternet to the world.

Today, however, the i-Mode services are suffering from two major limitations:

i-Mode services are limited by the low data rate of the PDC radio networks.

i-Mode users rely on proprietary protocols developed by NTT DoCoMo, ratherthan on standard IP-based protocols, to access i-Mode services. The i-Modeservices are provided by World-Wide Web (WWW) sites specifically designedfor mobile users. Mobile devices use a set of proprietary protocols developed by NTT DoCoMo to communicate with these WWW sites via a gateway.

Thegateway converts between the protocols over the radio access network and the protocols used by the WWW sites. The proprietary protocols make it difficult for i-Mode to be adopted by other countries.

Today, 2.5G and the early forms of 3G wireless networks allow mobile users toaccess the Internet via standard IP-based protocols. This is enabling mobile Internetservices similar to those provided by i-Mode to be provided to users around theworld and over radio channels of higher data rates than 2G radio systems.

Current Wave of Mobile Data Services: High-Speed and Multimedia Mobile Internet Services

The higher system capacities and data rates provided by 2.5G and 3G wireless networks plus the closer integration of 2.5G and 3G wireless networks with the Internet enabled by the IP technologies used in these wireless networks are enabling many new ways for people to communicate.

People are not only talking over mobile phones and using instant text-based messages. They can now use their mobile phones to take pictures, record videos, and send the pictures and videos to other people or devices; receive location-dependent services; and play sophisticated games in real time with remote users.

Examples of advanced mobile data and multimedia applications include the following:

- **Camera phones**: Mobile phones with integrated cameras that allow a user to take still pictures, record short videos with sound, and send the photos and videos as multimedia messages or email to other users.

- **Multimedia Messaging Services (MMS):** Send and receive messages with multimedia contents (data, voice, still pictures, videos, etc.).

- **Networked gaming**: People may download games to their mobile handsets and play the games locally. They may also use their mobile handsets to play games with remote users in real time.

- **Location-based services**: People may use their mobile devices to receive real time navigation services, local maps, and information on local points of interest (e.g., restaurants, tourist locations, cinemas, gas stations, shopping malls, hospitals, and vehicle repair shops).

- **Streaming videos to mobile devices**: People may use their mobile devices to view real-time and non-real-time videos, for example, short videos received from friends' camera phones, watch TV.

- **Vehicle information systems**: People on moving vehicles (e.g., cars, trains,boats, airplanes) may access the Internet or their enterprise networks the same way as when they are at their offices or homes.

- They may be able to surf the Internet, access their corporate networks, download games from the network, play games with remote users, obtain tour guidance information, obtain realtime traffic and route conditions information, etc.

- Many of the services described above are already available over 2.5G or early forms of 3G wireless networks and are changing the ways people communicate in a fundamental manner.

- Camera phones, for example, enabled the first multimedia mobile applications that are truly useful to and accepted by large populations of mobile users worldwide.

- The world's first commercial camera phones that allow users to take pictures and videos and allow them to send the pictures and videos over wireless networks to other users emerged in late 2001. Ever since their introduction, camera phones have been experiencing a spectacular growth.

- According to Strategy Analytics (January 2003), 10 million camera phones were sold worldwide in the first nine months of 2002. The number of camera phones sold in 2002 was estimated to be 16 million,which is approximately the same level of worldwide sales of PDAs in 2002.

- Within merely around a year of camera phones' existence, their annual worldwide sales in 2002 were already close to the approximately 22 million digital cameras sold in the same year.

- Today, camera phones are improving rapidly. For example, picture display screens are larger, picture resolutions are higher, and user applications for handling pictures and videos are richer (e.g., applications for editing pictures and videos, taking snapshot pictures from videos, and sending pictures and videos as instant messages or email attachments over wireless networks to other users).

- Although the functionalities of camera phones are improving, their prices are declining due to intensive competition as all major mobile phone manufacturers around the world are competing in the camera phone market now.

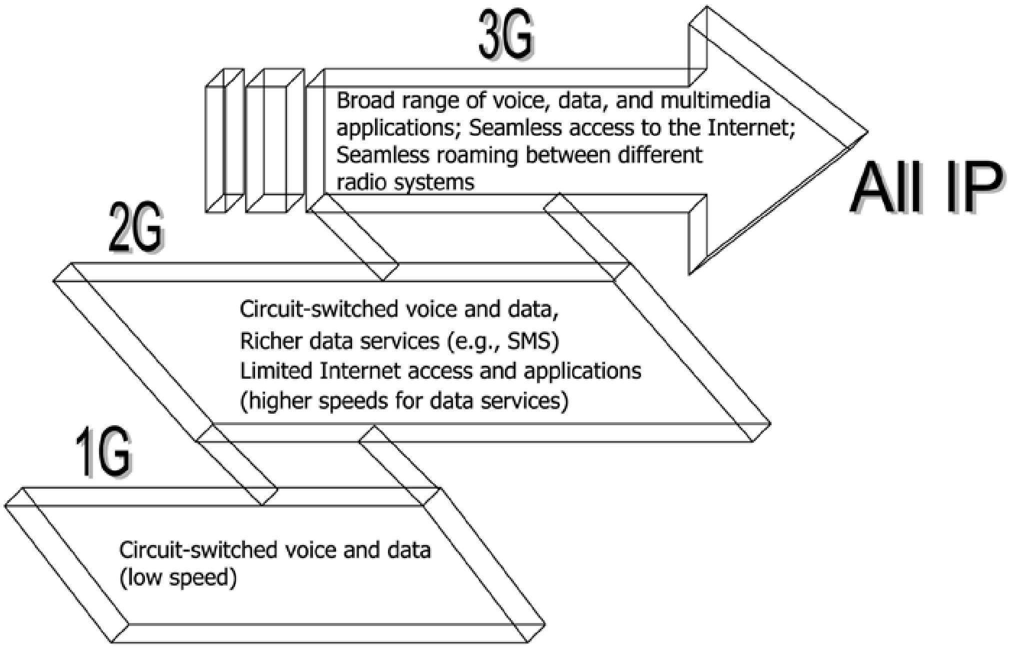


Figure 5.2.Evolution of mobile services

- The evolution of mobile services is illustrated in Figure 5.2.

- For example, camera phones priced at around US $100 became available in late 2002, whereas the average camera phone prices were around US $300 in early 2002.

- The improving functionalities and the falling prices of the camera phones, with the higher data rates of 2.5G and 3G wireless networks, make camera phones more useful and affordable to the consumers and will therefore further accelerate the growth of camera phones.

- Strategy Analytics expects that camera phones will outsell digital cameras by 2004.

- The widespread use of camera phones is not only changing the ways people communicate, but also it is changing the mix of mobile services and the nature of the network traffic generated by mobile services. In particular, camera phones help increase multimedia traffic over the wireless networks significantly.

IP-BASED WIRELESS NETWORKS:

Wireless networks are evolving into IP-based mobile networks.

IP-based wireless networks are better suited for supporting the rapidly growing mobile data and multimedia services. As mobile data and multimedia services continue to grow more rapidly than mobile voice services, they will overtake mobile voice services to become the dominant mobile services in the near future.

As mobile data services become increasingly important to consumers, the revenues generated by network operators from mobile data services will also surpass the revenues from mobile voice services.

Figure 5.3 illustrates the estimated and the forecast (by Analysis Research Limited) Average Revenue Per User (ARPU) for mobile voice and non-voice services in Western Europe, 2000–2007. It shows that non-voice mobile services will grow significantly and steadily over the next few years to account for over 35% of the total revenue from mobile services in 2007.

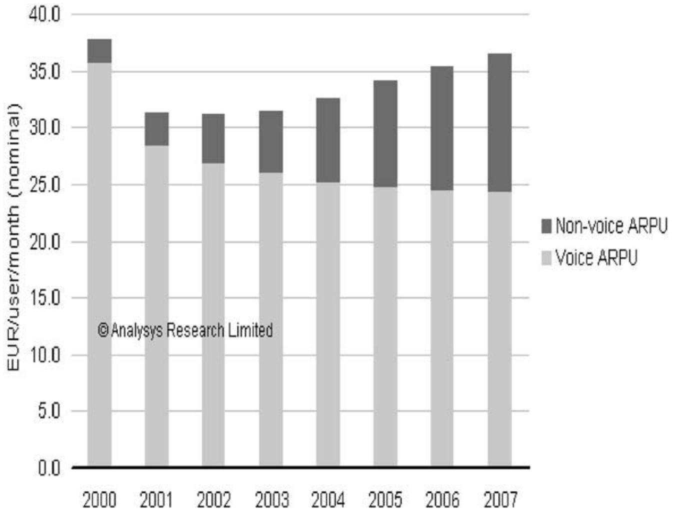


Figure 5.3 Growth of mobile voice and non-voice services

In the US, the penetration of mobile Internet services is expected to be even higher than in Western Europe, suggesting that non-voice mobile services could account for an even higherpercentage of mobile operators' total revenue.

IP-based wireless networks bring the successful Internet service paradigm to mobile providers and users. A key reason for the success of the Internet is that the IP-based Internet paradigm enables everyone in the world to create and offer services over the Internet anytime and anywhere, as long as they have a computer connected to the Internet.

An IP-based wireless network would bring the service innovation potentials of the Internet paradigm to future wireless networks.

IP-based wireless networks can integrate seamlessly with the Internet. Radio systems need to be connected to the Internet to allow mobile users to access the information, applications, and services available over the Internet.

Connecting an IP-based wireless network to the Internet is easier and more cost-effective than connecting a circuit-switched wireless network to the Internet.

Many mobile network operators also operate wireline networks. They have already built out IP core networks to support wireline IP services or as a backbone network for transporting circuit-switched voice traffic.

Mobile network operators could leverage their existing IP core networks to support radio access networks and provide services to mobile users.

IP-based radio access systems are becoming important components of public wireless networks.

IP technologies provide a better solution for making different radio technologies transparently to users.

Different radio technologies will continue to coexist in public wireless networks. These radio tech., include not only different wide-area radio technologies but also the fast growing IP-based public WLANs.

One radio technology (e.g., public WLANs) may meet communications needs other radio technologies (e.g., cellular radio systems) may not be able to meet easily. Therefore, heterogeneous radio systems are expected to coexist in the long run.

IP-based network services and applications could be provided to all users in a seamless manner, regardless of which specific radio systems or mobile devices (e.g., PDAs, laptops, phones, or any other special-purpose devices) they are using.

3GPP:

The Third Generation Partnership Project (3GPP™) was established in 1998 to develop specifications for advanced mobile communications. The 3GPP organizational partners are listed in table 5.1 below.

Table 5.1-3GPP Organizational Partners

Organization	***Base region***
Association of Radio Industries and Businesses (ARIB)	Japan
Alliance for Telecommunications Industry Solutions (ATIS)	USA
China Communications Standards Association (CCSA)	China
EuropeanTelecommunications Standards Institute (ETSI)	Europe
Telecommunications Technology Association (TTA)	Korea
Telecommunication Technology Committee (TTC)	Japan

It comprises:

- seven regional Standards Development Organizations (SDOs) including ETSI
- market associations
- several hundred companies

The original scope of 3GPP was to produce globally applicable reports and specifications for a third generation mobile system based on evolved Global System for Mobile communication (GSM™) core networks and the radio access technologies that they support.

Today the project provides complete system specifications for cellular telecommunications network technologies. It covers:

- radio access
- the core transport network
- service capabilities
- codecs
- security
- quality of service

The specifications also provide hooks for non-radio access to the core network, and for interworking with Wi-Fi networks.

1. TSG CN (Core Network): TSG CN is responsible for the specifications of the core network part of 3GPP systems, which is based on GSM and GPRS core networks. More specifically, TSG CN is responsible primarily for specifications of the layer-3 radio protocols (Call Control, Session Management, Mobility Management) between the user equipment and the core network, signaling between the core network nodes, interconnection with external networks.

Core network aspects of the interface between a radio access network and the core network, management of the core network, and matters related to supporting packet services (e.g., mapping of QoS).

2. TSG GERAN (GSM EDGE Radio Access Network): TSG GERAN is responsible for the specification of the radio access part of GSM/EDGE. This includes the RF layer; layer 1, 2, and 3 for the GERAN

Interfaces internal to the GERAN, interfaces between a GERAN and the core network, conformance test specifications for all aspects of GERAN base stations and terminals, and GERAN-specific network management specifications for the nodes in the GERAN.

3. TSG RAN (Radio Access Network): TSG RAN is responsible for the definition of the functions, requirements, and interfaces of the UTRAN. This includes radio performance; layer 1, 2, and 3 specifications in UTRAN;

specifications of the UTRAN internal interfaces and the interface between UTRAN and core networks; definition of the network management requirements in UTRAN and conformance testing for base stations.

4. TSG SA (Service and System Aspects): TSG SA is responsible for the overall architecture and service capabilities of systems based on 3GPP specifications.

This includes the definition and maintenance of the overall system architecture, definition of required bearers and services, development of service capabilities and a service architecture, as well as charging, security, and network management aspects of 3GPP system.

5. TSG T (Terminal): TSG T is responsible for specifying terminal interfaces (logical and physical), terminal capabilities (such as execution environments), and terminal performance/testing. 3GPP specifications produced in different time periods are published as Releases.

Each Release contains a set of Technical Specifications and Technical Reports. A Release is said to be frozen at a specific date if its content can only be revised in case a correction is needed after that date. Initially, 3GPP planned to standardize a new release each year.

The first release therefore is named as Release 99 (frozen in March 2000). Release 99 (R99 in short) mainly focuses on a new RAN based on WCDMA. It also emphasizes the interworking and backward compatibility with GSM. Due to a variety of modifications proposed, Release 00 (R00) was scheduled into two different releases, which are named as Release 4 (R4) and Release 5 (R5).

Release 4,frozen in March 2001, is a minor release with some enhancements to R99. IPtransport was also introduced into the core network. Release 5 was frozen in June2002. It comprises major changes in the core network based on IP protocols. Morespecifically, phase 1 of the IP Multimedia Subsystem (IMS) was defined.

In addition,IP transport in the UNTRAN was specified. Release 6 is expected to be frozen inMarch 2004. It will focus on IMS phase 2, harmonization of the IMS in 3GPP and3GPP2, interoperability of UMTS and WLAN, and multimedia broadcast andmulticast.

3GPP2

The 3GPP2, like 3GPP, is also an international collaboration to produce globalstandards for third-generation wireless networks. 3GPP2 was formed soon after3GPP when the American National Standards Institute (ANSI) failed to convince3GPP to include “non-GSM” technologies in 3G standards. 3GPP2 members arealso classified into Organizational Partners and Market Representation Partners.Today, 3GPP2 has five Organizational Partners: ARIB (Japan), CWTS (China), TIA (Telecommunications Industry Association) in North America, TTA (Korea), andTTC (Japan).

Standards produced by 3GPP2 are published as 3GPP2 Technical Specifications.Technical Working Groups (TSGs) are responsible for producing TechnicalSpecifications. A Steering Committee coordinates the works among different TSGs.Currently, 3GPP2 has the following TSGs:

TSG-A (Access Network Interfaces): TSG-A is responsible for the specifications of interfaces between the radio access network and core network, as well as within the access network. Specifically, it has a responsibility for the specifications of the following aspects of radio access network interfaces:

physical links, transports and signaling, support for access network mobility, 3G capability (e.g., high-speed data support), interfaces inside the radio access network, and interoperability specification.

TSG-C (cdma2000): TSG-C is responsible for the radio access part, including its internal structure, of systems based on 3GPP2 specifications. Specifically, it has a responsibility for the requirements, functions, and interfaces for the cdma2000 radio infrastructure and user terminal equipment.

These include specifications of radio layers 1 –3, radio link protocol, support for enhanced privacy, authentication and encryption, digital speech codecs, video codec selection and specification of related video services, data and other ancillary services support, conformance test plans, and location-based services support.

TSG-S (Service and System Aspects): TSG-S is responsible for the development of service capability requirements for systems based on 3GPP2 specifications.

It is also responsible for high-level architectural issues, as required to coordinate service development across the various TSGs.

Some specific responsibilities include

- Definition of services, network management, and system requirements.
- Development and maintenance of network architecture and associated system requirements and reference models.
- Management, technical coordination, as well as architectural and requirements development associated with all end-to-end features, services, and system capabilities, including, but not limited to, security and QoS.
- Requirements for international roaming.. TSG-X (Intersystem Operations): TSG-X is responsible for the specifications of the core network part of systems, based on 3GPP2 specifications. Specifically, it has a responsibility for:
- Core network internal interfaces for call associated and noncall associated signaling.
- IP technology to support wireless packet data services, including voice and other multimedia services.
- Core network internal interfaces for bearer transport.
- Charging, accounting, and billing specifications.
- Validation and verification of specification text it develops.
- Evolution of core network to support interoperability and intersystem operations, and international roaming.
- Network support for enhanced privacy, authentication, data integrity, and other security aspects.
- Wireless IP services.

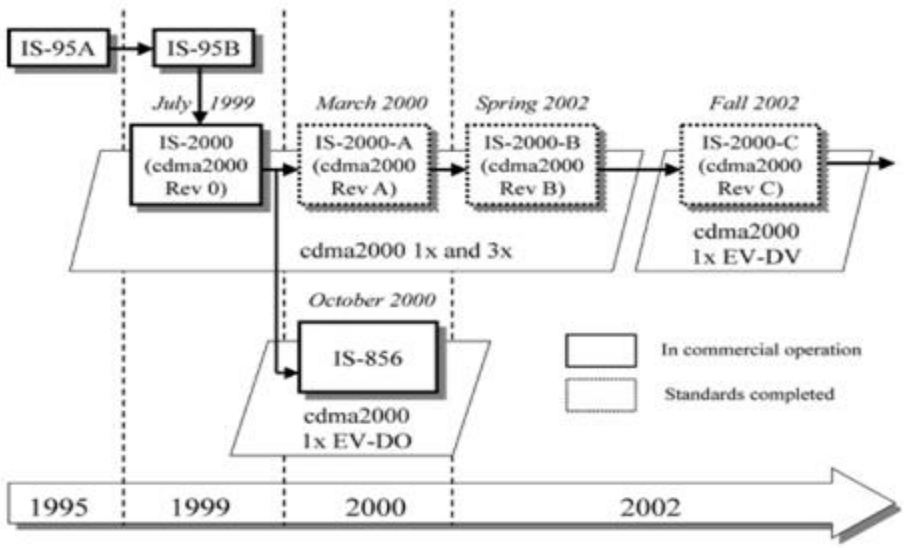


Figure 5.4 CDMA2000 family

Although 3GPP2 specifies standards for both core network and radio accessnetwork, revisions of 3GPP2 specifications are primary based on the cdma2000radio access network. As shown in Figure 5.4, there are three revisions incdma2000 1x and 3x.

They are specified by 3GPP2 C.S001-0005 Revision 0 [2, 3, 4,5, 6], C.S001-0005 Revision A [1, 10, 13, 16, 19], and C.S001-0005 Revision B [7,11, 14, 17, 20]. The specifications are based on the TIA IS-2000 series [35]. Thereare two evolutions (EV) of cdma2000 1x.

The cdma2000 1x EV-DO, specified byIS-856 [34]/3GPP2 C.S0024 [9], defined the enhancement of cdma2000 1x for dataonly (DO). It is based on the HDR developed by QUALCOMM for direct Internetaccess.

The specifications of 3GPP2 C.S001-0005 Revision C [8, 12, 15, 18, 21]specify cdma2000 1x EV-DV, the evolution of cdma2000 1x for both data and voice(DV) enhancement. In addition to conventional circuit-switching network, packetswitching network based on IP is also incorporated.