COURSE MATERIAL
(LECTURE NOTES)
On
Mobile Computing

UNIT 1

Introduction to Mobile Computing

The rapidly expanding technology of cellular communication, wireless LANs, and satellite services will make information accessible anywhere and at any time. Regardless of size, most mobile computers will be equipped with a wireless connection to the fixed part of the network, and, perhaps, to other mobile computers. The resulting computing environment, which is often referred to as ***mobile or nomadic computing***, no longer requires users to maintain a fixed and universally known position in the network and enables almost unrestricted mobility. Mobility and portability will create an entire new class of applications and, possibly, new massive markets combining personal computing and consumer electronics.

Mobile Computing is an umbrella term used to describe technologies that enable people to access network services anyplace, anytime, and anywhere.

A communication device can exhibit any one of the following characteristics:

1. Fixed **and wired**: This configuration describes the typical desktop computer in an office. Neither weight nor power consumption of the devices allow for mobile usage. The devices use fixed networks for performance reasons.

2. Mobile **and wired**: Many of today's laptops fall into this category; users carry the laptop from one hotel to the next, reconnecting to the company's network via the telephone network and a modem.

3. Fixed **and wireless**: This mode is used for installing networks, e.g., in historical buildings to avoid damage by installing wires, or at trade shows to ensure fast network setup.

4. Mobile **and wireless**: This is the most interesting case. No cable restricts the user, who can roam between different wireless networks. Most technologies discussed in this book deal with this type of device and the networks supporting them. Today's most successful example for this category is GSM with more than 800 million users.

APPLICATIONS OF MOBILE COMPUTING

In many fields of work, the ability to keep on the move is vital in order to utilise time efficiently. The importance of Mobile Computers has been highlighted in many fields of which a few are described below:

1. Vehicles: Music, news, road conditions, weather reports, and other broadcast information are received via digital audio broadcasting (DAB) with 1.5 Mbit/s. For personal communication, a universal mobile telecommunications system (UMTS) phone might be available offering voice and data connectivity with 384 kbit/s. The current position of the car is determined via the global positioning system (GPS). Cars driving in the same area build a local ad-hoc network for the fast exchange of information in emergency situations or to help each other keep a safe distance. In case of an accident, not only will the airbag be triggered, but the police and ambulance service will be informed via an emergency call to a service provider. Buses, trucks, and trains are already transmitting maintenance and logistic information to their home base, which helps to improve organization (fleet management), and saves time and money.

2. Emergencies: An ambulance with a high-quality wireless connection to a hospital can carry vital information about injured persons to the hospital from the scene of the accident. All the necessary steps for this particular type of accident can be prepared and specialists can be consulted for an early diagnosis. Wireless networks are the only means of communication in the case of natural disasters such as hurricanes or earthquakes. In the worst cases, only decentralized, wireless ad-hoc networks survive.

3. Business: Managers can use mobile computers say, critical presentations to major customers. They can access the latest market share information. At a small recess, they can revise the presentation to take advantage of this information. They can communicate with the office about possible new offers and call meetings for discussing responds to the new proposals. Therefore, mobile computers can leverage competitive advantages. A travelling salesman today needs instant access to the company's database: to ensure that files on his or her laptop reflect the current situation, to enable the company to keep track of all activities of their travelling employees, to keep databases consistent etc. With wireless access, the laptop can be turned into a true mobile office, but efficient and powerful synchronization mechanisms are needed to ensure data consistency.

4. Credit Card Verification: At Point of Sale (POS) terminals in shops and
Supermarkets, when customers use credit cards for transactions, the intercommunication required between the bank central computer and the POS terminal, in order to effect verification of the card usage, can take place quickly and securely over cellular channels using a mobile computer unit. This can speed up the transaction process and relieve congestion at the POS terminals.

5. Replacement of Wired Networks: wireless networks can also be used to replace wired networks, e.g., remote sensors, for tradeshows, or in historic buildings. Due to economic reasons, it is often impossible to wire remote sensors for weather forecasts, earthquake detection, or to provide environmental information. Wireless connections, e.g., via satellite, can help in this situation. Other examples for wireless networks are computers, sensors, or information displays in historical buildings, where excess cabling may destroy valuable walls or floors.

6. Infotainment: wireless networks can provide up-to-date information at any appropriate location. The travel guide might tell you something about the history of a building (knowing via GPS, contact to a local base station, or triangulation where you are) downloading information about a concert in the building at the same evening via a local wireless network. Another growing field of wireless network applications lies in entertainment and games to enable, e.g., ad-hoc gaming networks as soon as people meet to play together.

Limitations of Mobile Computing

Resource constraints: Battery

Interference: Radio transmission cannot be protected against interference using shielding and result in higher loss rates for transmitted data or higher bit error rates respectively

Bandwidth: Although they are continuously increasing, transmission rates are still very low for wireless devices compared to desktop systems. Researchers look for more efficient communication protocols with low overhead.

Dynamic changes in communication environment: variations in signal power within a region, thus link delays and connection losses

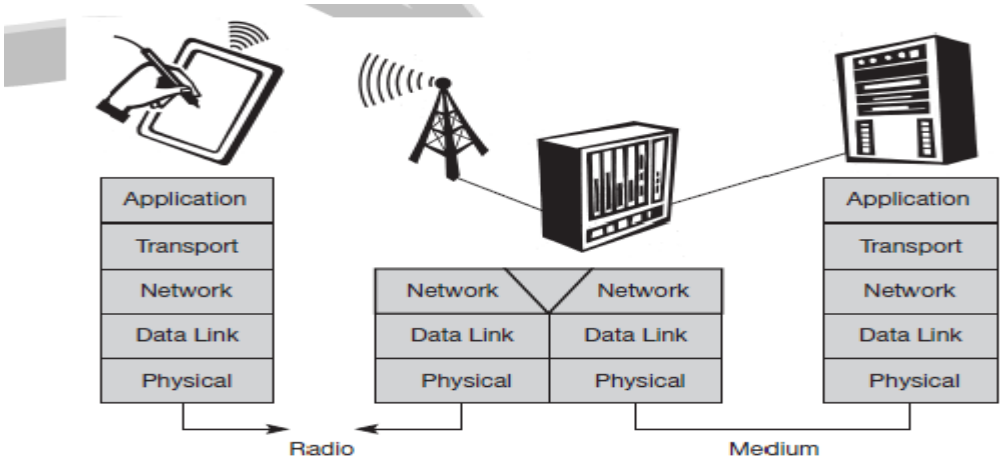
Network Issues: discovery of the connection-service to destination and connection stability

Interoperability issues: the varying protocol standards

Security constraints: Not only can portable devices be stolen more easily, but the radio interface is also prone to the dangers of eavesdropping. Wireless access must always include encryption, authentication, and other security mechanisms that must be efficient and simple to use.

A SIMPLIFIED REFERENCE MODEL

The figure shows the **protocol stack** implemented in the system according to the reference model. **End-systems**, such as the PDA and computer in the example, need a full protocol stack comprising the application layer, transport layer, network layer, data link layer, and physical layer. Applications on the end-systems communicate with each other using the lower layer services. **Intermediate systems**, such as the interworking unit, do not necessarily need all of the layers.



A Simplified Reference Model

a) **Physical layer**: This is the lowest layer in a communication system and is responsible for the conversion of a stream of bits into signals that can be transmitted on the sender side. The physical layer of the receiver then transforms the signals back into a bit stream. For wireless communication, the physical layer is responsible for frequency selection, generation of the carrier frequency, signal detection (although heavy interference may disturb the signal), modulation of data onto a carrier frequency and (depending on the transmission scheme) encryption.

b) **Data link layer**: The main tasks of this layer include accessing the medium, multiplexing of different data streams, correction of transmission errors, and synchronization (i.e., detection of a data frame). Altogether, the data link layer is responsible for a reliable point-to-point connection between two devices or a point-to-multipoint connection between one sender and several receivers.

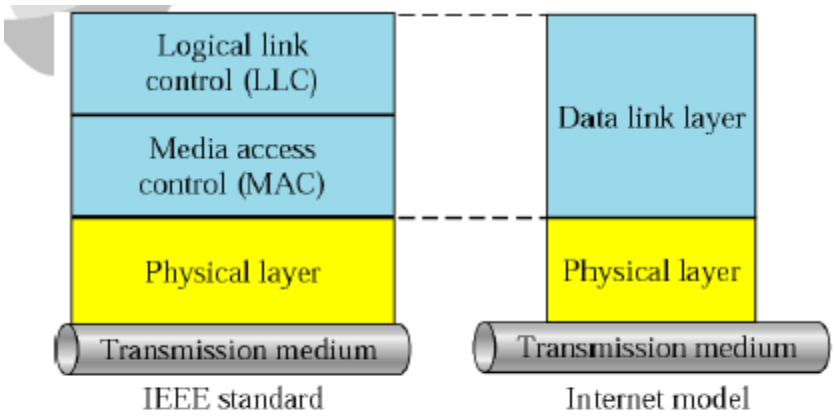
c) **Network layer**: This third layer is responsible for routing packets through a network or establishing a connection between two entities over many other intermediate systems. Important functions are addressing, routing, device location, and handover between different networks.

d) **Transport layer**: This layer is used in the reference model to establish an end-to-end connection

e) **Application layer**: Finally, the applications (complemented by additional layers that can support applications) are situated on top of all transmission oriented layers. Functions are service location, support for multimedia applications, adaptive applications that can handle the large variations in transmission characteristics, and wireless access to the world-wide web using a portable device.

Media Access Control (**MAC**)

The **Media Access Control** (**MAC**) data communication protocol sub-layer, also known as the Medium Access Control, is a sublayer of the Data Link Layer specified in the seven-layer OSI model (layer 2). The hardware that implements the MAC is referred to as a **Medium Access Controller**. The MAC sub-layer acts as an interface between the Logical Link Control (LLC) sublayer and the network's physical layer. The MAC layer emulates a full-duplex logical communication channel in a multi-point network. This channel may provide unicast, multicast or broadcast communication service.



LLC and MAC sublayers

Motivation for a specialized MAC

One of the most commonly used MAC schemes for wired networks is carrier sense multiple access with collision detection (CSMA/CD). In this scheme, a sender senses the medium (a wire or coaxial cable) to see if it is free. If the medium is busy, the sender waits until it is free. If the medium is free, the sender starts transmitting data and continues to listen into the medium. If the sender detects a collision while sending, it stops at once and sends a jamming signal. But this scheme doest work well with wireless networks. The problems are:

a) Signal strength decreases proportional to the square of the distance

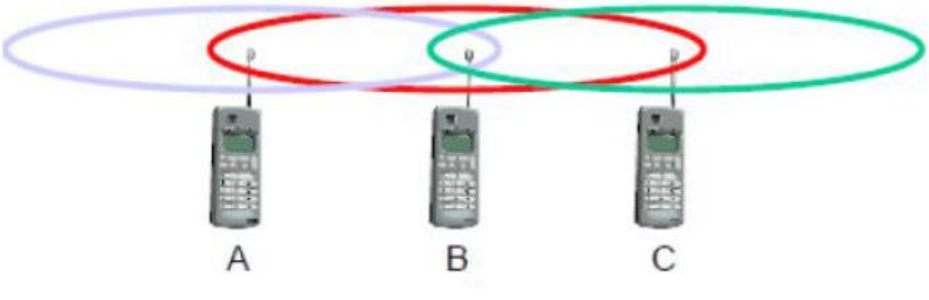
b) The sender would apply CS and CD, but the collisions happen at the receiver

c) It might be a case that a sender cannot “hear” the collision, i.e., CD does not work

d) Furthermore, CS might not work, if for e.g., a terminal is “hidden”

Hidden and Exposed Terminals

Consider the scenario with three mobile phones as shown below. The transmission range of A reaches B, but not C (the detection range does not reach C either). The transmission range of C reaches B, but not A. Finally, the transmission range of B reaches A and C, i.e., A cannot detect C and vice versa.



Hidden terminals

a) A sends to B, C cannot hear A

b) C wants to send to B, C senses a "free" medium (CS fails) and starts transmitting

c) Collision at B occurs, A cannot detect this collision (CD fails) and continues with its transmission to B

d) A is "hidden" from C and vice versa

Exposed terminals

a) B sends to A, C wants to send to another terminal (not A or B) outside the range

b) C senses the carrier and detects that the carrier is busy.

c) C postpones its transmission until it detects the medium as being idle again but A is outside radio range of C, waiting is **not** necessary

d) C is "exposed" to B

Hidden terminals cause collisions, where as Exposed terminals causes unnecessary delay.

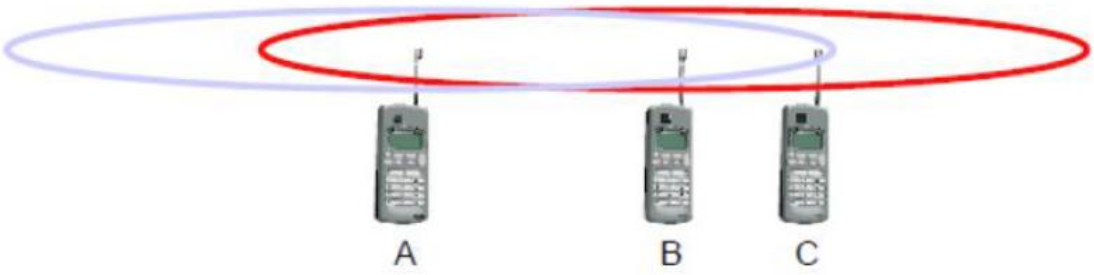
Near and far terminals

Consider the situation shown below. A and B are both sending with the same transmission power.

a) Signal strength decreases proportional to the square of the distance

b) So, B's signal drowns out A's signal making C unable to receive A's transmission

c) If C is an arbiter for sending rights, B drown out A's signal on the physical layer making C unable
to hear out A.



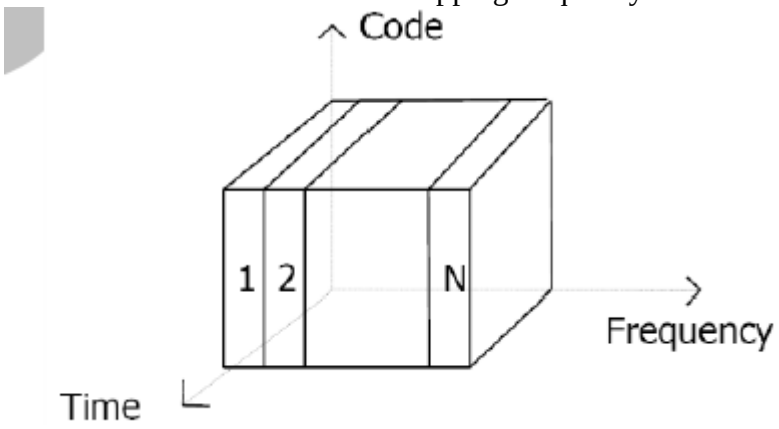
The **near/far effect** is a severe problem of wireless networks using CDM. All signals should arrive at the receiver with more or less the same strength for which Precise power control is to be implemented.

SDMA

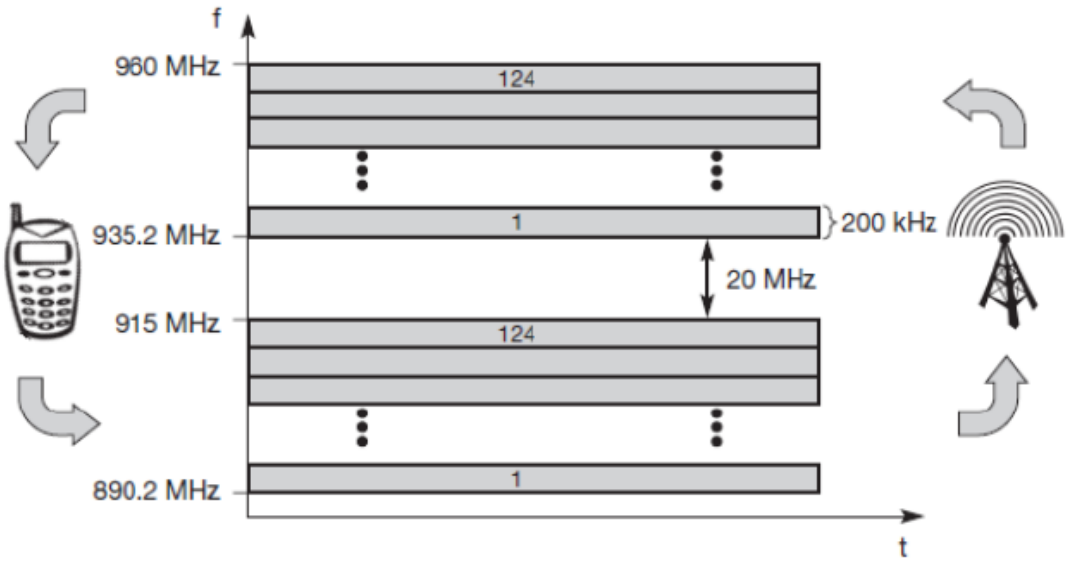
Space Division Multiple Access (SDMA) is used for allocating a separated space to users in wireless networks. A typical application involves assigning an optimal base station to a mobile phone user. The mobile phone may receive several base stations with different quality. A MAC algorithm could now decide which base station is best, taking into account which frequencies (FDM), time slots (TDM) or code (CDM) are still available. The basis for the SDMA algorithm is formed by cells and sectorized antennas which constitute the infrastructure implementing **space division multiplexing (SDM).** SDM has the unique advantage of not requiring any multiplexing equipment. It is usually combined with other multiplexing techniques to better utilize the individual physical channels.

FDMA

Frequency division multiplexing (FDM) describes schemes to subdivide the frequency dimension into several non-overlapping frequency bands.



Frequency Division Multiple Access is a method employed to permit several users to transmit simultaneously on one satellite transponder by assigning a specific frequency within the channel to each user. Each conversation gets its own, unique, radio channel. The channels are relatively narrow, usually 30 KHz or less and are defined as either transmit or receive channels. A full duplex conversation requires a transmit & receive channel pair. FDM is often used for simultaneous access to the medium by base station and mobile station in cellular networks establishing a duplex channel. A scheme called **frequency division duplexing (FDD)** in which the two directions, mobile station to base station and vice versa are now separated using different frequencies.



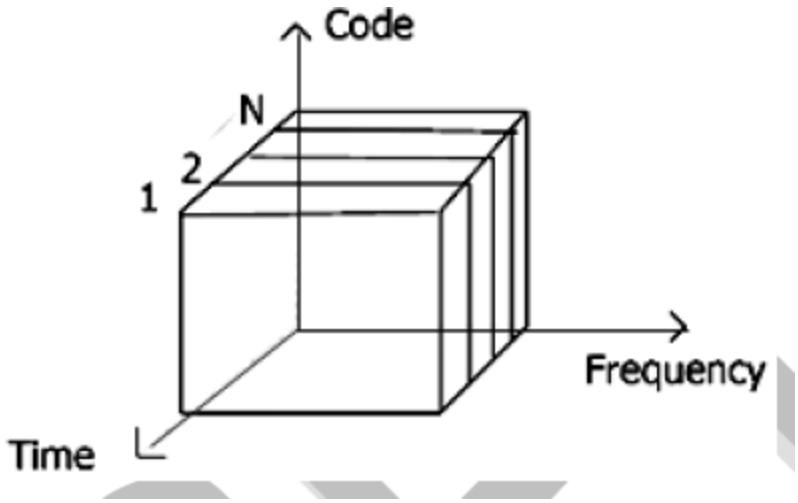
FDM for multiple access and duplex

The two frequencies are also known as **uplink**, i.e., from mobile station to base station or from ground control to satellite, and as **downlink**, i.e., from base station to mobile station or from satellite to ground control. The basic frequency allocation scheme for GSM is fixed and regulated by national authorities. All uplinks use the band between 890.2 and 915 MHz, all downlinks use 935.2 to 960 MHz. According to FDMA, the base station, shown on the right side, allocates a certain frequency for up- and downlink to establish a duplex channel with a mobile phone. Up- and downlink have a fixed relation. If the uplink frequency is fu = 890 MHz + n·0.2 MHz, the downlink frequency is fd = fu + 45 MHz, i.e**., fd = 935 MHz + n·0.2 MHz** for a certain channel n. The base station selects the channel. Each channel (uplink and downlink) has a bandwidth of 200 kHz.

This scheme also has disadvantages. While radio stations broadcast 24 hours a day, mobile communication typically takes place for only a few minutes at a time. Assigning a separate frequency for each possible communication scenario would be a tremendous waste of (scarce) frequency resources. Additionally, the fixed assignment of a frequency to a sender makes the scheme very inflexible and limits the number of senders.

TDMA

A more flexible multiplexing scheme for typical mobile communications is time division multiplexing (TDM). Compared to FDMA, time division multiple access (TDMA) offers a much more flexible scheme, which comprises all technologies that allocate certain time slots for communication. Now synchronization between sender and receiver has to be achieved in the time domain. Again this can be done by using a fixed pattern similar to FDMA techniques, i.e., allocating a certain time slot for a channel, or by using a dynamic allocation scheme.



Listening to different frequencies at the same time is quite difficult, but listening to many channels separated in time at the same frequency is simple. Fixed schemes do not need identification, but are not as flexible considering varying bandwidth requirements.

CARRIER SENSE MULTIPLE ACCESS

One improvement to the basic Aloha is sensing the carrier before accessing the medium. Sensing the carrier and accessing the medium only if the carrier is idle decreases the probability of a collision. But, as already mentioned in the introduction, hidden terminals cannot be detected, so, if a hidden terminal transmits at the same time as another sender, a collision might occur at the receiver. This basic scheme is still used in most wireless LANs. The different versions of CSMA are:

a) **1-persistent CSMA**: Stations sense the channel and listens if its busy and transmit immediately, when the channel becomes idle. It's called 1-persistent CSMA because the host transmits with a probability of 1 whenever it finds the channel idle.

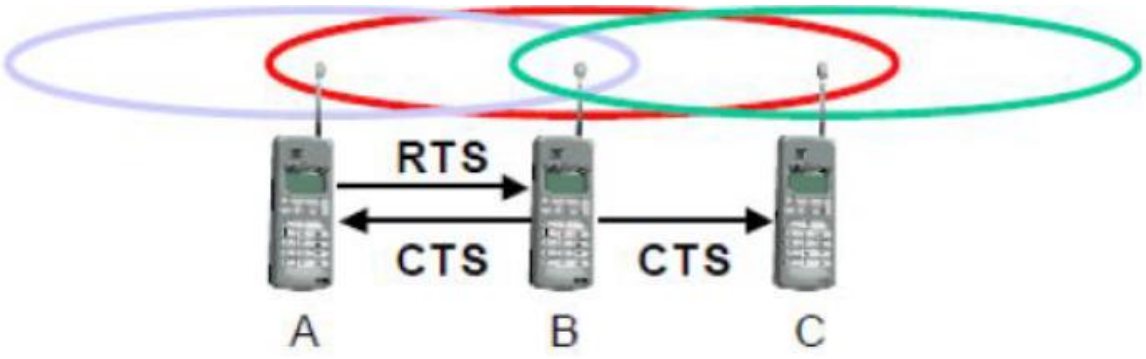
b) **non-persistent CSMA**: stations sense the carrier and start sending immediately if the medium is idle. If the medium is busy, the station pauses a random amount of time before sensing the medium again and repeating this pattern.

c) **p-persistent CSMA**: systems nodes also sense the medium, but only transmit with a probability of p, with the station deferring to the next slot with the probability 1-p, i.e., access is slotted in addition

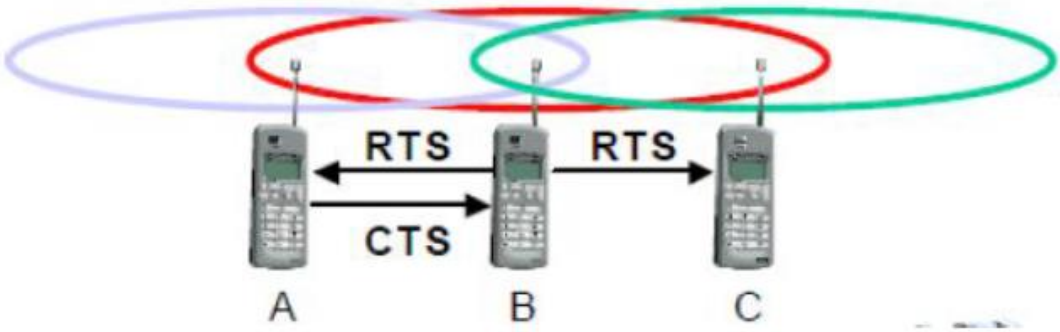
CSMA with collision avoidance (**CSMA/CA**) is one of the access schemes used in wireless LANs following the standard IEEE 802.11. Here sensing the carrier is combined with a back-off scheme in case of a busy medium to achieve some fairness among competing stations.

MULTIPLE ACCESS WITH COLLISION AVOIDANCE

Multiple access with collision avoidance (MACA) presents a simple scheme that solves the hidden terminal problem, does not need a base station, and is still a random access Aloha scheme – but with dynamic reservation. Consider the hidden terminal problem scenario.
A starts sending to B, C does not receive this transmission. C also wants to send something to B and senses the medium. The medium appears to be free, the carrier sense fails. C also starts sending causing a collision at B. But A cannot detect this collision at B and continues with its transmission. A is **hidden** for C and vice versa.
With MACA, A does not start its transmission at once, but sends a **request to send (RTS)** first. B receives the RTS that contains the name of sender and receiver, as well as the length of the future transmission. This RTS is not heard by C, but triggers an acknowledgement from B, called **clear to send (CTS)**. The CTS again contains the names of sender (A) and receiver (B) of the user data, and the length of the future transmission.



This CTS is now heard by C and the medium for future use by A is now reserved for the duration of the transmission. After receiving a CTS, C is not allowed to send anything for the duration indicated in the CTS toward B. A collision cannot occur at B during data transmission, and the hidden terminal problem is solved. Still collisions might occur when A and C transmits a RTS at the same time. B resolves this contention and acknowledges only one station in the CTS. No transmission is allowed without an appropriate CTS.
Now MACA tries to avoid the **exposed terminals** in the following way:



With MACA, B has to transmit an RTS first containing the name of the receiver (A) and the sender (B). C does not react to this message as it is not the receiver, but A acknowledges using a CTS which identifies B as the sender and A as the receiver of the following data transmission. C does not receive this CTS and concludes that A is outside the detection range. C can start its transmission assuming it will not cause a collision at A. The problem with exposed terminals is solved without fixed access patterns or a base station.

Polling

Polling schemes are used when one station wants to be heard by others. Polling is a strictly centralized scheme with one master station and several slave stations. The master can poll the slaves according to many schemes: round robin (only efficient if traffic patterns are similar over all stations), randomly, according to reservations (the classroom example with polite students) etc. The master could also establish a list of stations wishing to transmit during a contention phase. After this phase, the station polls each station on the list.

Example: Randomly Addressed Polling

- base station signals readiness to all mobile terminals
- terminals ready to send transmit random number without collision using CDMA or FDMA
- the base station chooses one address for polling from list of all random numbers (collision if two terminals choose the same address)
- the base station acknowledges correct packets and continues polling the next terminal
- this cycle starts again after polling all terminals of the list

CDMA

Code division multiple access systems apply codes with certain characteristics to the transmission to separate different users in code space and to enable access to a shared medium without interference.

All terminals send on the same frequency probably at the same time and can use the whole bandwidth of the transmission channel. Each sender has a unique random number, the sender XORs the signal with this random number. The receiver can "tune" into this signal if it knows the pseudo random number, tuning is done via a correlation function

Disadvantages:

1. higher complexity of a receiver (receiver cannot just listen into the medium and start receiving if there is a signal)
2. all signals should have the same strength at a receiver

Advantages:

1. all terminals can use the same frequency, no planning needed
2. huge code space (e.g. 232) compared to frequency space
3. interferences (e.g. white noise) is not coded
4. forward error correction and encryption can be easily integrated

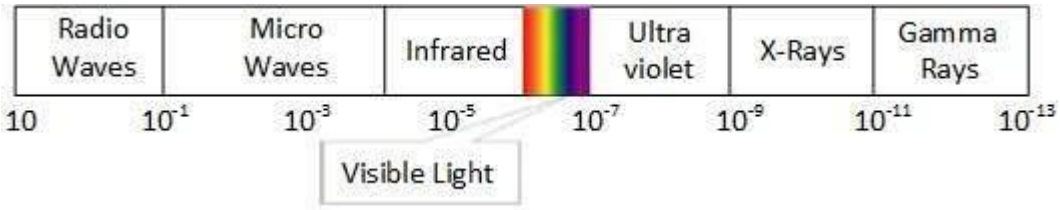
UNIT 2

Wireless Transmissions

Wireless transmission is a form of unguided media. Wireless communication involves no physical link established between two or more devices, communicating wirelessly. Wireless signals are spread over in the air and are received and interpreted by appropriate antennas.

When an antenna is attached to electrical circuit of a computer or wireless device, it converts the digital data into wireless signals and spread all over within its frequency range. The receptor on the other end receives these signals and converts them back to digital data.

A little part of electromagnetic spectrum can be used for wireless transmission.

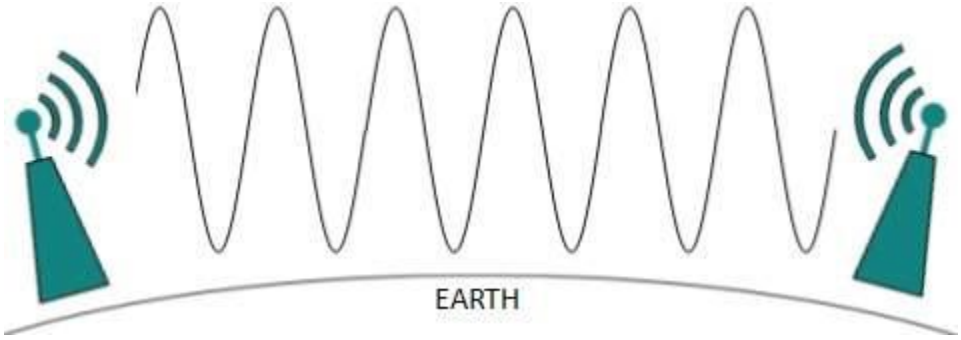


Radio Transmission

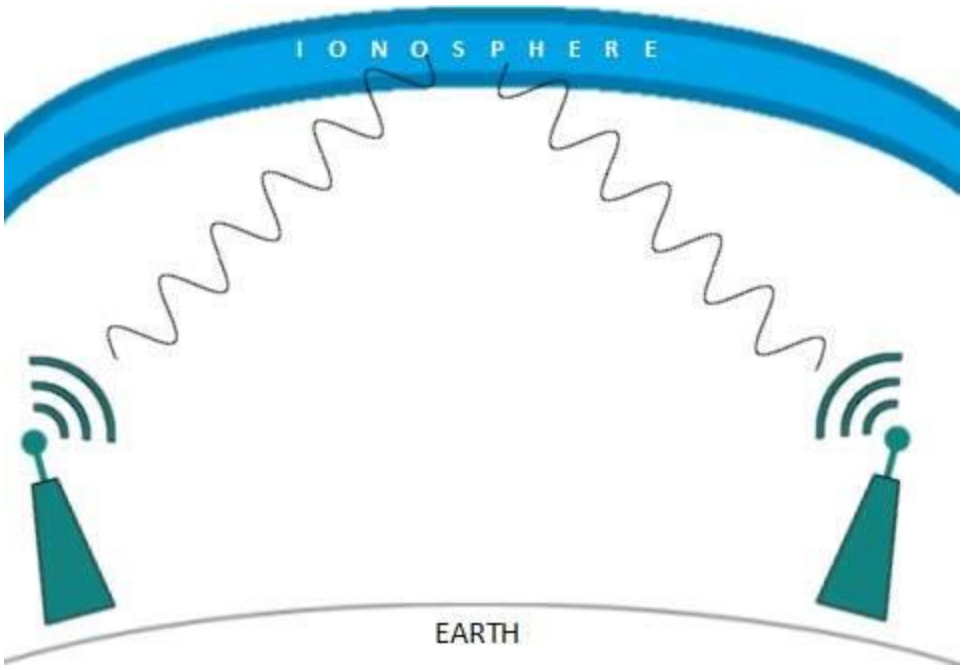
Radio frequency is easier to generate and because of its large wavelength it can penetrate through walls and structures alike.Radio waves can have wavelength from 1 mm – 100,000 km and have frequency ranging from 3 Hz (Extremely Low Frequency) to 300 GHz (Extremely High Frequency). Radio frequencies are sub-divided into six bands.

Radio waves at lower frequencies can travel through walls whereas higher RF can travel in straight line and bounce back.The power of low frequency waves decreases sharply as they cover long distance. High frequency radio waves have more power.

Lower frequencies such as VLF, LF, MF bands can travel on the ground up to 1000 kilometers, over the earth's surface.



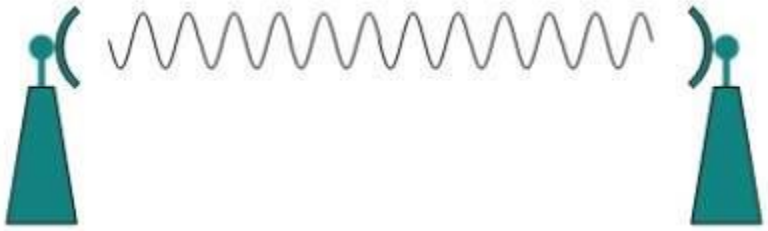
Radio waves of high frequencies are prone to be absorbed by rain and other obstacles. They use Ionosphere of earth atmosphere. High frequency radio waves such as HF and VHF bands are spread upwards. When they reach Ionosphere, they are refracted back to the earth.



Microwave Transmission

Electromagnetic waves above 100 MHz tend to travel in a straight line and signals over them can be sent by beaming those waves towards one particular station. Because Microwaves travels in straight lines, both sender and receiver must be aligned to be strictly in line-of-sight.

Microwaves can have wavelength ranging from 1 mm – 1 meter and frequency ranging from 300 MHz to 300 GHz.



Microwave antennas concentrate the waves making a beam of it. As shown in picture above, multiple antennas can be aligned to reach farther. Microwaves have higher frequencies and do not penetrate wall like obstacles.

Microwave transmission depends highly upon the weather conditions and the frequency it is using.

Infrared Transmission

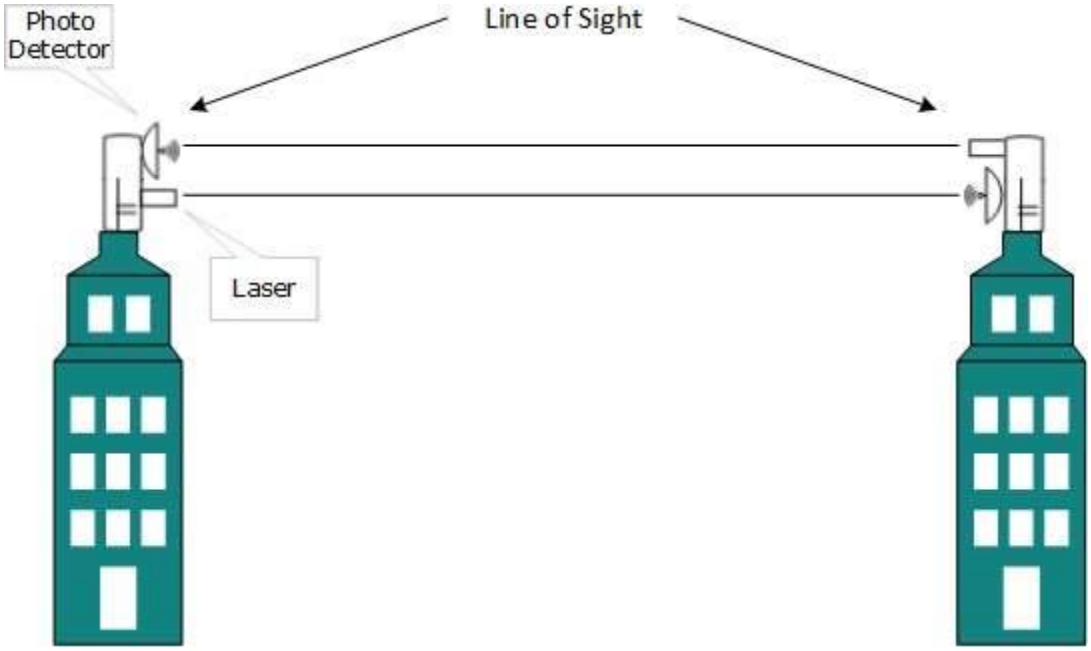
Infrared wave lies in between visible light spectrum and microwaves. It has wavelength of 700-nm to 1-mm and frequency ranges from 300-GHz to 430-THz.

Infrared wave is used for very short range communication purposes such as television and it's remote. Infrared travels in a straight line hence it is directional by nature. Because of high frequency range, Infrared cannot cross wall-like obstacles.

Light Transmission

Highest most electromagnetic spectrum which can be used for data transmission is light or optical signaling. This is achieved by means of LASER.

Because of frequency light uses, it tends to travel strictly in straight line.Hence the sender and receiver must be in the line-of-sight. Because laser transmission is unidirectional, at both ends of communication the laser and the photo-detector needs to be installed. Laser beam is generally 1mm wide hence it is a work of precision to align two far receptors each pointing to lasers source.



Laser works as Tx (transmitter) and photo-detectors works as Rx (receiver).

Lasers cannot penetrate obstacles such as walls, rain, and thick fog. Additionally, laser beam is distorted by wind, atmosphere temperature, or variation in temperature in the path.

Laser is safe for data transmission as it is very difficult to tap 1mm wide laser without interrupting the communication channel.

MULTIPLEXING

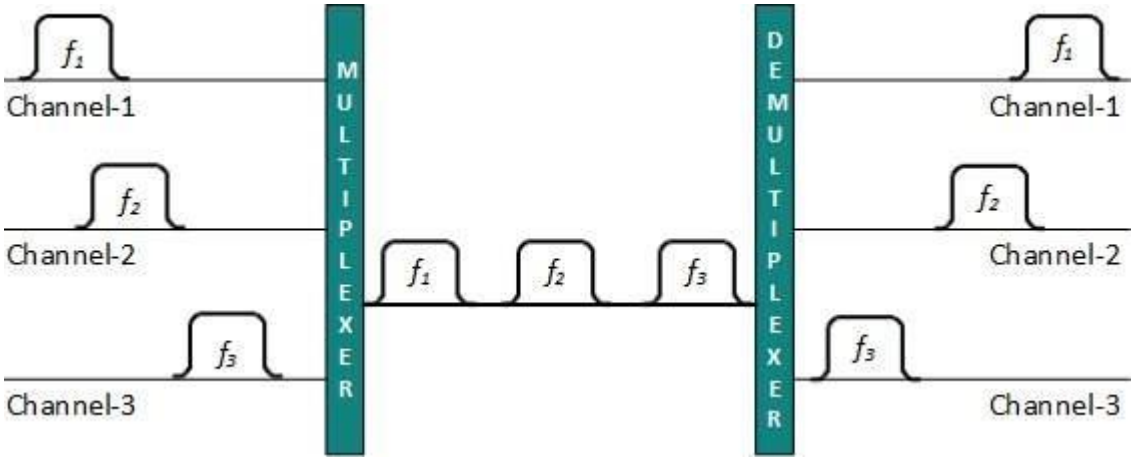
Multiplexing is a technique by which different analog and digital streams of transmission can be simultaneously processed over a shared link. Multiplexing divides the high capacity medium into low capacity logical medium which is then shared by different streams.

Communication is possible over the air (radio frequency), using a physical media (cable), and light (optical fiber). All mediums are capable of multiplexing.

When multiple senders try to send over a single medium, a device called Multiplexer divides the physical channel and allocates one to each. On the other end of communication, a De-multiplexer receives data from a single medium, identifies each, and sends to different receivers.

Frequency Division Multiplexing

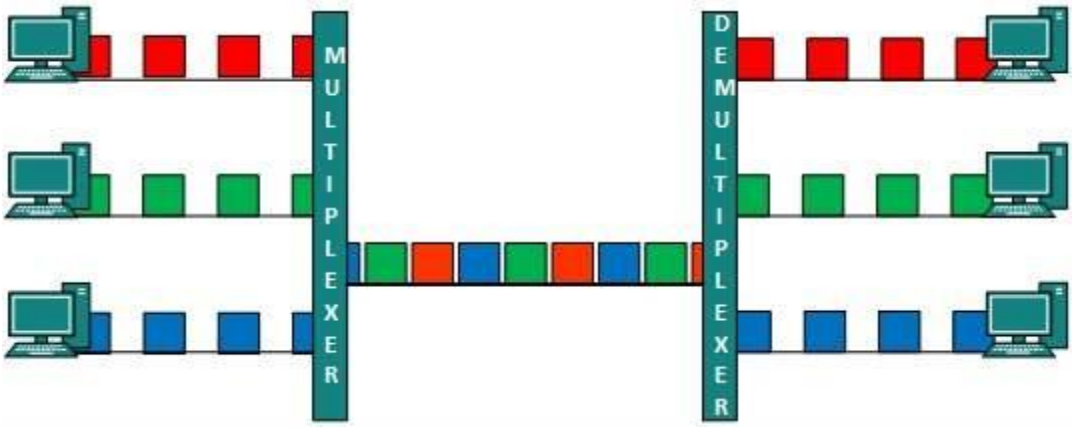
When the carrier is frequency, FDM is used. FDM is an analog technology. FDM divides the spectrum or carrier bandwidth in logical channels and allocates one user to each channel. Each user can use the channel frequency independently and has exclusive access of it. All channels are divided in such a way that they do not overlap with each other. Channels are separated by guard bands. Guard band is a frequency which is not used by either channel.



Time Division Multiplexing

TDM is applied primarily on digital signals but can be applied on analog signals as well. In TDM the shared channel is divided among its user by means of time slot. Each user can transmit data within the provided time slot only. Digital signals are divided in frames, equivalent to time slot i.e. frame of an optimal size which can be transmitted in given time slot.

TDM works in synchronized mode. Both ends, i.e. Multiplexer and De-multiplexer are timely synchronized and both switch to next channel simultaneously.



When channel A transmits its frame at one end,the De-multiplexer provides media to channel A on the other end.As soon as the channel A's time slot expires, this side switches to channel B. On the other end, the De-multiplexer works in a synchronized manner and provides media to channel B. Signals from different channels travel the path in interleaved manner.

Wavelength Division Multiplexing

Light has different wavelength (colors). In fiber optic mode, multiple optical carrier signals are multiplexed into an optical fiber by using different wavelengths. This is an analog multiplexing technique and is done conceptually in the same manner as FDM but uses light as signals.

Further, on each wavelength time division multiplexing can be incorporated to accommodate more data signals.

Code Division Multiplexing

Multiple data signals can be transmitted over a single frequency by using Code Division Multiplexing. FDM divides the frequency in smaller channels but CDM allows its users to full bandwidth and transmit signals all the time using a unique code. CDM uses orthogonal codes to spread signals.

Each station is assigned with a unique code, called chip. Signals travel with these codes independently, inside the whole bandwidth.The receiver knows in advance the chip code signal it has to receive.

SWITCHING

Switching is process to forward packets coming in from one port to a port leading towards the destination. When data comes on a port it is called ingress, and when data leaves a port or goes out it is called egress. A communication system may include number of switches and nodes. At broad level, switching can be divided into two major categories:

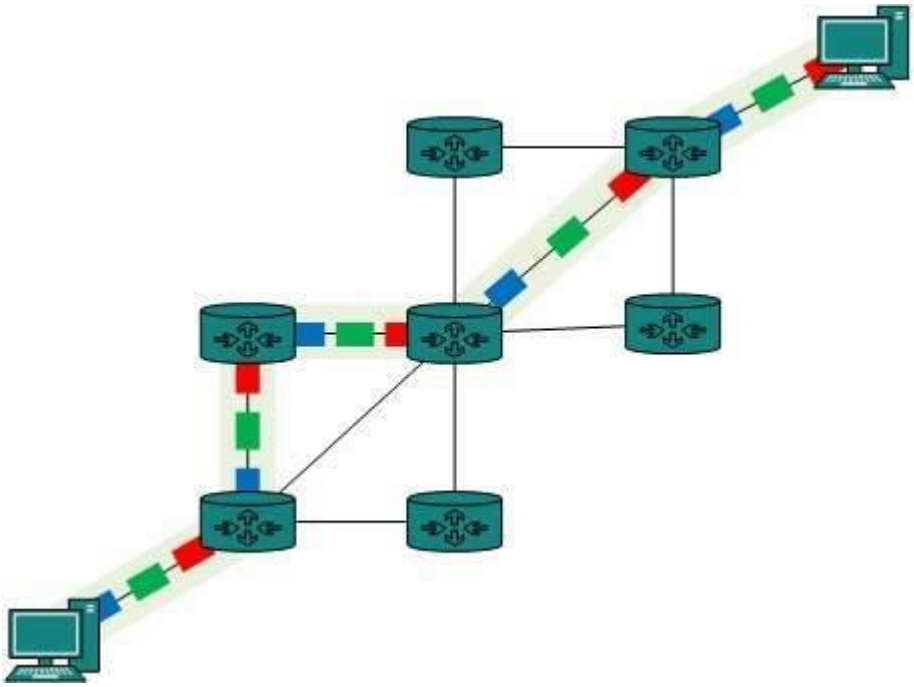
- **Connectionless:** The data is forwarded on behalf of forwarding tables. No previous handshaking is required and acknowledgements are optional.
- **Connection Oriented:** Before switching data to be forwarded to destination, there is a need to pre-establish circuit along the path between both endpoints. Data is then forwarded on that circuit. After the transfer is completed, circuits can be kept for future use or can be turned down immediately.

Circuit Switching

When two nodes communicate with each other over a dedicated communication path, it is called circuit switching. There 'is a need of pre-specified route from which data will travels and no other data is permitted. In circuit switching, to transfer the data, circuit must be established so that the data transfer can take place.

Circuits can be permanent or temporary. Applications which use circuit switching may have to go through three phases:

- Establish a circuit
- Transfer the data
- Disconnect the circuit

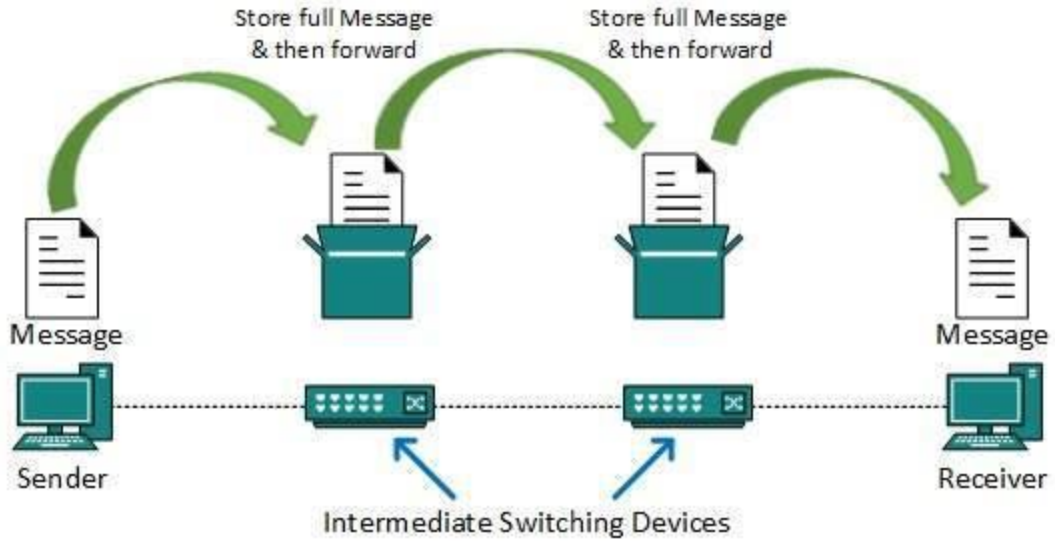


Circuit switching was designed for voice applications. Telephone is the best suitable example of circuit switching. Before a user can make a call, a virtual path between caller and callee is established over the network.

Message Switching

This technique was somewhere in middle of circuit switching and packet switching. In message switching, the whole message is treated as a data unit and is switching / transferred in its entirety.

A switch working on message switching, first receives the whole message and buffers it until there are resources available to transfer it to the next hop. If the next hop is not having enough resource to accommodate large size message, the message is stored and switch waits.



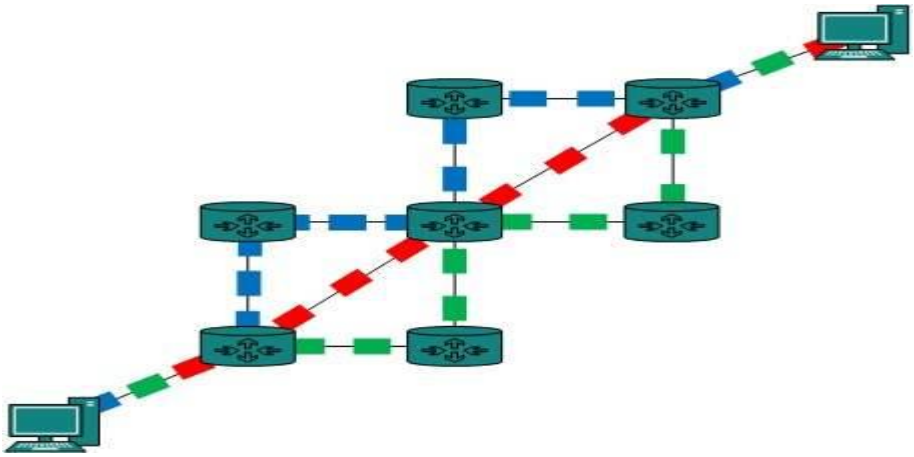
This technique was considered substitute to circuit switching. As in circuit switching the whole path is blocked for two entities only. Message switching is replaced by packet switching. Message switching has the following drawbacks:

- Every switch in transit path needs enough storage to accommodate entire message.
- Because of store-and-forward technique and waits included until resources are available, message switching is very slow.
- Message switching was not a solution for streaming media and real-time applications.

Packet Switching

Shortcomings of message switching gave birth to an idea of packet switching. The entire message is broken down into smaller chunks called packets. The switching information is added in the header of each packet and transmitted independently.

It is easier for intermediate networking devices to store small size packets and they do not take much resources either on carrier path or in the internal memory of switches.



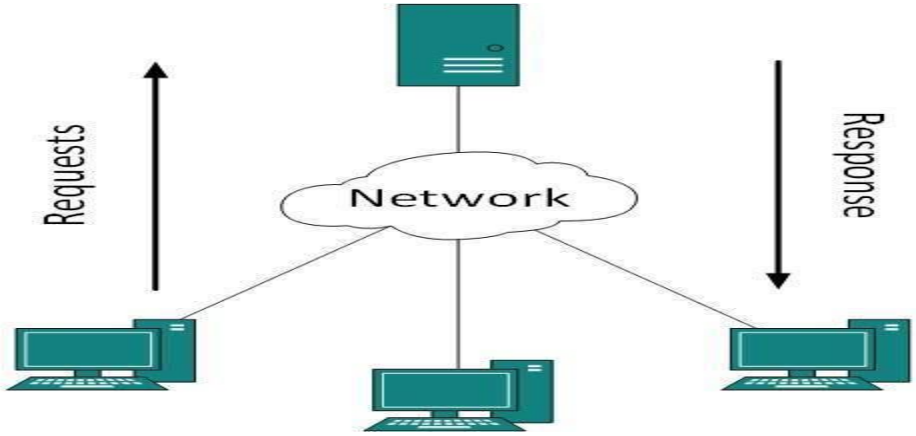
Packet switching enhances line efficiency as packets from multiple applications can be multiplexed over the carrier. The internet uses packet switching technique. Packet switching enables the user to differentiate data streams based on priorities. Packets are stored and forwarded according to their priority to provide quality of service.

CLIENT SERVER MODEL

Two remote application processes can communicate mainly in two different fashions:

- **Peer-to-peer:** Both remote processes are executing at same level and they exchange data using some shared resource.
- **Client-Server:** One remote process acts as a Client and requests some resource from another application process acting as Server.

In client-server model, any process can act as Server or Client. It is not the type of machine, size of the machine, or its computing power which makes it server; it is the ability of serving request that makes a machine a server.



A system can act as Server and Client simultaneously. That is, one process is acting as Server and another is acting as a client. This may also happen that both client and server processes reside on the same machine.

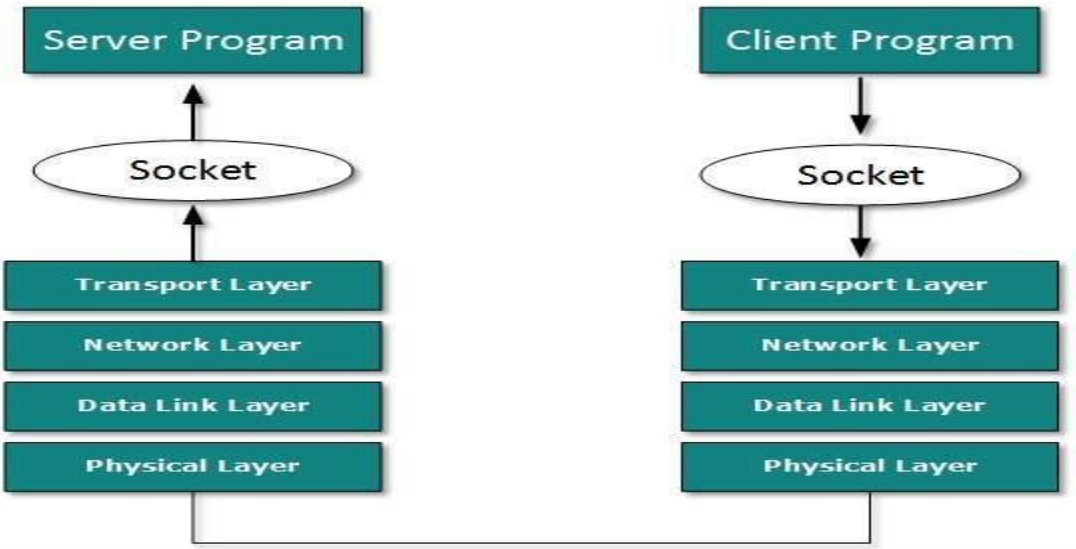
Communication

Two processes in client-server model can interact in various ways:

- Sockets
- Remote Procedure Calls (RPC)

Sockets

In this paradigm, the process acting as Server opens a socket using a well-known (or known by client) port and waits until some client request comes. The second process acting as a Client also opens a socket but instead of waiting for an incoming request, the client processes 'requests first'.



When the request is reached to server, it is served. It can either be an information sharing or resource request.

Remote Procedure Call

This is a mechanism where one process interacts with another by means of procedure calls. One process (client) calls the procedure lying on remote host. The process on remote host is said to be Server. Both processes are allocated stubs. This communication happens in the following way:

- The client process calls the client stub. It passes all the parameters pertaining to program local to it.
- All parameters are then packed (marshalled) and a system call is made to send them to other side of the network.
- Kernel sends the data over the network and the other end receives it.
- The remote host passes data to the server stub where it is unmarshalled.
- The parameters are passed to the procedure and the procedure is then executed.
- The result is sent back to the client in the same manner.

UNIT 3

UMTS (Universal Mobile Telecommunications System)

UMTS will allow a future mass market for high-quality wireless multimedia communications that will approach two billion users worldwide by the year 2010.

This new technology will deliver low-cost, high-capacity wireless communications, offering data rates of 1Mbps to 2Mbps with global roaming and other advanced UMTS services.

UMTS (Universal Mobile Telecommunications Service)

UMTS (Universal Mobile Telecommunications Service) is a third-generation (3G) broadband, packet-based transmission of text, digitized voice, video, and multimedia at data rates up to 2 megabits per second (Mbps). UMTS offers a consistent set of services to mobile computer and phone users, no matter where they are located in the world. UMTS is based on the Global System for Mobile (GSM) communication standard. It is also endorsed by major standards bodies and manufacturers as the planned standard for mobile users around the world. Once UMTS is fully available, computer and phone users can be constantly attached to the Internet wherever they travel and, as they roam, will have the same set of capabilities. Users will have access through a combination of terrestrial wireless and satellite transmissions. Until UMTS is fully implemented, users can use multi-mode devices that switch to the currently available technology (such as GSM 900 and 1800) where UMTS is not yet available.

First Generation Systems

All these systems were analog systems, using FDMA technology. They are also known as First Generation (1G) systems. Different systems came into use based on the cellular principle. They are listed below.

Year	Mobile System
1981	Nordic Mobile Telephone(NMT)450
1982	American Mobile Phone System(AMPS)
1985	Total Access Communication System(TACS)
1986	Nordic Mobile Telephony(NMT)900

Disadvantages of 1G systems

- They were analog and hence are were not robust to interference.
- Different countries followed their own standards, which were incompatible.

To overcome the difficulties of 1G, digital technology was chosen by most of the countries and a new era, called 2G, started.

Advantages of 2G

- Improved Spectral Utilization achieved by using advanced modulation techniques.
- Lower bit rate voice coding enabled more users getting the services simultaneously.
- Reduction of overhead in signaling paved way for capacity enhancement.
- Good source and channel coding techniques make the signal more robust to Interference.
- New services like SMS were included.
- Improved efficiency of access and hand-off control were achieved.

Name of the Systems	**Country**
DAMPS-Digital Advanced Mobile Phone System	North America
GSM-Global System for Mobile communication	European Countries and International applications
JDC - Japanese Digital Cellular	Japan
CT-2 Cordless Telephone–2	UK
DECT-Digital European Cordless Telephone	European countries

History of GSM

GSM standard is a European standard, which has addressed many problems related to compatibility, especially with the development of digital radio technology.

Milestones of GSM

- 1982 - Confederation of European Post and Telegraph (CEPT) establishes Group Special Mobile.

- 1985 - Adoption of list of recommendation was decided to be generated by the group.
- 1986 - Different field tests were done for radio technique for the common air interface.
- 1987 - TDMA was chosen as the Access Standard. MoU was signed between 12 operators.
- 1988 - Validation of system was done.
- 1989 - Responsibility was taken up by European Telecommunication Standards Institute (ETSI).
- 1990 - First GSM specification was released.
- 1991 - First commercial GSM system was launched.

Frequency Range of GSM

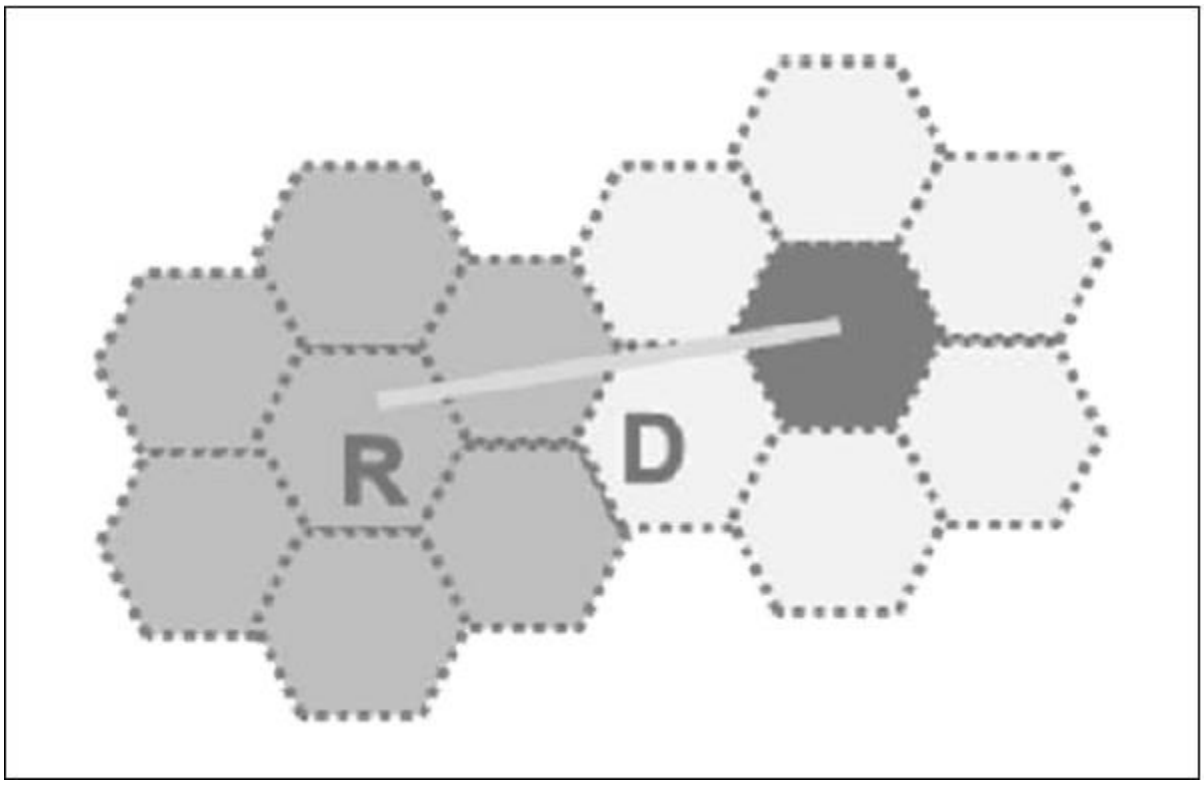
GSM works on four different frequency ranges with FDMA-TDMA and FDD. They are as follows −

System	**P-GSM (Primary)**	**E-GSM (Extended)**	**GSM 1800**	**GSM 1900**
Freq Uplink	890-915MHz	880-915MHz	1710-1785Mhz	1850-1910MHz
Freq Downlink	935-960MHz	925-960MHz	1805-1880Mhz	1930-1990MHz

Cellular Approach

With limited frequency resource, cellular principle can serve thousands of subscribers at an affordable cost. In a cellular network, total area is subdivided into smaller areas called “cells”. Each cell can cover a limited number of mobile subscribers within its boundaries. Each cell can have a base station with a number of RF channels.

Frequencies used in a given cell area will be simultaneously reused at a different cell which is geographically separated. For example, a typical seven-cell pattern can be considered.



Total available frequency resources are divided into seven parts, each part consisting of a number of radio channels and allocated to a cell site. In a group of 7 cells, available frequency spectrum is consumed totally. The same seven sets of frequency can be used after certain distance.

The group of cells where the available frequency spectrum is totally consumed is called a cluster of cells.

Two cells having the same number in the adjacent cluster, use the same set of RF channels and hence are termed as “Co-channel cells”. The distance between the cells using the same frequency should be sufficient to keep the co-channel (co-chl) interference to an acceptable level. Hence, the cellular systems are limited by Co-channel interference.

Hence a cellular principle enables the following.

- More efficient usage of available limited RF source.
- Manufacturing of every piece of subscriber's terminal within a region with the same set of channels so that any mobile can be used anywhere within the region.

Shape of Cells

For analytical purposes a “Hexagon” cell is preferred to other shapes on paper due to the following reasons.

- A hexagon layout requires fewer cells to cover a given area. Hence, it envisages fewer base stations and minimum capital investment.

- Other geometrical shapes cannot effectively do this. For example, if circular shaped cells are there, then there will be overlapping of cells.
- Also for a given area, among square, triangle and hexagon, radius of a hexagon will be the maximum which is needed for weaker mobiles.

In reality cells are not hexagonal but irregular in shape, determined by factors like propagation of radio waves over the terrain, obstacles, and other geographical constraints. Complex computer programs are required to divide an area into cells. One such program is "Tornado" from Siemens.

Operating Environment

Due to mobility, the radio signals between a base station and mobile terminals undergo a variety of alterations as they travel from transmitter to receiver, even within the same cell. These changes are due to −

- Physical separation of transmitter and receiver.
- Physical environment of the path i.e. terrain, buildings, and other obstacles.

Slow Fading

- In free space conditions (or) LOS, RF signal propagation constant is considered as two i.e. r = 2. This is applicable for static radio systems.
- In mobile environment, these variations are appreciable and normally 'r' is taken as 3 to 4.

Rayleigh Fading

The direct line of sight in mobile environment, between base station and the mobile is not ensured and the signal received at the receiver is the sum of a number of signals reaching through different paths (multipath). Multipath propagation of RF waves is due to the reflection of RF energy from a hill, building, truck, or aero plane etc.; the reflected energy undergoes a phase change also.

If there are 180 out-of phase with direct path signals, they tend to cancel out each other. So the multipath signals tend to reduce the signal strength. Depending upon the location of the transmitter and receiver and various reflecting obstacles along the path length, signal fluctuates. The fluctuations occur fast and it is known as "Rayleigh fading".

In addition, multipath propagation leads to "pulse widening" and "Inter symbol Interference".

Doppler Effect

Due to the mobility of the subscriber, a change occurs in the frequency of the received RF signals. Cellular mobile systems use following techniques to counter these problems.

- Channel coding
- Interleaving

- Equalization
- Rake receivers
- Slow frequency hopping
- Antennae diversity

2G networks follow digital system for communication. This improved the audio quality in transmission. In 2G networks phone conversations are digitally encrypted.These networks provided far greater mobile phone coverage. 2G networks also introduced data services for mobile. Picture messages and MMS (Multimedia Messaging Service) were introduced. The two popular standards introduced by 2G systems are GSM and CDMA.

3G wireless network technology provides high data transfer rates for handheld devices. The high data transfer rates allows 3G networks to offer multimedia services combining voice and data. 3G is also referred to as wireless broadband as it has the facility to send and receive large amounts of data using a mobile phone. The access part in 3G networks uses WCDMA (Wideband Code Division Multiple Access). It requires upgrading the base stations (mobile towers) and mobile phones. Also the base stations need to be close to each other.

A 4G system, also called Long Term Evolution (L.T.E.), provides mobile ultra-broadband Internet access to mobile devices. 4G networks offer very high speeds and provides excellent performance for bandwidth intensive applications such as high quality streaming video. One of the key requirements for 4G is a wireless IP- based access system. The access part in 4G networks uses OFDMA (Orthogonal Frequency Division Multiple Access). 4G provides good quality images and videos than TV.

UNIT 4

Mobile IP

The IP addresses are designed to work with stationary hosts because part of the address defines the network to which the host is attached**.** A host cannot change its IP address without terminating on-going sessions and restarting them after it acquires a new address. Other link layer mobility solutions exist but are not sufficient enough for the global Internet.

Mobility is the ability of a node to change its point-of-attachment while maintaining all existing communications and using the same IP address.

Nomadicity allows a node to move but it must terminate all existing communications and then can initiate new connections with a new address. Mobile IP is a network layer solution for homogenous and heterogeneous mobility on the global Internet which is scalable, robust, secure and which allows nodes to maintain all ongoing communications while moving.

Design Goals:

Mobile IP was developed as a means for transparently dealing with problems of mobile users. Mobile IP was designed to make the size and the frequency of required routing updates as small as possible. It was designed to make it simple to implement mobile node software. It was designed to avoid solutions that require mobile nodes to use multiple addresses.

Requirements:
There are several requirements for Mobile IP to make it as a standard. Some of them are:
1. *Compatibility*: The whole architecture of internet is very huge and a new standard cannot introduce changes to the applications or network protocols already in use. Mobile IP is to be integrated into the existing operating systems. Also, for routers also it may be possible to enhance its capabilities to support mobility instead of changing the routers which is highly impossible. Mobile IP must not require special media or MAC/LLC protocols, so it must use the same interfaces and mechanisms to access the lower layers as IP does. Finally, end-systems enhanced with a mobile IP implementation should still be able to communicate with fixed systems without mobile IP.

2. *Transparency*: Mobility remains invisible for many higher layer protocols and applications. Higher layers continue to work even if the mobile computer has changed its point of attachment to the network and even notice a lower bandwidth and some interruption in the service. As many of today's applications have not been designed to use in mobile environments, the effects of mobility will be higher delay and lower bandwidth.

3. *Scalability and efficiency*: The efficiency of the network should not be affected even if a new mechanism is introduced into the internet. Enhancing IP for mobility must not generate many new messages flooding the whole network. Special care is necessary to be taken considering the lower bandwidth of wireless links. Many mobile systems have a wireless link to an attachment point. Therefore, only some additional packets must be necessary between a mobile system and a node in the network. It is indispensable for a mobile IP to be scalable over a large number of participants in the whole internet, throughout the world.

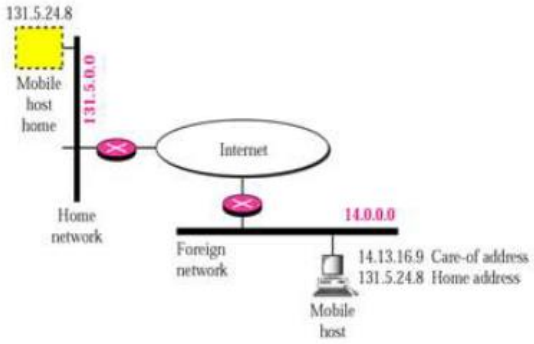
4. *Security*: Mobility possesses many security problems. A minimum requirement is the authentication of all messages related to the management of mobile IP. It must be sure for the IP layer if it forwards a packet to a mobile host that this host really is the receiver of the packet. The IP layer can only guarantee that the IP address of the receiver is correct. There is no way to prevent faked IP addresses and other attacks.

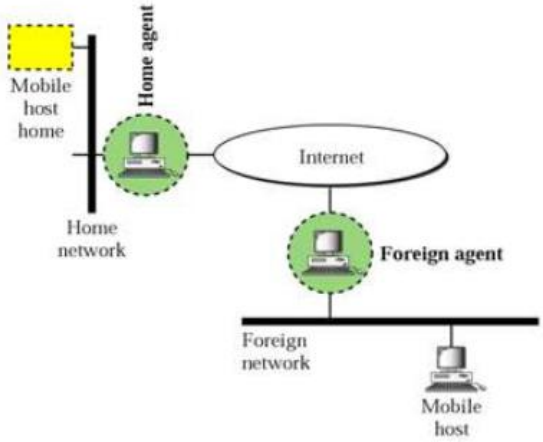
The goal of a mobile IP can be summarized as: 'supporting end-system mobility while maintaining scalability, efficiency, and compatibility in all respects with existing applications and Internet protocols'.

Entities and terminology
The following defines several entities and terms needed to understand mobile IP as defined in RFC 3344.

Mobile Node (MN):
A mobile node is an end-system or router that can change its point of attachment to the internet using mobile IP. The MN keeps its IP address and can continuously communicate with any other system in the internet as long as link-layer connectivity is given. Examples are laptop, mobile phone, router on an aircraft etc.

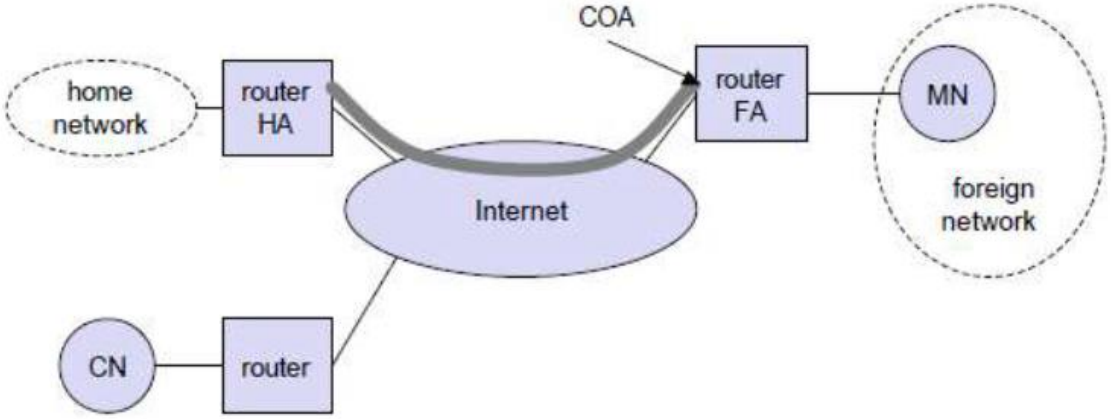
Correspondent node (CN):
At least one partner is needed for communication. In the following the CN represents this partner for the MN. The CN can be a fixed or mobile node.
Home network:
The home network is the subnet the MN belongs to with respect to its IP address. No mobile IP support is needed within the home network.
Foreign network:
The foreign network is the current subnet the MN visits and which is not the home Network





Foreign agent (FA):
The FA can provide several services to the MN during its visit to the foreign network. The FA can have the COA, acting as tunnel endpoint and forwarding packets to the MN. The FA can be the default router for the MN. FAs can also provide security services because they belong to the foreign network as opposed to the MN which is only visiting. FA is implemented on a router for the subnet the MN attaches to.
Care-of address (COA):
The COA defines the current location of the MN from an IP point of view. All IP packets sent to the MN are delivered to the COA, not directly to the IP address of the MN. Packet delivery toward the MN is done using a tunnel, i.e., the COA marks the tunnel endpoint, i.e., the address where packets exit the tunnel. There are two different possibilities for the location of the COA:
Foreign agent COA:
The COA could be located at the FA, i.e., the COA is an IP address of the FA. The FA is the tunnel end-point and forwards packets to the MN. Many MN using the FA can share this COA as common COA.
Co-located COA:
The COA is co-located if the MN temporarily acquired an additional IP address which acts as COA. This address is now topologically correct, and the tunnel endpoint is at the MN. Co-located addresses can be acquired using services such as DHCP.
Home agent (HA):
The HA provides several services for the MN and is located in the home network. The tunnel for packets toward the MN starts at the HA. The HA maintains a location registry, i.e., it is informed of the MN's location by the current COA. Three alternatives for the implementation of an HA exist.

1. The HA can be implemented on a router that is responsible for the home network. This is obviously the best position, because without optimizations to mobile IP, all packets for the MN have to go through the router anyway.
2. If changing the router's software is not possible, the HA could also be implemented on an arbitrary node in the subnet. One disadvantage of this solution is the double



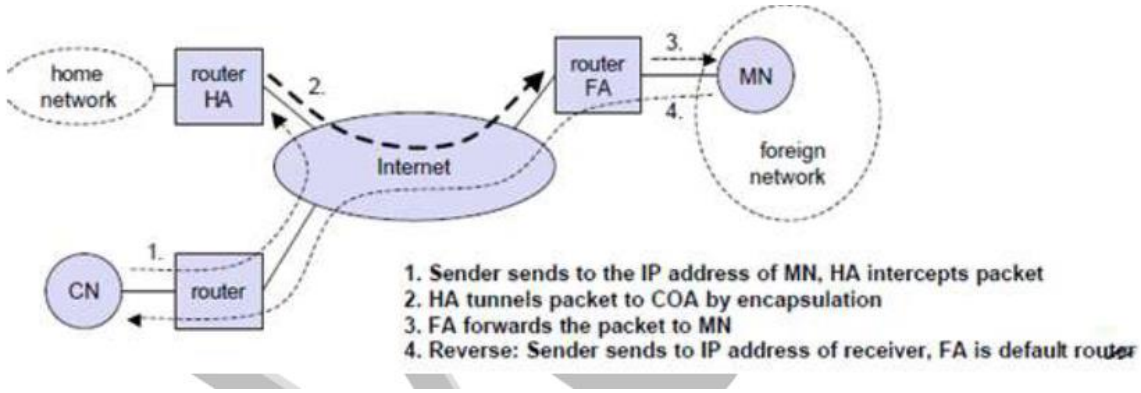
crossing of the router by the packet if the MN is in a foreign network. A packet for the MN comes in via the router; the HA sends it through the tunnel which again crosses the router. Finally, a home network is not necessary at all. The HA could be again on the 'router' but this time only acting as a manager for MNs belonging to a virtual home network. All MNs are always in a foreign network with this solution. A CN is connected via a router to the internet, as are the home network and the foreign network. The HA is implemented on the router connecting the home network with the internet, an FA is implemented on the router to the foreign network. The MN is currently in the foreign network. The tunnel for packets toward the MN starts at the HA and ends at the FA, for the FA has the COA in the above example.

IP packet delivery

Consider the above example in which a correspondent node (CN) wants to send an IP packet to the MN. One of the requirements of mobile IP was to support hiding the mobility of the MN. CN does not need to know anything about the MN's current location and sends the packet as usual to the IP address of MN as shown below.

CN sends an IP packet with MN as a destination address and CN as a source address. The internet, not having information on the current location of MN, routes the packet to the router responsible for the home network of MN. This is done using the standard routing mechanisms of the internet. The HA now intercepts the packet, knowing that MN is currently not in its home network. The packet is not forwarded into the subnet as usual, but encapsulated and tunnelled to the COA. A new header is put in front of the old IP header showing the COA as new destination and HA as source of the encapsulated packet (step 2).

The foreign agent now decapsulates the packet, i.e., removes the additional header, and forwards the original packet with CN as source and MN as destination to the MN (step 3). Again, for the MN mobility is not visible. It receives the packet with the same sender and receiver address as it would have done in the home network.



Sending packets from the mobile node (MN) to the CN is comparatively simple. The MN sends the packet as usual with its own fixed IP address as source and CN's address as destination (step 4).

The router with the FA acts as default router and forwards the packet in the same way as it would do for any other node in the foreign network. As long as CN is a fixed node the remainder is in the fixed internet as usual. If CN were also a mobile node residing in a foreign network, the same mechanisms as described in steps 1 through 3 would apply now in the other direction.

Working of Mobile IP:-

Mobile IP has two addresses for a mobile host: one home address and one care-of address. The home address is permanent; the care-of address changes as the mobile host moves from one network to another. To make the change of address transparent to the rest of the Internet requires a home agent and a foreign agent. The specific function of an agent is performed in the application layer. When the mobile host and the foreign agent are the same, the care-of address is called a co-located care-of address. To communicate with a remote host, a mobile host goes through three phases: agent discovery, registration, and data transfer.

Agent Discovery

A mobile node has to find a foreign agent when it moves away from its home network. To solve this problem, mobile IP describes two methods: agent advertisement and agent solicitation.

Agent advertisement

For this method, foreign agents and home agents advertise their presence periodically using special **agent advertisement** messages, which are broadcast into the subnet. Mobile IP does not use a new packet type for agent advertisement; it uses the router advertisement packet of ICMP, and appends an agent advertisement message.

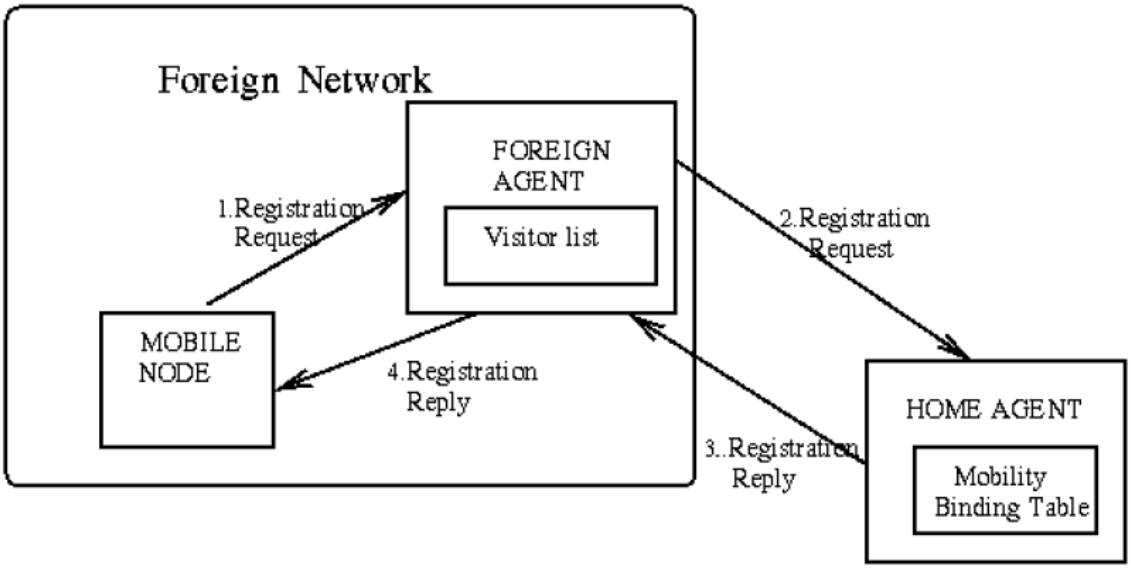
Agent Solicitation

If no agent advertisements are present or the inter-arrival time is too high, and an MN has not received a COA by other means, the mobile node must send **agent solicitations**. Care must be taken to ensure that these solicitation messages do not flood the network, but basically an MN can search for an FA endlessly sending out solicitation messages. If a node does not receive an answer to its solicitations it must decrease the rate of solicitations exponentially to avoid flooding the network until it reaches a maximum interval between solicitations (typically one minute). Discovering a new agent can be done anytime, not just if the MN is not connected to one.

After these steps of advertisements or solicitations the MN can now receive a COA, either one for an FA or a co-located COA.

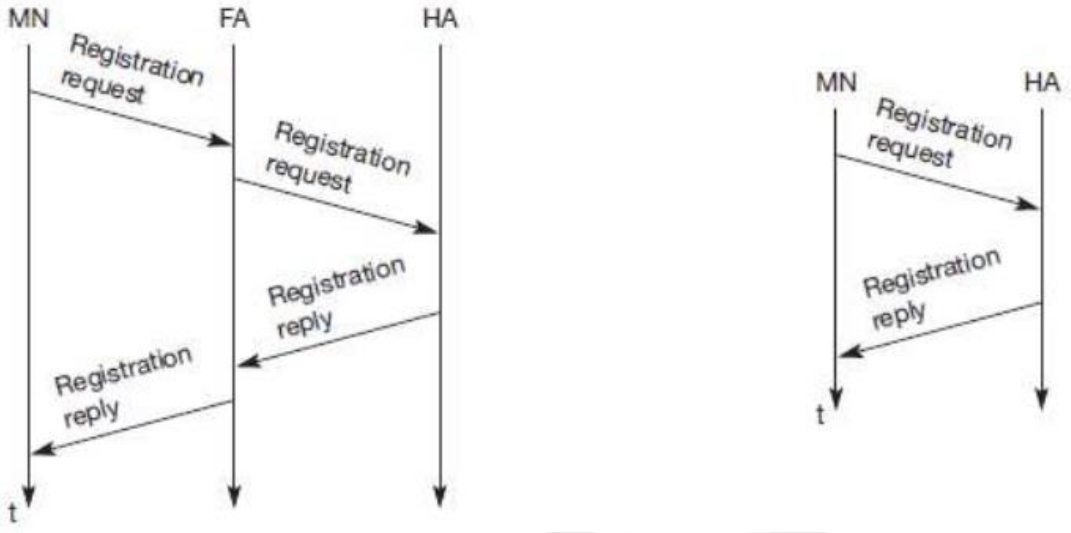
Agent Registration

Having received a COA, the MN has to register with the HA. The main purpose of the registration is to inform the HA of the current location for correct forwarding of packets.



Registration can be done in two different ways depending on the location of the COA. If the COA is at the FA, the MN sends its registration request containing the COA to the FA which forwards the request to the HA. The HA now sets up a **mobility binding,** containing the mobile node's home IP address and the current COA. It also contains the lifetime of the registration which is negotiated during the registration process. Registration expires automatically after the lifetime and is deleted; so, an MN should reregister before expiration.

This mechanism is necessary to avoid mobility bindings which are no longer used. After setting up the mobility binding, the HA sends a reply message back to the FA which forwards it to the MN.



Registration of a mobile node via the FA or directly with the HA

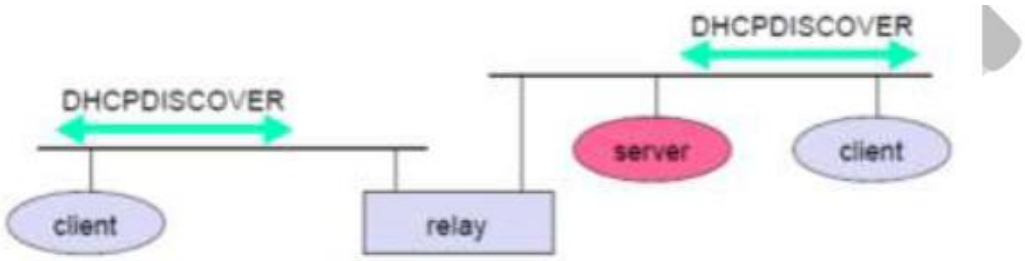
If the COA is co-located, registration can be simpler, the MN sends the request directly to the HA and vice versa. This is also the registration procedure for MNs returning to their home network to register directly with the HA.
UDP packets are used for the registration requests using the port no 434. The IP source address of the packet is set to the interface address of the MN, the IP destination address is that of the FA or HA.

IPv6
The design of Mobile IP support in IPv6 (Mobile IPv6) benefits both from the experiences gained from the development of Mobile IP support in IPv4, and from the opportunities provided by IPv6. Mobile IPv6 thus shares many features with Mobile IPv4, but is integrated into IPv6 and offers many other improvements. This section summarizes the major differences between Mobile IPv4 and Mobile IPv6:
There is no need to deploy special routers as "foreign agents", as in Mobile IPv4. Mobile IPv6 operates in any location without any special support required from the local router.
Support for route optimization is a fundamental part of the protocol, rather than a nonstandard set of extensions.
Mobile IPv6 route optimization can operate securely even without pre-arranged security associations. It is expected that route optimization can be deployed on a global scale between all mobile nodes and correspondent nodes.
Support is also integrated into Mobile IPv6 for allowing route optimization to coexist efficiently with routers that perform "ingress filtering"
The IPv6 Neighbor Unreachability Detection assures symmetric reachability between the mobile node and its default router in the current location.
Most packets sent to a mobile node while away from home in Mobile IPv6 are sent using an IPv6 routing header rather than IP encapsulation, reducing the amount of resulting overhead compared to Mobile IPv4.

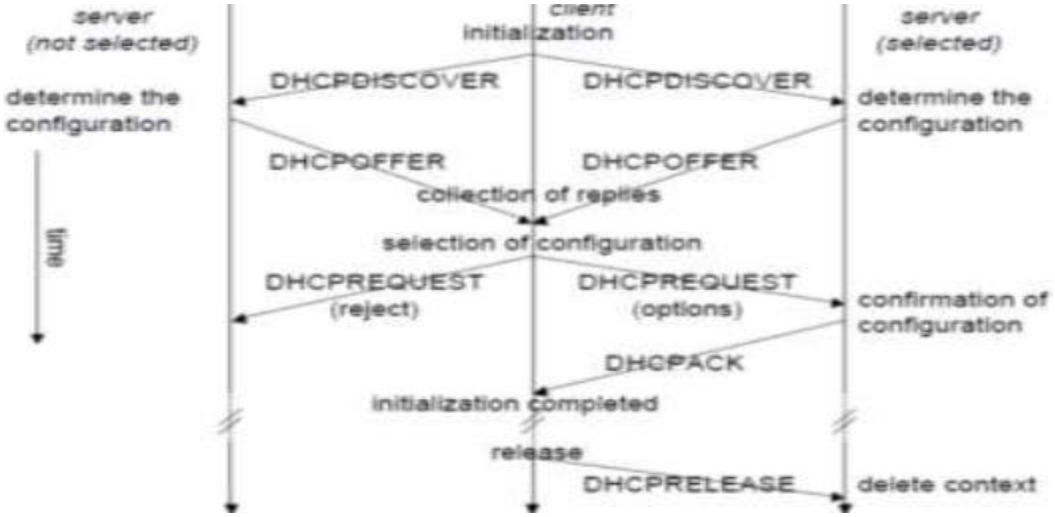
Mobile IPv6 is decoupled from any particular link layer, as it uses IPv6 Neighbor Discovery instead of ARP. This also improves the robustness of the protocol.
The use of IPv6 encapsulation (and the routing header) removes the need in Mobile IPv6 to manage "tunnel soft state".
The dynamic home agent address discovery mechanism in Mobile IPv6 returns a single reply to the mobile node. The directed broadcast approach used in IPv4 returns separate replies from each home agent.

Dynamic Host Configuration Protocol (DHCP)

DHCP is an automatic configuration protocol used on IP networks. **DHCP** allows a computer to join an IP-based network without having a pre-configured IP address. DHCP is a protocol that assigns unique IP addresses to devices, then releases and renews these addresses as devices leave and re-join the network. If a new computer is connected to a network, DHCP can provide it with all the necessary information for full system integration into the network, e.g., addresses of a DNS server and the default router, the subnet mask, the domain name, and an IP address. Providing an IP address makes DHCP very attractive for mobile IP as a source of careof-addresses.



DHCP is based on a client/server model as shown below. DHCP clients send a request to a server (DHCPDISCOVER in the example) to which the server responds. A client sends requests using MAC broadcasts to reach all devices in the LAN. A DHCP relay might be needed to forward requests across inter-working units to a DHCP server.
Consider the scenario where there is one client and two servers are present. A typical initialization of a DHCP client is shown below:



the client broadcasts a DHCPDISCOVER into the subnet. There might be a relay to forward this broadcast. In the case shown, two servers receive this broadcast and determine the configuration they can offer to the client. Servers reply to the client's request with DHCPOFFER and offer a list of configuration parameters. The client can now choose one of the configurations offered.

The client in turn replies to the servers, accepting one of the configurations and rejecting the others using DHCPREQUEST. If a server receives a DHCPREQUEST with a rejection, it can free the reserved configuration for other possible clients. The server with the configuration accepted by the client now confirms the configuration with DHCPACK. This completes the initialization phase. If a client leaves a subnet, it should release the configuration received by the server using DHCPRELEASE. Now the server can free the context stored for the client and offer the configuration again. The configuration a client gets from a server is only leased for a certain amount of time, it has to be reconfirmed from time to time. Otherwise the server will free the configuration. This timeout of configuration helps in the case of crashed nodes or nodes moved away without releasing the context.

DHCP is a good candidate for supporting the acquisition of care-of addresses for mobile nodes. The same holds for all other parameters needed, such as addresses of the default router, DNS servers, the timeserver etc. A DHCP server should be located in the subnet of the access point of the mobile node, or at least a DHCP relay should provide forwarding of the messages. RFC 3118 specifies authentication for DHCP messages so as to provide protection from malicious DHCP servers. Without authentication, a DHCP server cannot trust the mobile node and vice versa.

The **Transmission Control Protocol** (**TCP**) is one of the core protocols of the Internet protocol suite, often simply referred to as TCP/IP. TCP is reliable, guarantees in-order delivery of data and incorporates congestion control and flow control mechanisms.
TCP supports many of the Internet's most popular application protocols and resulting applications, including the World Wide Web, e-mail, File Transfer Protocol and Secure Shell. In the Internet protocol suite, TCP is the intermediate layer between the Internet layer and application layer.
The major responsibilities of TCP in an active session are to:
• **Provide reliable in-order transport of data**: to not allow losses of data.
• **Control congestions in the networks**: to not allow degradation of the network performance,
• **Control a packet flow between the transmitter and the receiver**: to not exceed the receiver's capacity.
TCP uses a number of mechanisms to achieve high performance and avoid 'congestion collapse', where network performance can fall by several orders of magnitude. These mechanisms control the rate of data entering the network, keeping the data flow below a rate that would trigger collapse. There are several mechanisms of TCP that influence the efficiency of TCP in a mobile environment. Acknowledgments for data sent, or lack of acknowledgments, are used by senders to implicitly interpret network conditions between the TCP sender and receiver.
Congestion Control
A transport layer protocol such as TCP has been designed for fixed networks with fixed end-systems. Congestion may appear from time to time even in carefully designed networks. The packet buffers of a router are filled and the router cannot forward the packets fast enough because the sum of the input rates of packets destined for one output link is higher than the capacity of the output link. The only thing a router can do in this situation is to drop packets.

A dropped packet is lost for the transmission, and the receiver notices a gap in the packet stream. Now the receiver does not directly tell the sender which packet is missing, but continues to acknowledge all in-sequence packets up to the missing one. The sender notices the missing acknowledgement for the lost packet and assumes a packet loss due to congestion. Retransmitting the missing packet and continuing at full sending rate would now be unwise, as this might only increase the congestion. To mitigate congestion, TCP slows down the transmission rate dramatically. All other TCP connections experiencing the same congestion do exactly the same so the congestion is soon resolved. Slow start TCP's reaction to a missing acknowledgement is quite drastic, but it is necessary to get rid of congestion quickly.
The behavior TCP shows after the detection of congestion is called **slow start.** The sender always calculates a **congestion window** for a receiver. The start size of the congestion window is one segment (TCP packet). The sender sends one packet and waits for acknowledgement. If this acknowledgement arrives, the sender increases the congestion window by one, now sending two packets (congestion window = 2). This scheme doubles the congestion window every time the

acknowledgements come back, which takes one round trip time (RTT). This is called the exponential growth of the congestion window in the slow start mechanism.

But doubling the congestion window is too dangerous. The exponential growth stops at the **congestion threshold**. As soon as the congestion window reaches the congestion threshold, further increase of the transmission rate is only linear by adding 1 to the congestion window each time the acknowledgements come back. Linear increase continues until a time-out at the sender occurs due to a missing acknowledgement, or until the sender detects a gap in transmitted data because of continuous acknowledgements for the same packet. In either case the sender sets the congestion threshold to half of the current congestion window. The congestion window itself is set to one segment and the sender starts sending a single segment. The exponential growth starts once more up to the new congestion threshold, then the window grows in linear fashion.

Fast Retransmit/Fast Recovery

The congestion threshold can be reduced because of two reasons. First one is if the sender receives continuous acknowledgements for the same packet. It informs the sender that the receiver has got all the packets upto the acknowledged packet in the sequence and also the receiver is receiving something continuously from the sender. The gap in the packet stream is not due to congestion, but a simple packet loss due to a transmission error. The sender can now retransmit the missing packet(s) before the timer expires. This behavior is called **fast retransmit**. It is an early enhancement for preventing slow-start to trigger on losses not caused by congestion. The receipt of acknowledgements shows that there is no congestion to justify a slow start. The sender can continue with the current congestion window. The sender performs a **fast recovery** from the packet loss. This mechanism can improve the efficiency of TCP dramatically. The other reason for activating slow start is a time-out due to a missing acknowledgement. TCP using fast retransmit/fast recovery interprets this congestion in the network and activates the slow start mechanism. The advantage of this method is its simplicity. Minor changes in the MH's software results in performance increase. No changes are required in FA or CH. The disadvantage of this scheme is insufficient isolation of packet losses. It mainly focuses on problems regarding Handover. Also it effects the efficiency when a CH transmits already delivered packets.

Mobile TCP

Both I-TCP and Snooping TCP does not help much, if a mobile host gets disconnected.

The **M-TCP (mobile TCP)** approach has the same goals as I-TCP and snooping TCP: to prevent the sender window from shrinking if bit errors or disconnection but not congestion cause current problems. M-TCP wants to improve overall throughput, to lower the delay, to maintain end to end semantics of TCP, and to provide a more efficient handover. Additionally, M-TCP is especially adapted to the problems arising from lengthy or frequent disconnections. M-TCP splits the TCP connection into two parts as I-TCP does. An unmodified TCP is used on the standard host-**supervisory host (SH)** connection, while an optimized TCP is used on the SH-MH connection.

The SH monitors all packets sent to the MH and ACKs returned from the MH. If the SH does not receive an ACK for some time, it assumes that the MH is disconnected. It then chokes the sender by setting the sender's window size to 0. Setting the window size to 0 forces the sender to go into **persistent mode**, i.e., the state of the sender will not change no matter how long the receiver is

disconnected. This means that the sender will not try to retransmit data. As soon as the SH (either the old SH or a new SH) detects connectivity again, it reopens the window of the sender to the old value. The sender can continue sending at full speed. This mechanism does not require changes to the sender's TCP. The wireless side uses an adapted TCP that can recover from packet loss much faster. This modified TCP does not use slow start, thus, M-TCP needs a **bandwidth manager** to implement fair sharing over the wireless link.

Advantages of M-TCP:
It maintains the TCP end-to-end semantics. The SH does not send any ACK itself butforwards the ACKs from the MH.
If the MH is disconnected, it avoids useless retransmissions, slow starts or breakingconnections by simply shrinking the sender's window to 0.
As no buffering is done as in I-TCP, there is no need to forward buffers to a new SH.
Lost packets will be automatically retransmitted to the SH.
Disadvantages of M-TCP:
As the SH does not act as proxy as in I-TCP, packet loss on the wireless link due to bit errors is propagated to the sender. M-TCP assumes low bit error rates, which is not always avalid assumption.
A modified TCP on the wireless link not only requires modifications to the MH protocol software but also new network elements like the bandwidth manager.

Transmission/time-out freezing
Often, MAC layer notices connection problems even before the connection is actually interrupted from a TCP point of view and also knows the real reason for the interruption. The MAC layer can inform the TCP layer of an upcoming loss of connection or that the current interruption is not caused by congestion. TCP can now stop sending and 'freezes' the current state of its congestion window and further timers. If the MAC layer notices the upcoming interruption early enough, both the mobile and correspondent host can be informed. With a fast interruption of the wireless link, additional mechanisms in the access point are needed to inform the correspondent host of the reason for interruption. Otherwise, the correspondent host goes into slow start assuming congestion and finally breaks the connection.

As soon as the MAC layer detects connectivity again, it signals TCP that it can resumeoperation at exactly the same point where it had been forced to stop. For TCP time simply doesnot advance, so no timers expire.
Advantages:
It offers a way to resume TCP connections even after long interruptions of the connection.
It can be used together with encrypted data as it is independent of other TCP mechanisms such as sequence no or acknowledgements **Disadvantages:**
Lots of changes have to be made in software of MH, CH and FA.
Selective retransmission
A very useful extension of TCP is the use of selective retransmission. TCP acknowledgements are cumulative, i.e., they acknowledge in-order receipt of packets up to a certain packet. A single acknowledgement confirms reception of all packets upto a certain packet. If a single packet is lost, the sender has to retransmit everything starting from the lost packet (go-back-n retransmission). This obviously wastes bandwidth, not just in the case of a mobile network, but

for any network. Using selective retransmission, TCP can indirectly request a selective retransmission of packets. The receiver can acknowledge single packets, not only trains of in-sequence packets. The sender can now determine precisely which packet is needed and can retransmit it. The **advantage** of this approach is obvious: a sender retransmits only the lost packets. This lowers bandwidth requirements and is extremely helpful in slow wireless links. The disadvantage is that a more complex software on the receiver side is needed. Also more buffer space is needed to resequence data and to wait for gaps to be filled.

Transaction-oriented TCP

Assume an application running on the mobile host that sends a short request to a server from time to time, which responds with a short message and it requires reliable TCP transport of the packets. For it to use normal TCP, it is inefficient because of the overhead involved. Standard TCP is made up of three phases: setup, data transfer and release. First, TCP uses a three-way handshake to establish the connection. At least one additional packet is usually needed for transmission of the request, and requires three more packets to close the connection via a three way handshake. So, for sending one data packet, TCP may need seven packets altogether. This kind of overhead is acceptable for long sessions in fixed networks, but is quite inefficient for short messages or sessions in wireless networks. This led to the development of transaction oriented
TCP (T/TCP).

T/TCP can combine packets for connection establishment and connection release with user data packets. This can reduce the number of packets down to two instead of seven. The obvious **advantage** for certain applications is the reduction in the overhead which standard TCP has for connection setup and connection release. Disadvantage is that it requires changes in the software in mobile host and all correspondent hosts. This solution does not hide mobility anymore. Also, T/TCP exhibits several security problems.

UNIT 5

GSM

GSM is the most successful digital mobile telecommunication system in the world today. It is used by over 800 million people in more than 190 countries. GSM permits the integration of different voice and data services and the interworking with existing networks. Services make a network interesting for customers. GSM has defined three different categories of services:
Bearer Services, Tele and **Supplementary Services**.

Bearer services:

GSM specifies different mechanisms for data transmission, the original GSM allowing for data rates of up to 9600 bit/s for non-voice services. Bearer services permit transparent and non-transparent, synchronous or asynchronous data transmission.

Transparent bearer services only use the functions of the physical layer (layer 1) to transmit data. Data transmission has a constant delay and throughput if no transmission errors occur. Transmission quality can be improved with the use of **forward error correction (FEC)**, which codes redundancy into the data stream and helps to reconstruct the original data in case of

transmission errors. Transparent bearer services do not try to recover lost data in case of, for example, shadowing or interruptions due to handover. **Non-transparent bearer services** use protocols of layers two and three to implement error correction and flow control. These services use the transparent bearer services, adding a **radio link protocol (RLP)**. This protocol comprises mechanisms of **high-level data link control (HDLC)**, and special selective-reject mechanisms to trigger retransmission of erroneous data. Using transparent and non-transparent services, GSM specifies several bearer services for interworking with PSTN, ISDN, and packet switched public data networks (PSPDN) like X.25, which is available worldwide. Data transmission can be fullduplex, synchronous with data rates of 1.2, 2.4, 4.8, and 9.6 kbit/s or full-duplex, asynchronous from 300 to 9,600 bit/s.

Tele services: GSM mainly focuses on voice-oriented tele services. These comprise encrypted voice transmission, message services, and basic data communication with terminals as known from the PSTN or ISDN (e.g., fax). The primary goal of GSM was the provision of high-quality digital voice transmission. Special codecs (coder/decoder) are used for voice transmission, while other codecs are used for the transmission of analog data for communication with traditional computer modems used in, e.g., fax machines. Another service offered by GSM is the

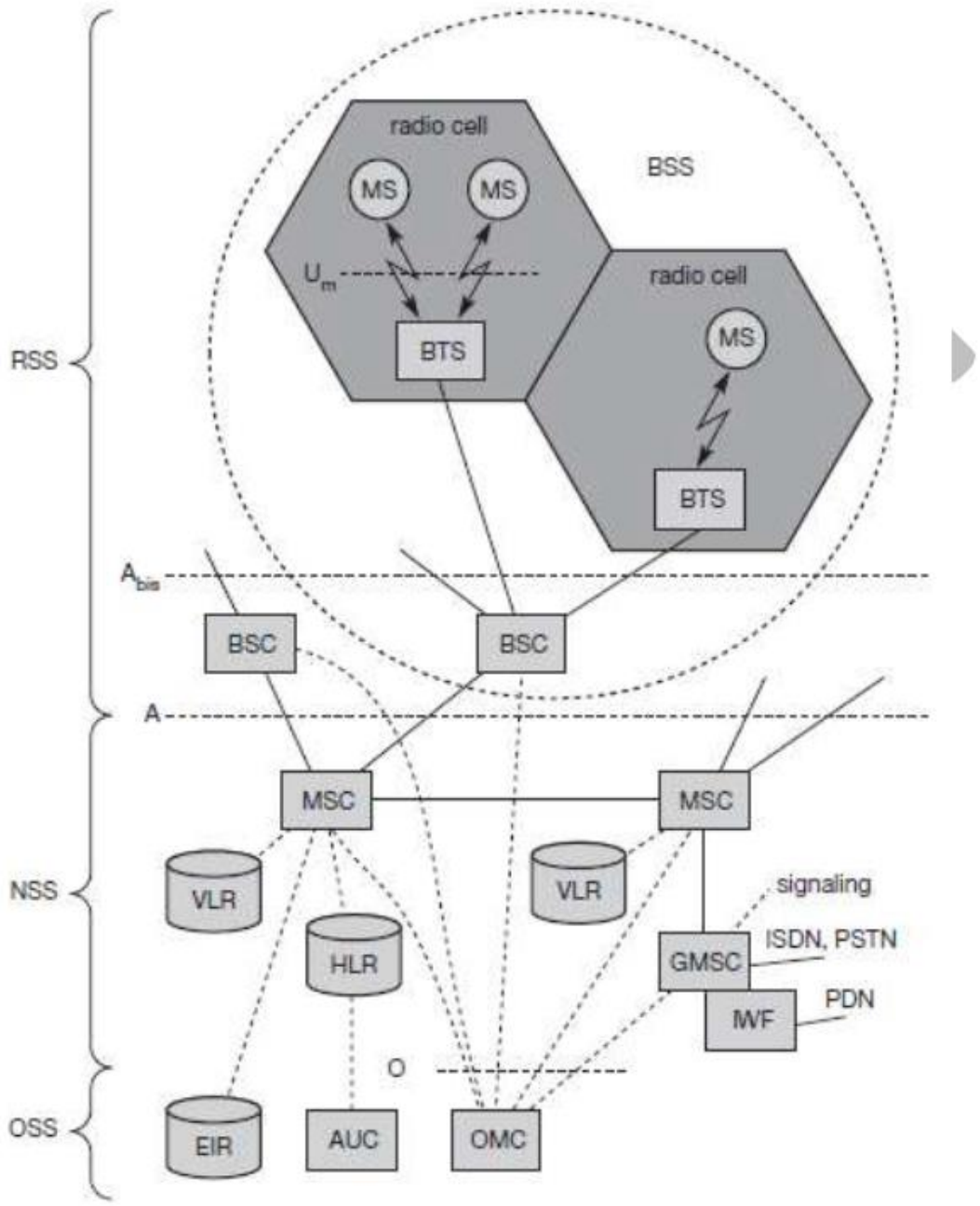
emergency number (eg 911, 999). This service is mandatory for all providers and free of charge. This connection also has the highest priority, possibly pre-empting other connections, and will automatically be set up with the closest emergency center. A useful service for very simple message transfer is the **short message service (SMS)**, which offers transmission of messages of up to 160 characters. Sending and receiving of SMS is possible during data or voice transmission. It can be used for "serious" applications such as displaying road conditions, e-mail headers or stock quotes, but it can also transfer logos, ring tones, horoscopes and love letters. The successor of SMS, the **enhanced message service (EMS)**, offers a larger message size, formatted text, and the transmission of animated pictures, small images and ring tones in a standardized way. But with MMS, EMS was hardly used. MMS offers the transmission of larger pictures (GIF, JPG, WBMP), short video clips etc. and comes with mobile phones that integrate small cameras. Another non-voice tele service is **group 3 fax**, which is available worldwide. In this service, fax data is transmitted as digital data over the analog telephone network according to the ITU-T standards T.4 and T.30 using modems.

Supplementary services:

In addition to tele and bearer services, GSM providers can offer **supplementary services**. these services offer various enhancements for the standard telephony service, and may vary from provider to provider. Typical services are user **identification**, call **redirection**, or **forwarding** of ongoing calls, barring of incoming/outgoing calls, Advice of Charge (AoC) etc. Standard ISDN features such as **closed user groups** and **multiparty** communication may be available.

GSM Architecture

A GSM system consists of three subsystems, the radio sub system (RSS), the network and switching subsystem (NSS), and the operation subsystem (OSS).



Functional Architecture of a GSM System

Network Switching Subsystem: The NSS is responsible for performing call processing andsubscriber related functions. The switching system includes the following functional units:
Home location register (HLR): It is a database used for storage and management of subscriptions. HLR stores permanent data about subscribers, including a subscribers service profile, location information and activity status. When an individual buys a subscription from the PCS provider, he or she is registered in the HLR of that operator.

Visitor location register (VLR): It is a database that contains temporary information about subscribers that is needed by the MSC in order to service visiting subscribers. VLR is always

integrated with the MSC. When a MS roams into a new MSC area, the VLR connected to that MSC will request data about the mobile station from the HLR. Later if the mobile station needs to make a call, VLR will be having all the information needed for call setup. Authentication center (AUC): A unit called the AUC provides authentication and encryption parameters that verify the users identity and ensure the confidentiality of each call.

Equipment identity register (EIR): It is a database that contains information about the identity of mobile equipment that prevents calls from stolen, unauthorized or defective mobile stations.

Mobile switching center (MSC): The MSC performs the telephony switching functions of the system. It controls calls to and from other telephone and data systems.

Radio Subsystem (RSS):
The radio **subsystem (RSS)** comprises all radio specific entities, i.e., the mobile **stations (MS)** and the **base station subsystem (BSS)**. The figure shows the connection between the RSS and the NSS via the **A interface** (solid lines) and the connection to the OSS via the **O interface** (dashed lines).
Base station subsystem (BSS): A GSM network comprises many BSSs, each controlled by a base station controller (BSC). The BSS performs all functions necessary to maintain radio connections to an MS, coding/decoding of voice, and rate adaptation to/from the wireless network part. Besides a BSC, the BSS contains several BTSs.

Base station controllers (BSC): The BSC provides all the control functions and physical links between the MSC and BTS. It is a high capacity switch that provides functions such as handover, cell configuration data, and control of radio frequency (RF) power levels in BTS. A number of BSC's are served by and MSC.

Base transceiver station (BTS): The BTS handles the radio interface to the mobile station. A BTS can form a radio cell or, using sectorized antennas, several and is connected to MS via the **Um interface**, and to the BSC via the **Abis interface**. The Um interface contains all the mechanisms necessary for wireless transmission (TDMA, FDMA etc.)The BTS is the radio equipment (transceivers and antennas) needed to service each cell in the network. A group of BTS's are controlled by an BSC.

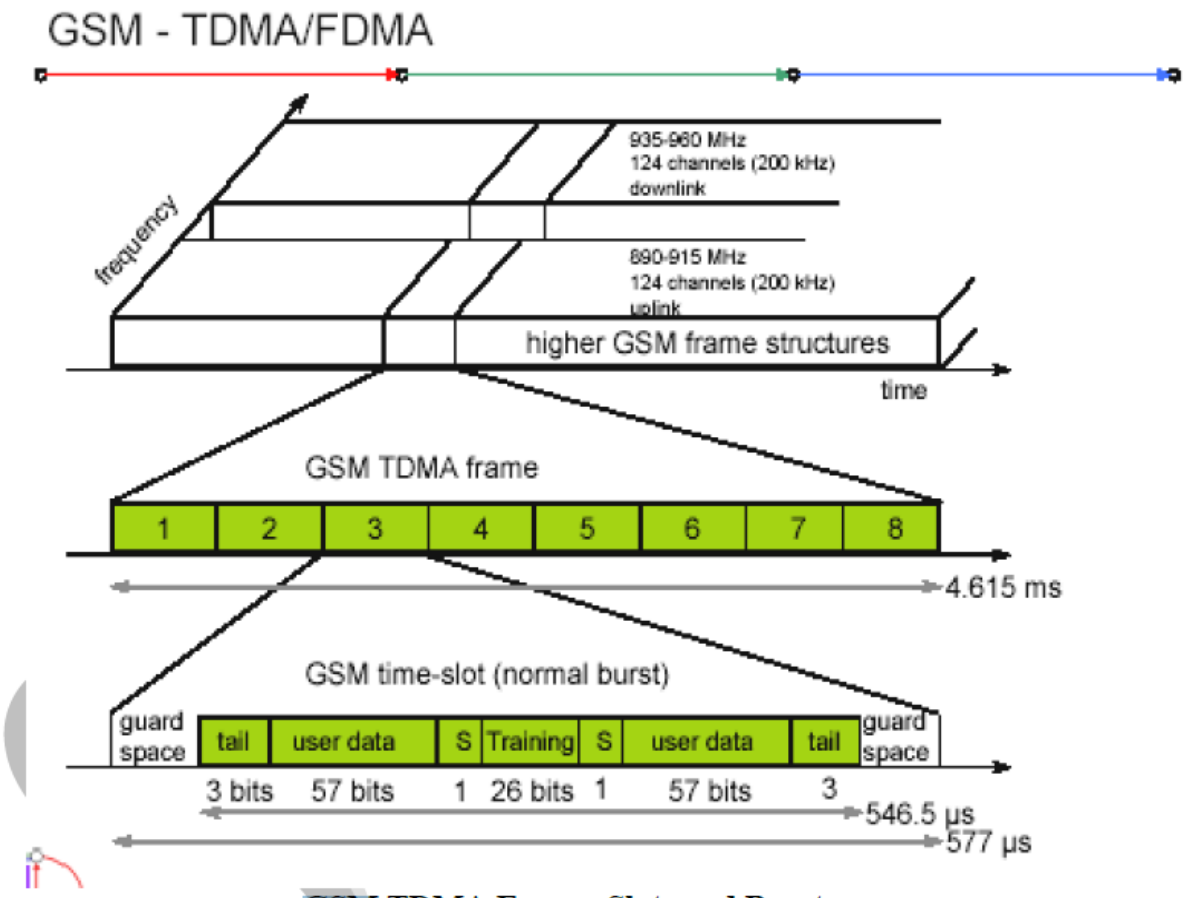
Operation and Support system:
The operations and maintenance center (OMC) is connected to all equipment in the switching system and to the BSC. Implementation of OMC is called operation and support system (OSS). The OSS is the functional entity from which the network operator monitors and controls the system. The purpose of OSS is to offer the customer cost-effective support for centralized, regional and local operational and maintenance activities that are required for a GSM network. OSS provides a network overview and allows engineers to monitor, diagnose and troubleshoot every aspect of the GSM network.

The mobile station (MS) consists of the mobile equipment (the terminal) and a smart card called the Subscriber Identity Module (SIM). The SIM provides personal mobility, so that the user can have access to subscribed services irrespective of a specific terminal. By inserting the SIM card

into another GSM terminal, the user is able to receive calls at that terminal, make calls from that terminal, and receive other subscribed services. The mobile equipment is uniquely identified by the International Mobile Equipment Identity (IMEI). The SIM card contains the International Mobile Subscriber Identity (IMSI) used to identify the subscriber to the system, a secret key for authentication, and other information. The IMEI and the IMSI are independent, thereby allowing personal mobility. The SIM card may be protected against unauthorized use by a password or personal identity number.

Radio Interface

The most interesting interface in a GSM system is Um, the radio interface, as it comprises various multiplexing and media access mechanisms. GSM implements SDMA using cells with BTS and assigns an MS to a BTS.



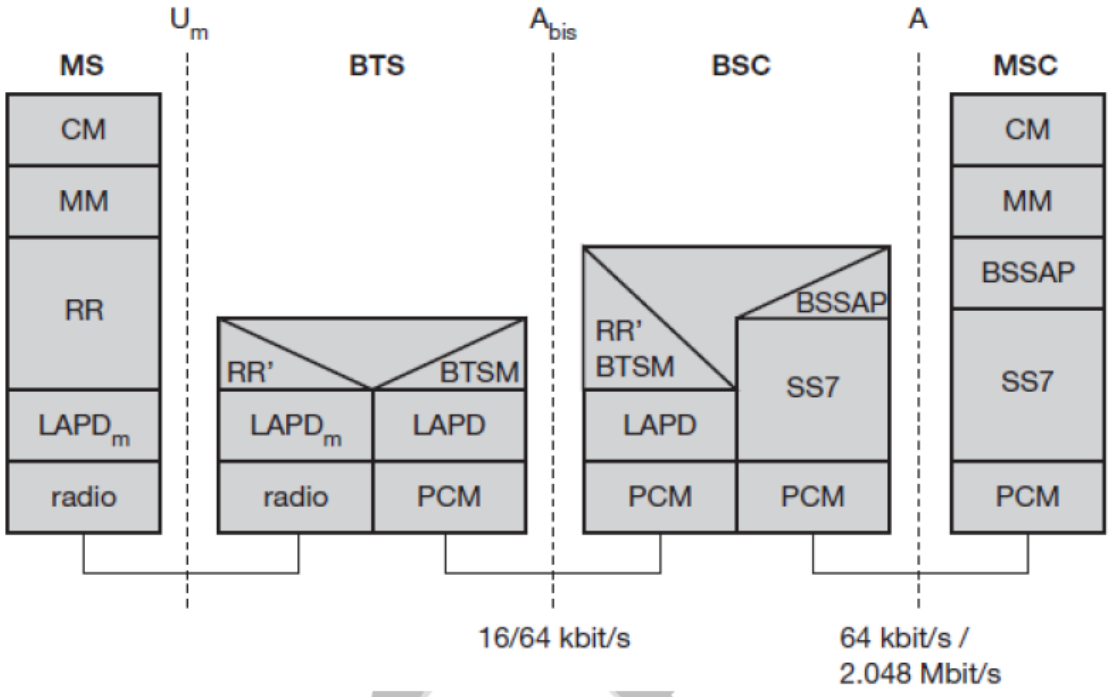
GSM TDMA Frame, Slots and Bursts

Each of the 248 channels is additionally separated in time via a **GSM TDMA frame**, i.e., each 200 kHz carrier is subdivided into frames that are repeated continuously. The duration of a frame is 4.615 ms. A frame is again subdivided into 8 **GSM time slots**, where each slot represents a physical TDM channel and lasts for 577 μs. Each TDM channel occupies the 200 kHz carrier for 577 μs every 4.615 ms. Data is transmitted in small portions, called **bursts**. The following figure shows a so called **normal burst** as used for data transmission inside a time slot.

As shown, the burst is only 546.5 μs long and contains 148 bits. The remaining 30.5 μs are used as **guard space** to avoid overlapping with other bursts due to different path delays and to give

the transmitter time to turn on and off. The first and last three bits of a normal burst (**tail**) are all set to 0 and can be used to enhance the receiver performance. The **training** sequence in the middle of a slot is used to adapt the parameters of the receiver to the current path propagation characteristics and to select the strongest signal in case of multi-path propagation. A flag **S** indicates whether the **data** field contains user or network control data.
Apart from the normal burst, ETSI (1993a) defines four more bursts for data transmission: a **frequency correction** burst allows the MS to correct the local oscillator to avoid interference with neighbouring channels, a **synchronization burst** with an extended training sequence synchronizes the MS with the BTS in time, an **access burst** is used for the initial connection setup between MS and BTS, and finally a **dummy burst** is used if no data is available for a slot.

GSM Protocols
The signalling protocol in GSM is structured into three general layers depending on the interface, as shown below. Layer 1 is the physical layer that handles all **radio**-specific functions. This includes the creation of bursts according to the five different formats, **multiplexing** of bursts into a TDMA frame, **synchronization** with the BTS, detection of idle channels, and measurement of the **channel quality** on the downlink. The physical layer at Um uses GMSK for digital **modulation** and performs **encryption/decryption** of data, i.e., encryption is not performed end-to-end, but only between MS and BSS over the air interface.



Protocol Architecture for Signaling

The main tasks of the physical layer comprise **channel coding** and **error detection/correction**, which is directly combined with the coding mechanisms. Channel coding makes extensive use of different **forward error correction (FEC)** schemes.
Signaling between entities in a GSM network requires higher layers. For this purpose, the **LAPDm** protocol has been defined at the Um interface for **layer two**. LAPDm has been derived

from link access procedure for the D-channel (**LAPD**) in ISDN systems, which is a version of HDLC. LAPDm is a lightweight LAPD because it does not need synchronization flags or checksumming for error detection. LAPDm offers reliable data transfer over connections, resequencing of data frames, and flow control.

The network layer in GSM, layer three, comprises several sublayers. The lowest sublayer is the radio resource management (RR). Only a part of this layer, RR', is implemented in the BTS, the remainder is situated in the BSC. The functions of RR' are supported by the BSC via the BTS management (BTSM). The main tasks of RR are setup, maintenance, and release of radio channels. Mobility management (MM) contains functions for registration, authentication, identification, location updating, and the provision of a temporary mobile subscriber identity (TMSI). Finally, the call management (CM) layer contains three entities: call control (CC), short message service (SMS), and supplementary service (SS). SMS allows for message transfer using the control channels SDCCH and SACCH, while SS offers the services like user identification, call redirection, or forwarding of ongoing calls. CC provides a point-to-point connection between two terminals and is used by higher layers for call establishment, call clearing and change of call parameters. This layer also provides functions to send in-band tones, called dual tone multiple frequency (DTMF), over the GSM network. These tones are used, e.g., for the remote control of answering machines or the entry of PINs in electronic banking and are, also used for dialing in traditional analog telephone systems. Additional protocols are used at the Abis and A interfaces.

Data transmission at the physical layer typically uses **pulse code modulation (PCM)** systems. LAPD is used for layer two at Abis, BTSM for BTS management. **Signaling system No. 7 (SS7)** is used for signaling between an MSC and a BSC. This protocol also transfers all management information between MSCs, HLR, VLRs, AuC, EIR, and OMC. An MSC can also control a BSS via a **BSS application part (BSSAP)**.

Localization and Calling

The fundamental feature of the GSM system is the automatic, worldwide localization of users for which, the system performs periodic location updates. The HLR always contains information about the current location and the VLR currently responsible for the MS informs the HLR about the location changes. Changing VLRs with uninterrupted availability is called roaming. Roaming can take place within a network of one provider, between two providers in a country and also between different providers in different countries.

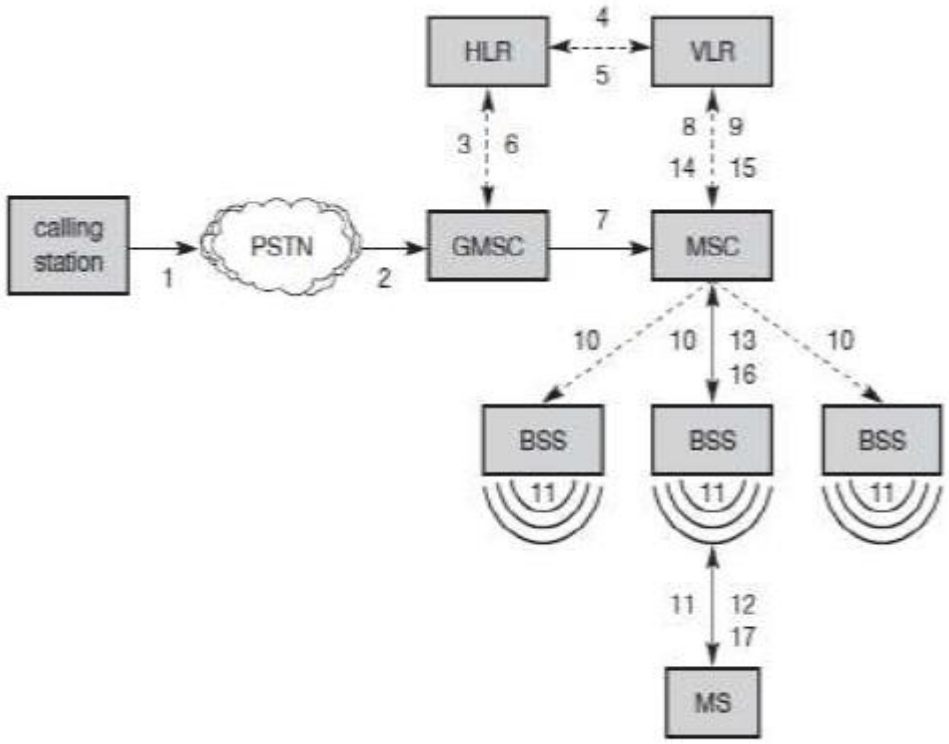
To locate and address an MS, several numbers are needed:

Mobile station international ISDN number (MSISDN):- The only important number for a user of GSM is the phone number. This number consists of the country code (CC), the national destination code (NDC) and the subscriber number (SN).

International mobile subscriber identity (IMSI): GSM uses the IMSI for internal unique identification of a subscriber. IMSI consists of a mobile country code (MCC), the mobile network code (MNC), and finally the mobile subscriber identification number (MSIN).

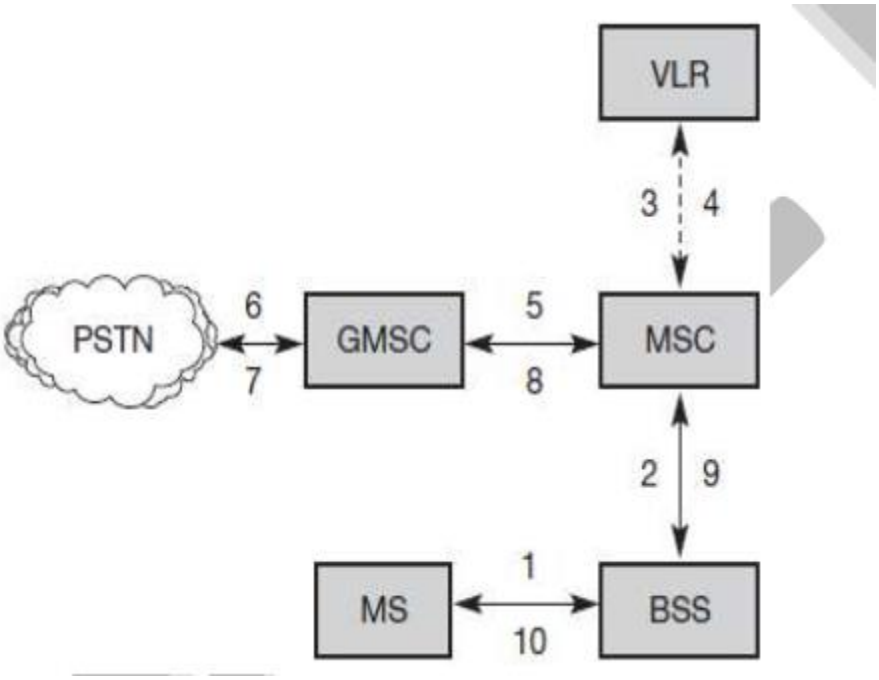
Temporary mobile subscriber identity (TMSI): To hide the IMSI, which would give away the exact identity of the user signalling over the air interface, GSM uses the 4 byte TMSI for local subscriber identification.

Mobile station roaming number (MSRN): Another temporary address that hides the identity and location of a subscriber is MSRN. The VLR generates this address on request from the MSC, and the address is also stored in the HLR. MSRN contains the current visitor country code (VCC), the visitor national destination code (VNDC), the identification of the current MSC together with the subscriber number. The MSRN helps the HLR to find a subscriber for an incoming call. For *a mobile terminated call (MTC),* the following figure shows the different steps that take place:

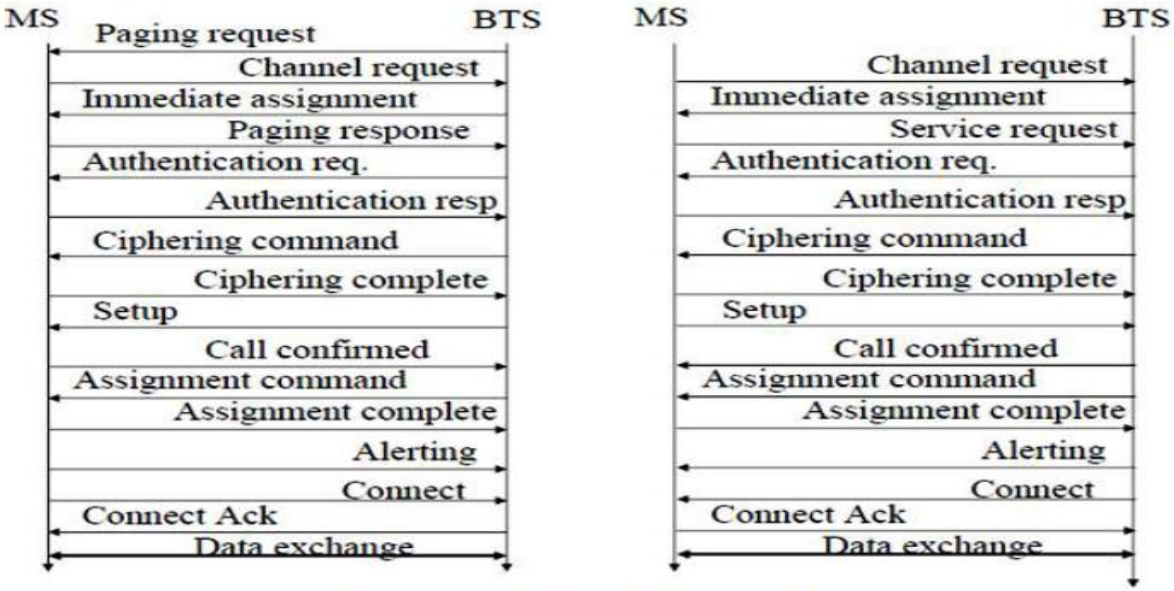


Mobile Terminated Call (MTC)

step 1: User dials the phone number of a GSM subscriber.
step 2: The fixed network (PSTN) identifies the number belongs to a user in GSM network and forwards the call setup to the Gateway MSC (GMSC).
step 3: The GMSC identifies the HLR for the subscriber and signals the call setup to HLR
step 4: The HLR checks for number existence and its subscribed services and requests an MSRN from the current VLR.
step 5: VLR sends the MSRN to HLR
step 6: Upon receiving MSRN, the HLR determines the MSC responsible for MS and forwards the information to the GMSC
step 7: The GMSC can now forward the call setup request to the MSC indicated
step 8: The MSC requests the VLR for the current status of the MS
step 9: VLR sends the requested information
step 10: If MS is available, the MSC initiates paging in all cells it is responsible for.
step 11: The BTSs of all BSSs transmit the paging signal to the MS
step 12: Step 13: If MS answers, VLR performs security checks
step 15: Till step 17: Then the VLR signals to the MSC to setup a connection to the MS. For a mobile originated call (MOC), the following steps take place:



step 1: The MS transmits a request for a new connection
step 2: The BSS forwards this request to the MSC
step 3: Step 4: The MSC then checks if this user is allowed to set up a call with the requested and checks the availability of resources through the GSM network and into the PSTN. If all resources are available, the MSC sets up a connection between the MS and the fixed network. In addition to the steps mentioned above, other messages are exchanged between an MS and BTS during connection setup (in either direction).



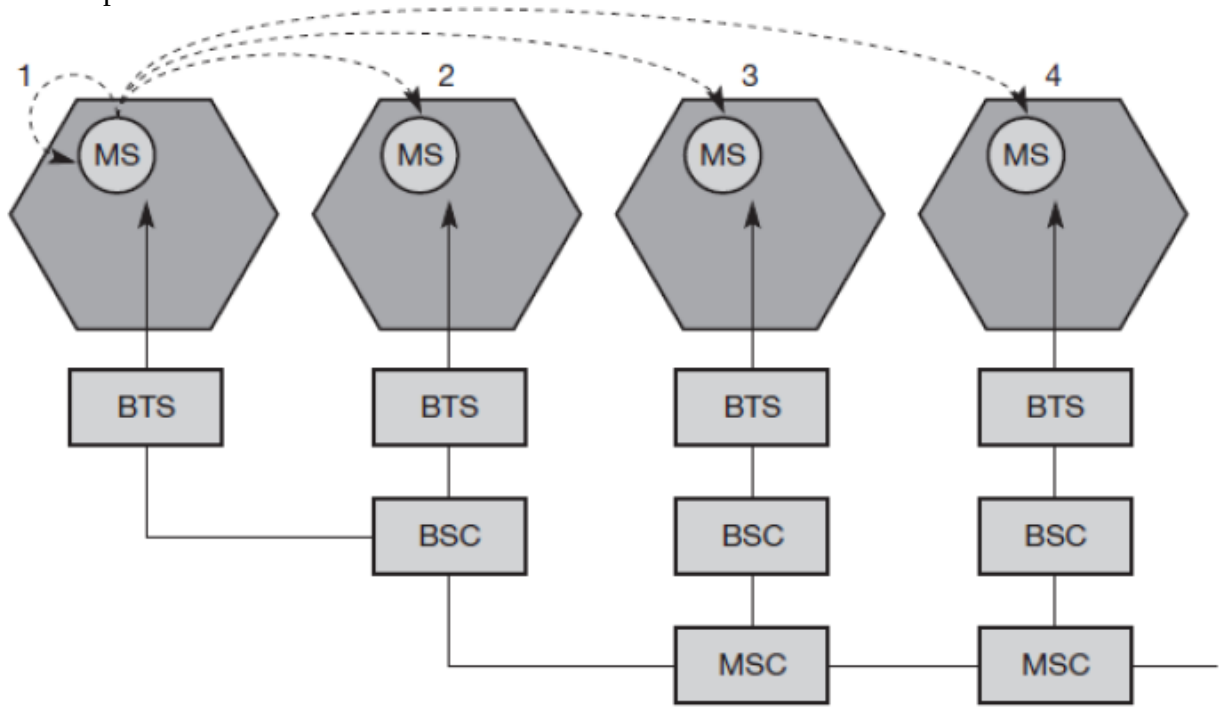
Message flow for MTC and MOC

HANDOVER

Cellular systems require **handover** procedures, as single cells do not cover the whole service area. However, a handover should not cause a cut-off, also called **call drop**. GSM aims at maximum handover duration of 60 ms. There are two basic reasons for a handover:

1. The mobile station **moves out of the range** of a BTS, decreasing the received **signal level** increasing the **error rate** thereby diminishing the **quality of the radio link.**
2. Handover may be due to **load balancing,** when an MSC/BSC decides the traffic is too high in one cell and shifts some MS to other cells with a lower load.

The four possible handover scenarios of GSM are shown below:



Intra-cell handover: Within a cell, narrow-band interference could make transmission at a certain frequency impossible. The BSC could then decide to change the carrier frequency (scenario 1).
Inter-cell, intra-BSC handover: This is a typical handover scenario. The mobile station moves from one cell to another, but stays within the control of the same BSC. The BSC then performs a handover, assigns a new radio channel in the new cell and releases the old one (scenario 2).
Inter-BSC, intra-MSC handover: As a BSC only controls a limited number of cells; GSM also has to perform handovers between cells controlled by different BSCs. This handover then has to be controlled by the MSC (scenario 3).
Inter MSC handover: A handover could be required between two cells belonging to different MSCs. Now both MSCs perform the handover together (scenario 4). To provide all the necessary information for a handover due to a weak link, MS and BTS both perform periodic measurements of the downlink and uplink quality respectively.

Measurement reports are sent by the MS about every half-second and contain the quality of the current link used for transmission as well as the quality of certain channels in neighboring cells (the BCCHs).

Security

GSM offers several security services using confidential information stored in the AuC and in the individual SIM. The SIM stores personal, secret data and is protected with a PIN against unauthorized use. Three algorithms have been specified to provide security services in GSM. **Algorithm A3** is used for **authentication**, **A5** for **encryption**, and **A8** for the **generation of a cipher key**. The various security services offered by GSM are:

Access control and authentication: The first step includes the authentication of a valid user for the SIM. The user needs a secret PIN to access the SIM. The next step is the subscriber authentication.

Authentication is based on the SIM, which stores the **individual authentication key Ki**, the **user identification IMSI**, and the algorithm used for authentication **A3**. The AuC performs the basic generation of random values RAND, signed responses SRES, and cipher keys Kc for each IMSI, and then forwards this information to the HLR. The current VLR requests the appropriate values for RAND, SRES, and **Kc** from the HLR. For authentication, the VLR sends the random value RAND to the SIM. Both sides, network and subscriber module, perform the same operation with RAND and the key **Ki**, called **A3**. The MS sends back the SRES generated by the SIM; the VLR can now compare both values. If they are the same, the VLR accepts the subscriber, otherwise the subscriber is rejected.

Confidentiality:

All user-related data is encrypted.

To ensure privacy, all messages containing user-related information are encrypted in GSM over the air interface. After authentication, MS and BSS can start using encryption by applying the cipher key **Kc**, which is generated using the individual key Ki and a random value by applying the algorithm A8. Note that the SIM in the MS and the network both calculate the same **Kc** based on the random value RAND. The key Kc itself is not transmitted over the air interface. MS and BTS can now encrypt and decrypt data using the algorithm A5 and the cipher key Kc.

Anonymity:

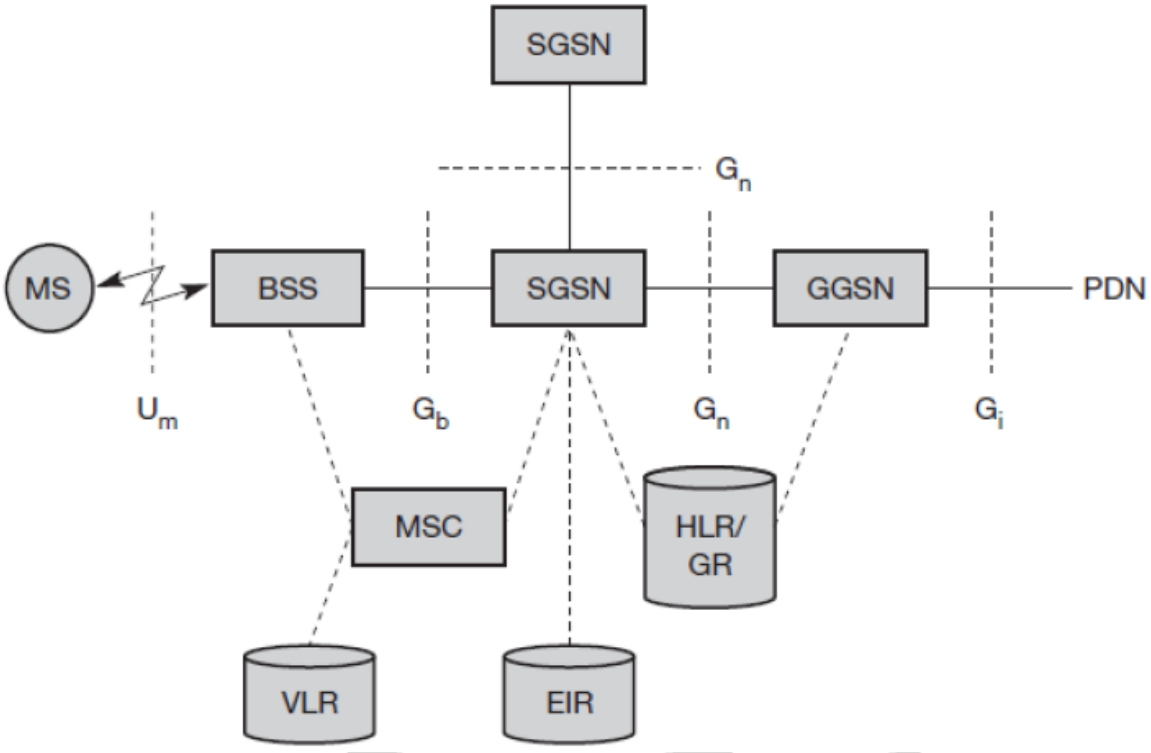
To provide user anonymity, all data is encrypted before transmission, and user identifiers are not used over the air. Instead, GSM transmits a temporary identifier (TMSI), which is newly assigned by the VLR after each location update. Additionally, the VLR can change the TMSI at any time.

GPRS:

The next step toward more flexible and powerful data transmission avoids the problems of HSCSD by being fully packet-oriented. The **general packet radio service (GPRS)** provides packet mode transfer for applications that exhibit traffic patterns such as frequent transmission of small volumes (e.g., typical web requests) or infrequent transmissions of small or medium volumes (e.g., typical web responses) according to the requirement specification. For the new GPRS radio channels, the GSM system can allocate between one and eight time slots within a TDMA frame. Time slots are not allocated in a fixed, pre-determined manner but on demand. All

time slots can be shared by the active users; up- and downlink are allocated separately. Allocation of the slots is based on current load and operator preferences. The GPRS concept is independent of channel characteristics and of the type of channel (traditional GSM traffic or control channel), and does not limit the maximum data rate (only the GSM transport system limits the rate). All GPRS services can be used in parallel to conventional services. GPRS includes several **security services** such as authentication, access control, user identity confidentiality, and user information confidentiality.

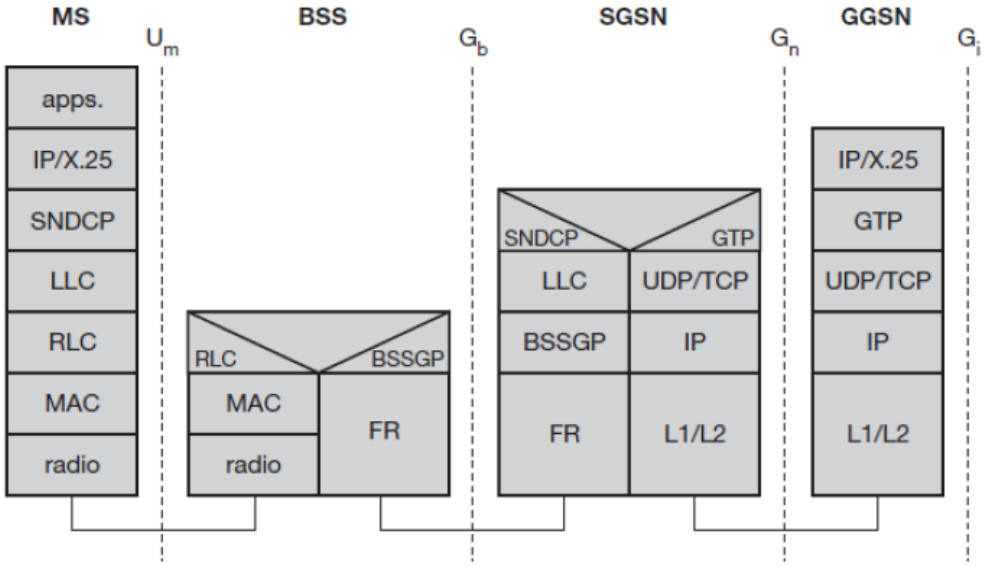
The GPRS architecture introduces two new network elements, which are called GPRS support nodes (GSN) and are in fact routers. All GSNs are integrated into the standard GSM architecture, and many new interfaces have been defined. The gateway GPRS support node (GGSN) is the interworking unit between the GPRS network and external packet data networks (PDN). This node contains routing information for GPRS users, performs address conversion, and tunnels data to a user via encapsulation. The GGSN is connected to external networks (e.g., IP or X.25) via the Gi interface and transfers packets to the SGSN via an IPbased GPRS backbone network (Gn interface). The other new element is the **serving GPRS support node (SGSN)** which supports the MS via the Gb interface. The SGSN, for example, requests user addresses from the **GPRS register (GR)**, keeps track of the individual MSs' location, is responsible for collecting billing information (e.g., counting bytes), and performs several security functions such as access control. The SGSN is connected to a BSC via frame relay and is basically on the same hierarchy level as an MSC. The GR, which is typically a part of the
HLR, stores all GPRS-relevant data.



GPRS Architecture Reference Model

As shown above, packet data is transmitted from a PDN, via the GGSN and SGSN directly to the BSS and finally to the MS. The MSC, which is responsible for data transport in the traditional circuit-switched GSM, is only used for signaling in the GPRS scenario. Before sending any data over the GPRS network, an MS must attach to it, following the procedures of the **mobility**

management. The attachment procedure includes assigning a temporal identifier, called a **temporary logical link identity (TLLI)**, and a **ciphering key sequence number (CKSN)** for data encryption. For each MS, a **GPRS context** is set up and stored in the MS and in the corresponding SGSN. Besides attaching and detaching, mobility management also comprises functions for authentication, location management, and ciphering.
The following figure shows the protocol architecture of the transmission plane for GPRS.
All data within the GPRS backbone, i.e., between the GSNs, is transferred using the **GPRS tunnelling protocol (GTP)**. GTP can use two different transport protocols, either the reliable **TCP** (needed for reliable transfer of X.25 packets) or the non-reliable **UDP** (used for IP packets). The network protocol for the GPRS backbone is **IP** (using any lower layers). To adapt to the different characteristics of the underlying networks, the **subnetwork dependent convergence protocol (SNDCP)** is used between an SGSN and the MS. On top of SNDCP and GTP, user packet data is tunneled from the MS to the GGSN and vice versa. To achieve a high reliability of packet transfer between SGSN and MS, a special LLC is used, which comprises ARQ and FEC mechanisms for PTP (and later PTM) services.



GPRS Transmission Plane Protocol Reference Model

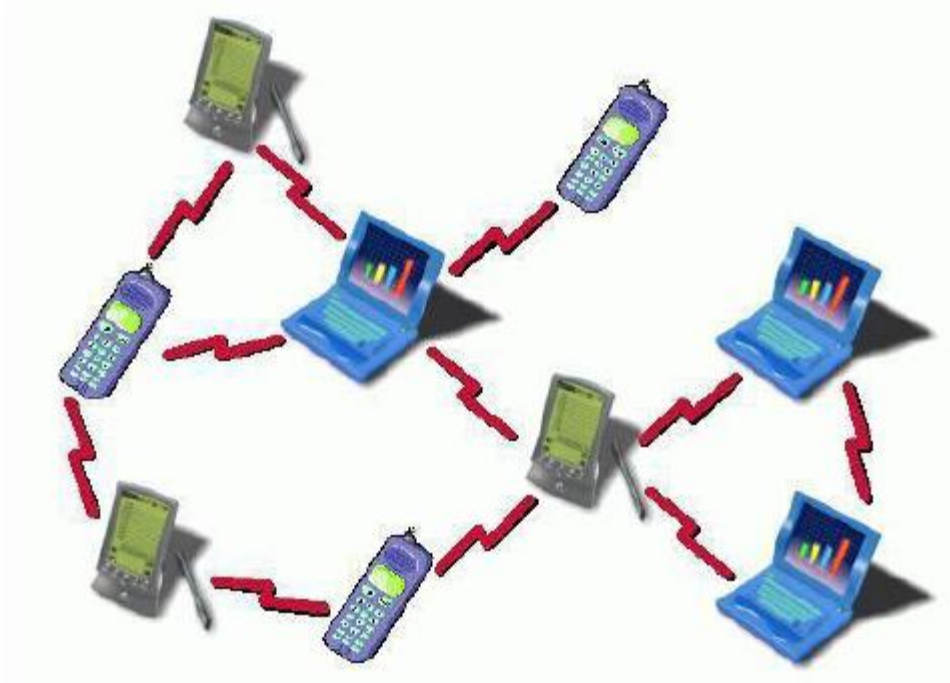
A base station subsystem GPRS protocol (BSSGP) is used to convey routing and QoSrelated information between the BSS and SGSN. BSSGP does not perform error correction and works on top of a frame relay (FR) network. Finally, radio link dependent protocols are needed to transfer data over the Um interface. The radio link protocol (RLC) provides a reliable link, while the MAC controls access with signalling procedures for the radio channel and the mapping of LLC frames onto the GSM physical channels. The radio interface at Um needed for GPRS does not require fundamental changes compared to standard GSM.

UNIT 6

HISTORICAL DEVELOPMENTS OF MANET In early 1970s, the Mobile Ad hoc Network (MANET) was called packet radio network, which was sponsored by Defense Advanced Research Projects Agency (DARPA). They had a project named packet radio having several wireless terminals that could communication with each other on battlefields. “It is interesting to note that these early packet radio systems predict the Internet and indeed were part of the motivation of the original Internet Protocol suite” . The whole life cycle of Ad hoc networks could be categorized into the first, second, and the third generation Ad hoc networks systems. Present Ad hoc networks systems are considered the third generation . The first generation goes back to 1972. At the time, they were called PRNET (Packet Radio Networks). In conjunction with ALOHA (Arial Locations of Hazardous Atmospheres) and CSMA (Carrier Sense Medium Access), approaches for medium access control and a kind of distance-vector routing PRNET were used on a trial basis to provide different networking capabilities in a combat environment. The second generation of Ad hoc networks emerged in 1980s, when the Ad hoc network systems were further enhanced and implemented as a part of the SURAN (Survivable Adaptive Radio Networks) program. This provided a packet-switched network to the mobile battlefield in an environment without infrastructure. This On-Demand Routing In Multi-Hop Wireless Mobile Ad hoc Networks Overview of Mobile Ad hoc Networks 20 program proved to be beneficial in improving the radios' performance by making them smaller, cheaper, and resilient to electronic attacks. In the 1990s (Third generation), the concept of commercial Ad hoc networks arrived with notebook computers and other viable communication equipments. At the same time, the idea of a collection of mobile nodes was proposed at several researchers gatherings. The IEEE 802.11 subcommittee had adopted the term "Ad hoc networks" and the research community had started to look into the possibility of deploying Ad hoc networks in other areas of application

BASIC CONCEPTS OF MOBILE AD HOC NETWORKS

An Ad hoc network is a collection of mobile nodes, which forms a temporary network without the aid of centralized administration or standard support devices regularly available as conventional networks. These nodes generally have a limited transmission range and, so, each node seeks the assistance of its neighboring nodes in forwarding packets and hence the nodes in an Ad hoc network can act as both routers and hosts. Thus a node may forward packets between other nodes as well as run user applications. By nature these types of networks are suitable for situations where either no fixed infrastructure exists or deploying network is not possible. Ad hoc mobile networks have found many applications in various fields like military, emergency, conferencing and sensor networks. Each of these application areas has their specific requirements for routing protocols. Since the network nodes are mobile, an Ad hoc network will typically have a dynamic topology, which will have profound effects on network characteristics. Network nodes will often be battery powered, which limits the capacity of CPU, memory, and bandwidth. This will require network functions that are resource effective. Furthermore, the wireless (radio) media will also affect the behavior of the network due to fluctuating link bandwidths resulting from relatively high error rates. These unique desirable features pose several new challenges in the design of wireless Ad hoc networking protocols.



Network functions such as routing, address allocation, authentication and authorization must be designed to cope with a dynamic and volatile On-Demand Routing In Multi-Hop Wireless Mobile Ad hoc Networks Overview of Mobile Ad hoc Networks 21 network topology. In order to establish routes between nodes, which are farther than a single hop, specially configured routing protocols are engaged. The unique feature of these protocols is their ability to trace routes in spite of a dynamic topology. In the simplest scenarios, nodes may be able to communicate directly with each other, for example, when they are within wireless transmission range of each other. However, Ad hoc networks must also support communication between nodes that are only indirectly connected by a series of wireless hops through other nodes. For example, to establish communication between nodes A and C the network must enlist the aid of node B to relay packets between them. The circles indicate the nominal range of each node's radio transceiver. Nodes A and C are not in direct transmission range of each other, since A's circle does not cover C. A Mobil Ad hoc network of three nodes, where nodes A and C must discover the route through B in order to communicate. In general, an Ad hoc network is a network in which every node is potentially a router and every node is potentially mobile. The presence of wireless communication and mobility make an Ad hoc network unlike a traditional wired network and requires that the routing protocols used in an Ad hoc network be based on new and different principles. Routing protocols for traditional wired networks are designed to support tremendous numbers of nodes, but they assume that the relative position of the nodes will generally remain unchanged.

Characteristics of MANET

In MANET, each node act as both host and router. That is it is autonomous in behavior. Multi-hop radio relaying- When a source node and destination node for a message is out of the radio range, the MANETs are capable of multi-hop routing.

Distributed nature of operation for security, routing and host configuration. A centralized firewall is absent here. The nodes can join or leave the network anytime, making the network topology dynamic in nature.
Mobile nodes are characterized with less memory, power and light weight features.
The reliability, efficiency, stability and capacity of wireless links are often inferior when compared with wired links. This shows the fluctuating link bandwidth of wireless links.
Mobile and spontaneous behavior which demands minimum human intervention to configure the network.
All nodes have identical features with similar responsibilities and capabilities and hence it forms a completely symmetric environment.
High user density and large level of user mobility.
Nodal connectivity is intermittent.

MANETs Applications ***1)Military battlefield:*** Ad-Hoc networking would allow the military to take advantage of commonplace network technology to maintain an information network between the soldiers, vehicles, and military information head quarter.
Collaborative work: For some business environments, the need for collaborative computing might be more important outside office environments than inside and where people do need to have outside meetings to cooperate and exchange information on a given project.

Local level: Ad-Hoc networks can autonomously link an instant and temporary multimedia network using notebook computers to spread and share information among participants at a e.g. conference or classroom. Another appropriate local level application might be in home networks where devices can communicate directly to exchange information.

Personal area network and bluetooth : A personal area network is a short range, localized network where nodes are usually associated with a given person. Short-range MANET such as Bluetooth can simplify the inter communication between various mobile devices such as a laptop, and a mobile phone.

Commercial Sector**:** Ad hoc can be used in emergency/rescue operations for disaster relief efforts, e.g. in fire, flood, or earthquake. Emergency rescue operations must take place where non-existing or damaged communications infrastructure and rapid deployment of a communication network is needed.

Emergency services
• Search and rescue operations

• Disaster recovery

• Replacement of fixed infrastructure in case of environmental disasters

• Policing and fire fighting

• Supporting doctors and nurses in hospitals

Commercial and civilian
• E-commerce: electronic payments anytime and anywhere environments
• Business: dynamic database access, mobile offices

• Vehicular services: road or accident guidance, transmission of road and weather conditions, taxi cab network, inter-vehicle networks
• Sports stadiums, trade fairs, shopping malls

• Networks of visitors at airports

Home and enterprise
• Home/office wireless networking networking
• Conferences, meeting rooms

• Personal area networks (PAN), Personal networks (PN)

• Networks at construction sites Education

• Universities and campus settings

• Virtual classrooms

• Ad hoc communications during meetings or lectures

Entertainment
• Multi-user games

•Wireless P2P networking

• Outdoor Internet access

• Robotic pets

• Theme parks

Sensor networks
• Home applications: smart sensors and actuators embedded in consumer electronics
• Body area networks (BAN)

• Data tracking of environmental conditions, animal movements, chemical/biological detection

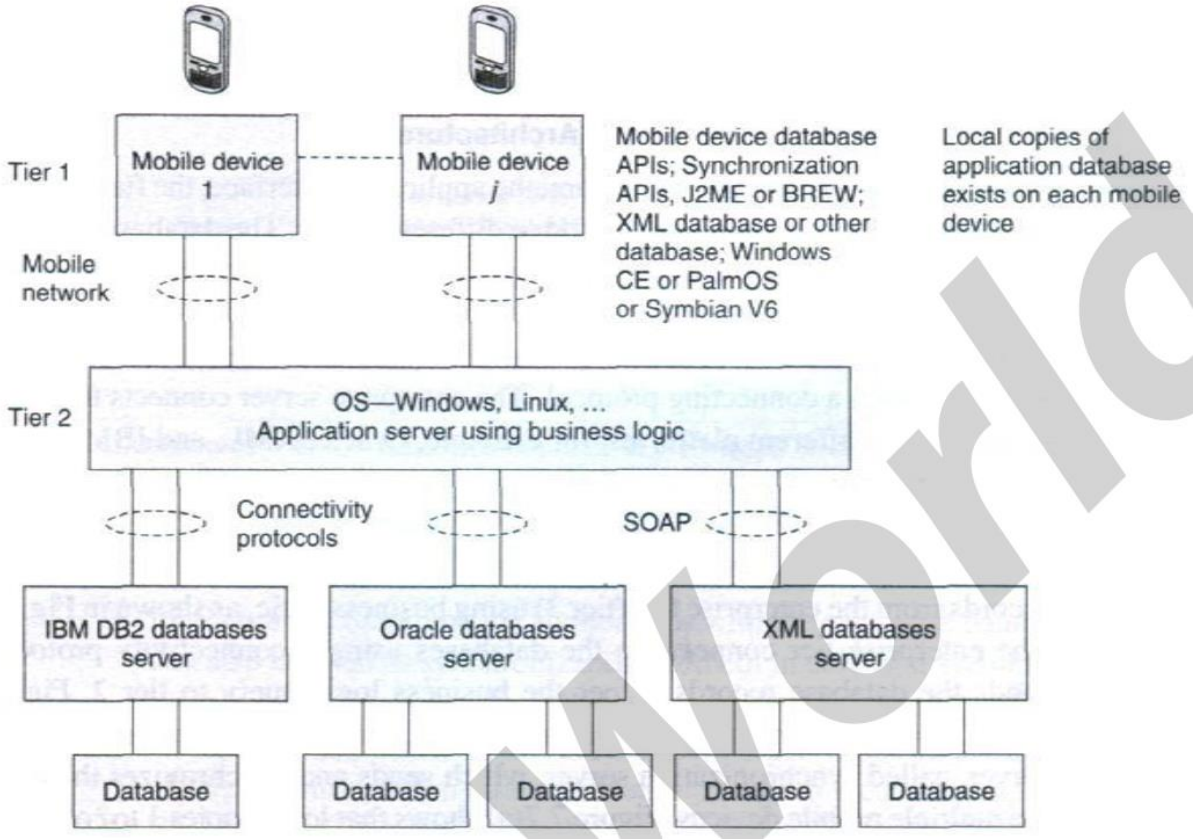
Context aware services
• Follow-on services: call-forwarding, mobile workspace

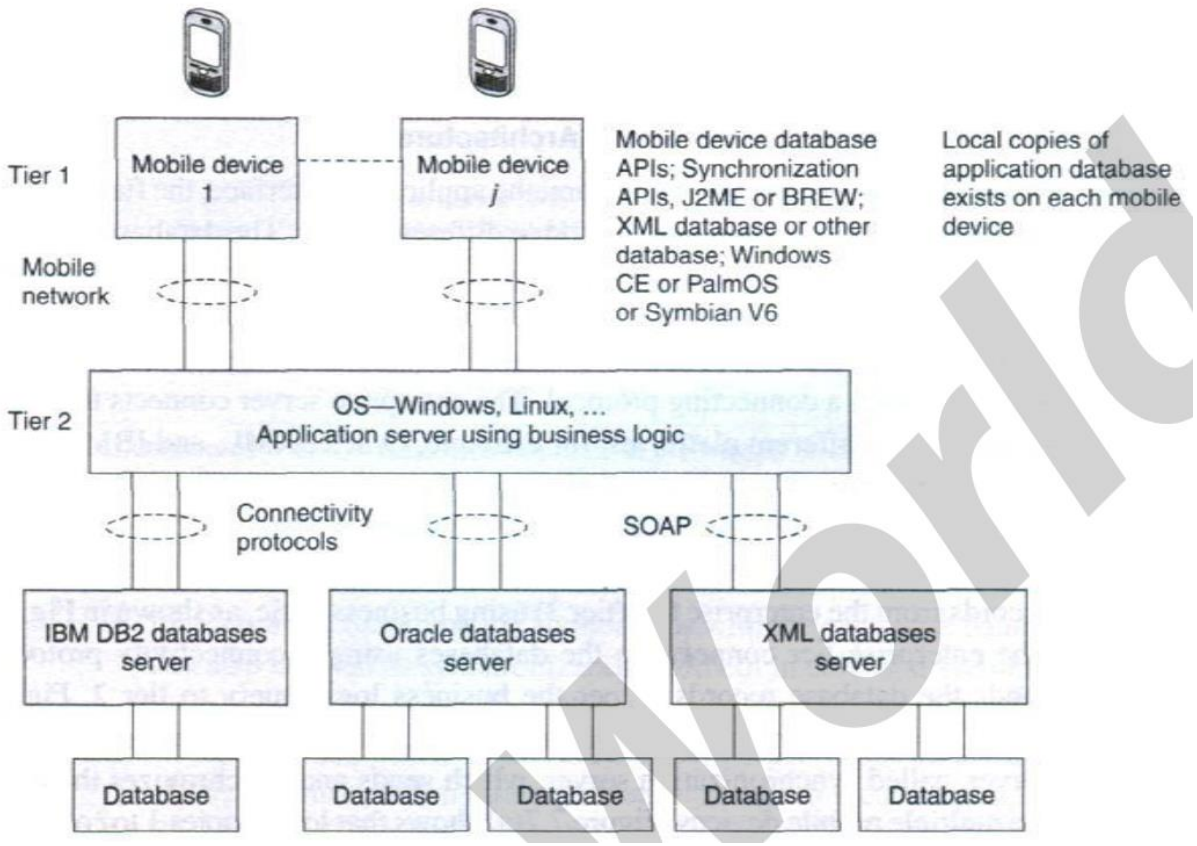
• Information services: location specific services, time dependent services

• Infotainment: touristic information

Coverage extension
• Extending cellular network access
• Linking up with the Internet, intranets, etc.

CLIENT SERVER ARCHITECTURE





In client server computing, the clients requests a resource and the server provides that resource. A server may serve multiple clients at the
same time while a client is in contact with only one server. Both the client and server usually communicate via a computer network but sometimes they may reside in the same system.

Characteristics of Client Server Computing

The salient points for client server computing are as follows:

- The client server computing works with a system of request and response. The client sends a request to the server and the server responds with the desired information.
- The client and server should follow a common communication protocol so they can easily interact with each other. All the communication protocols are available at the application layer.
- A server can only accommodate a limited number of client requests at a time. So it uses a system based to priority to respond to the requests.
- Denial of Service attacks hindera servers ability to respond to authentic client requests by inundating it with false requests.
- An example of a client server computing system is a web server. It returns the web pages to the clients that requested them.

Difference between Client Server Computing and Peer to Peer Computing

The major differences between client server computing and peer to peer computing are as follows:

- In client server computing, a server is a central node that services many client nodes. On the other hand, in a peer to peer system, the nodes collectively use their resources and communicate with each other.
- In client server computing the server is the one that communicates with the other nodes. In peer to peer to computing, all the nodes are equal and share data with each other directly.
- Client Server computing is believed to be a subcategory of the peer to peer computing.

Advantages of Client Server Computing

The different advantages of client server computing are:

- All the required data is concentrated in a single place i.e. the server. So it is easy to protect the data and provide authorisation and authentication.
- The server need not be located physically close to the clients. Yet the data can be accessed efficiently.
- It is easy to replace, upgrade or relocate the nodes in the client server model because all the nodes are independent and request data only from the server.
- All the nodes i.e clients and server may not be build on similar platforms yet they can easily facilitate the transfer of data.

Disadvantages of Client Server Computing

The different disadvantages of client server computing are:

- If all the clients simultaneously request data from the server, it may get overloaded. This may lead to congestion in the network.
- If the server fails for any reason, then none of the requests of the clients can be fulfilled. This leads of failure of the client server network.
- The cost of setting and maintaining a client server model are quite high.

WAP Architecture

WAP stands for **W**ireless **A**pplication **P**rotocol. The dictionary definition of these terms are as follows −

- **Wireless** − Lacking or not requiring a wire or wires pertaining to radio transmission.
- **Application** − A computer program or piece of computer software that is designed to do a specific task.
- **Protocol** − A set of technical rules about how information should be transmitted and received using computers.

WAP is the set of rules governing the transmission and reception of data by computer applications on or via wireless devices like mobile phones. WAP allows wireless devices to

view specifically designed pages from the Internet using only plain text and very simple black-and-white pictures.

WAP is a standardized technology for cross-platform, distributed computing very similar to the Internet's combination of Hypertext Markup Language (HTML) and Hypertext Transfer Protocol (HTTP), except that it is optimized for:

- low-display capability
- low-memory
- low-bandwidth devices, such as personal digital assistants (PDAs), wireless phones, and pagers.

WAP is designed to scale across a broad range of wireless networks like GSM, IS-95, IS-136, and PDC.

Who is behind WAP?

The Wireless Application Protocol (WAP) is a result of joint efforts taken by companies teaming up in an industry group called WAP Forum (www.wapforum.org).

On June 26, 1997, Ericsson, Motorola, Nokia, and Unwired Planet took the initiative to start a rapid creation of a standard for making advanced services within the wireless domain a reality. In December 1997, WAP Forum was formally created and after the release of the WAP 1.0 specifications in April 1998, WAP Forum membership was opened to all.

The WAP Forum now has over 500 members and represents over 95 percent of the global handset market. Companies such as Nokia, Motorola and Ericsson are all members of the forum.

The objective of the forum is to create a license-free standard that brings information and telephony services to wireless devices.

Why is WAP Important?

Until the first WAP devices emerged, the Internet was a Internet and a mobile phone was a mobile phone. You could surf the Net, do serious research, or be entertained on the Internet using your computer, but this was limited to your computer.

Now with the appearance of WAP, the scene is that we have the massive information, communication, and data resources of the Internet becoming more easily available to anyone with a mobile phone or communications device.

WAP being open and secure, is well suited for many different applications including, but not limited to stock market information, weather forecasts, enterprise data, and games.

Despite the common misconception, developing WAP applications requires only a few modifications to existing web applications. The current set of web application development tools will easily support WAP development, and in the future more development tools will be announced.

WAP Microbrowser

To browse a standard internet site you need a web browser. Similar way to browse a WAP enables website, you would need a micro browser. A Micro Browser is a small piece of software that makes minimal demands on hardware, memory and CPU. It can display information written in a restricted mark-up language called WML. Although, tiny in memory footprint it supports many features and is even scriptable.

Today, all the WAP enabled mobile phones or PDAs are equipped with these micro browsers so that you can take full advantage of WAP technology.

A programming model similar to the Internet's

Though WAP is a new technology, but it reuse the concepts found on the Internet. This reuse enables a quick introduction of WAP-based services, since both service developers and manufacturers are familiar with these concepts today.

Wireless Markup Language (WML)

You must be using HTML language to develop your web-based application. Same way, WML is a markup language used for authoring WAP services, fulfilling the same purpose as HTML does on the Web. In contrast to HTML, WML is designed to fit small handheld devices.

WMLScript

Once again, you must be using Java Script or VB script to enhance the functionality of your web applications. Same way, WMLScript can be used to enhance the functionality of a service, just as Java script can be utilized in HTML. It makes it possible to add procedural logic and computational functions to WAPbased services.

Wireless Telephony Application Interface (WTAI)

The WTAI is an application framework for telephony services. WTAI user agents are able to make calls and edit the phone book by calling special WMLScript functions or by accessing special URLs. If one writes WML decks containing names of people and their phone numbers, you may add them to your phone book or call them right away just by clicking the appropriate hyperlink on the screen.

Optimized protocol stack

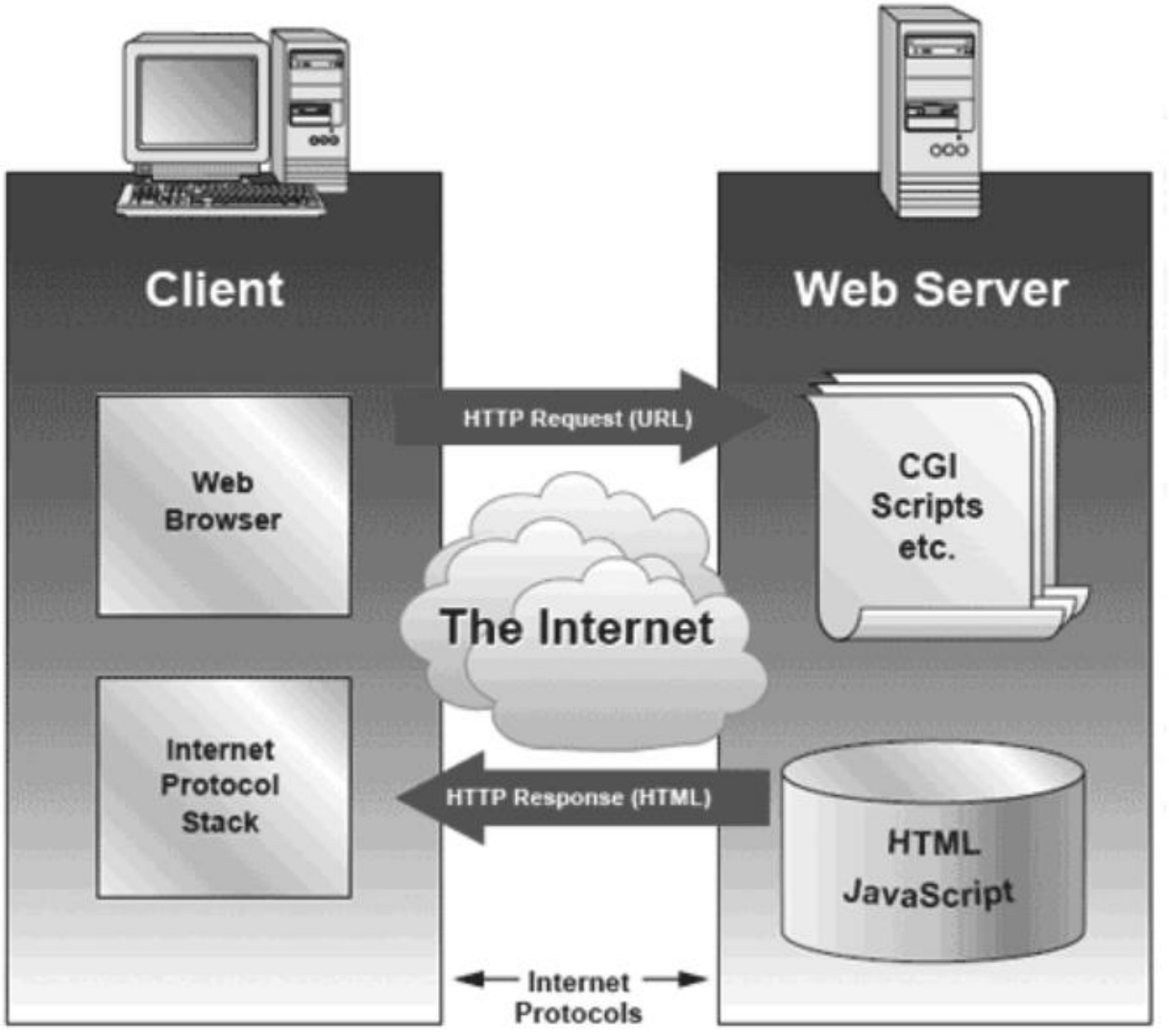
The protocols used in WAP are based on well-known Internet protocols, such as **HTTP** and **Transmission Control Protocol** (TCP), but they have been optimized to address the constraints of a wireless environment, such as low bandwidth and high latency.

The Internet Model

The Internet model makes it possible for a client to reach services on a large number of origin servers, each addressed by a **unique Uniform Resource Locator** (URL).

The content stored on the servers is of various formats, but HTML is the predominant. HTML provides the content developer with a means to describe the appearance of a service in a flat document structure. If more advanced features like procedural logic are needed, then scripting languages such as JavaScript or VB Script may be utilised.

The figure below shows how a WWW client request a resource stored on a web server. On the Internet standard communication protocols, like HTTP and Transmission Control Protocol/Internet Protocol (TCP/IP) are used.

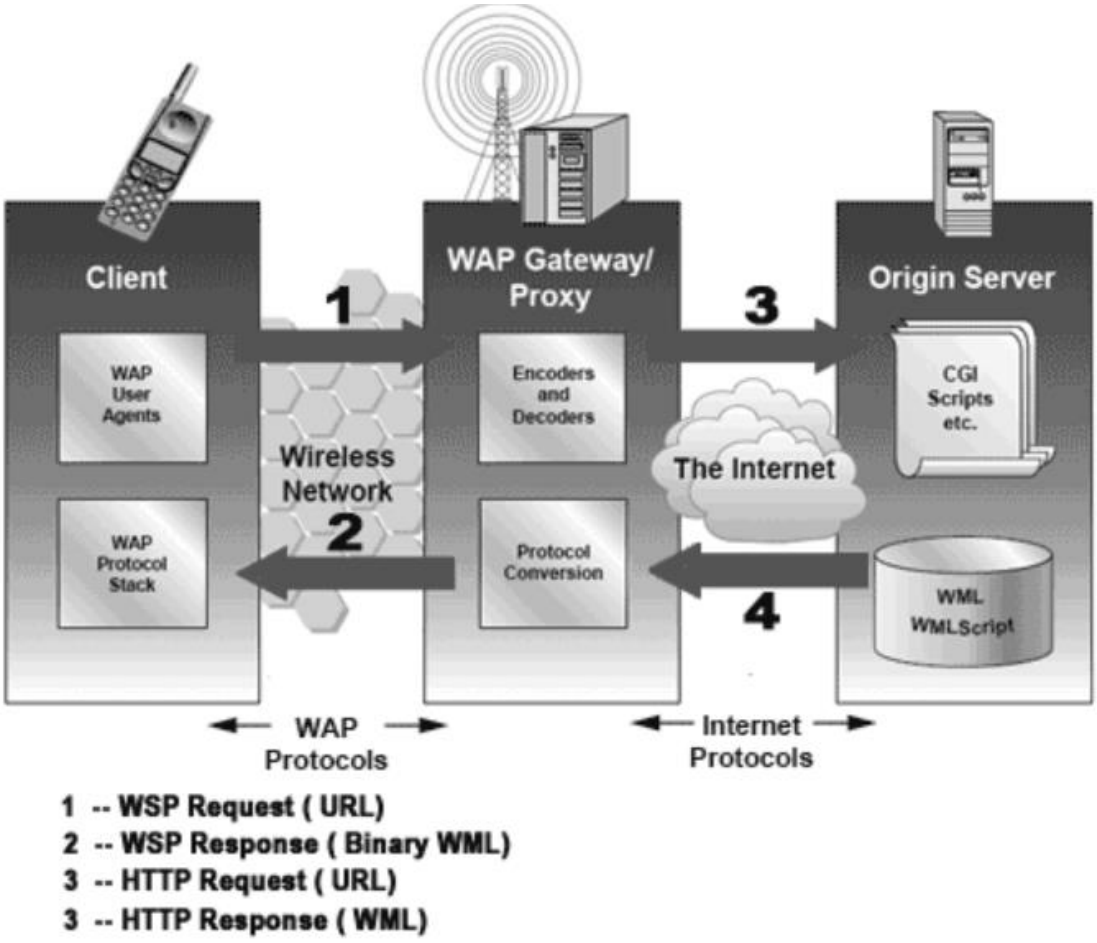


The content available at the web server may be static or dynamic. Static content is produced once and not changed or updated very often; for example, a company presentation. Dynamic

content is needed when the information provided by the service changes more often; for example, timetables, news, stock quotes, and account information. Technologies such as Active Server Pages (ASP), Common Gateway Interface (CGI), and Servlets allow content to be generated dynamically.

The WAP Model

The figure below shows the WAP programming model. Note, the similarities with the Internet model. Without the WAP Gateway/Proxy, the two models would have been practically identical.



WAP Gateway/Proxy is the entity that connects the wireless domain with the Internet. You should make a note that the request that is sent from the wireless client to the WAP Gateway/Proxy uses the Wireless Session Protocol (WSP). In its essence, WSP is a binary version of HTTP.

A **markup language** − the Wireless Markup Language (WML) has been adapted to develop optimized WAP applications. In order to save valuable bandwidth in the wireless network, WML can be encoded into a compact binary format. Encoding WML is one of the tasks performed by the WAP Gateway/Proxy.

How WAP Model Works?

When it comes to actual use, WAP works as follows −

- The user selects an option on their mobile device that has a URL with Wireless Markup language (WML) content assigned to it.
- The phone sends the URL request via the phone network to a WAP gateway using the binary encoded WAP protocol.
- The gateway translates this WAP request into a conventional HTTP request for the specified URL and sends it on to the Internet.
- The appropriate Web server picks up the HTTP request.
- The server processes the request just as it would any other request. If the URL refers to a static WML file, the server delivers it. If a CGI script is requested, it is processed and the content returned as usual.
- The Web server adds the HTTP header to the WML content and returns it to the gateway.
- The WAP gateway compiles the WML into binary form.
- The gateway then sends the WML response back to the phone.
- The phone receives the WML via the WAP protocol.
- The micro-browser processes the WML and displays the content on the screen.

Layers of WAP Protocol

Application Layer

Wireless Application Environment (WAE). This layer is of most interest to content developers because it contains among other things, device specifications, and the content development programming languages, WML, and WMLScript.

Session Layer

Wireless Session Protocol (WSP). Unlike HTTP, WSP has been designed by the WAP Forum to provide fast connection suspension and reconnection.

Transaction Layer

Wireless Transaction Protocol (WTP). The WTP runs on top of a datagram service, such as User Datagram Protocol (UDP) and is part of the standard suite of TCP/IP protocols used to provide a simplified protocol suitable for low bandwidth wireless stations.

Security Layer

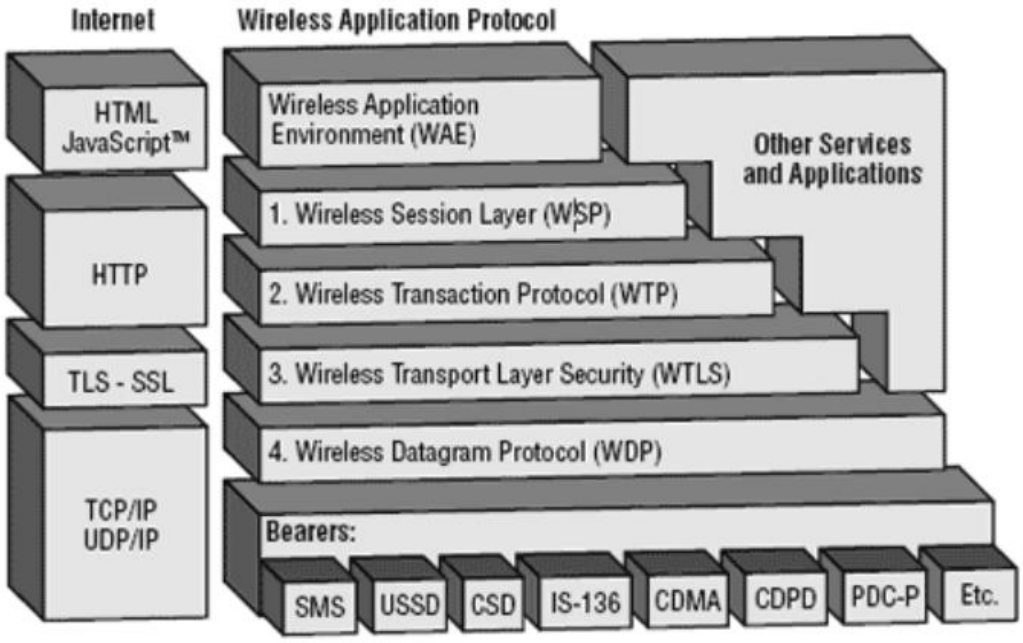
Wireless Transport Layer Security (WTLS). WTLS incorporates security features that are based upon the established Transport Layer Security (TLS) protocol standard. It includes data integrity checks, privacy, service denial, and authentication services.

Transport Layer

Wireless Datagram Protocol (WDP). The WDP allows WAP to be bearer-independent by adapting the transport layer of the underlying bearer. The WDP presents a consistent data format to the higher layers of the WAP protocol stack, thereby offering the advantage of bearer independence to application developers.

Each of these layers provides a well-defined interface to the layer above it. This means that the internal workings of any layer are transparent or invisible to the layers above it. The layered architecture allows other applications and services to utilise the features provided by the WAP-stack as well. This makes it possible to use the WAP-stack for services and applications that currently are not specified by WAP.

The WAP protocol architecture is shown below alongside a typical Internet Protocol stack.



ireless Application Environment (WAE), the uppermost layer in the WAP stack, provides an environment that enables a wide range of applications to be used on the wireless devices. We have earlier discussed about the WAP WAE programming model. In this chapter, we will focus on the various components of WAE.

Components of WAE

Addressing Model

A syntax suitable for naming resources stored on servers. WAP use the same addressing model as the one used on the Internet that is Uniform Resource Locators (URL).

Wireless Markup Language (WML)

A lightweight markup language designed to meet the constraints of a wireless environment with low bandwidth and small handheld devices. The Wireless Markup Language is WAP's analogy to HTML used on the WWW. WML is based on the Extensible Markup Language (XML).

WMLScript

A lightweight scripting language. WMLScript is based on ECMAScript, the same scripting language that JavaScript is based on. It can be used for enhancing services written in WML in the way that it to some extent adds intelligence to the services; for example, procedural logic, loops, conditional expressions, and computational functions.

Wireless Telephony Application (WTA, WTAI)

A framework and programming interface for telephony services. The Wireless Telephony Application (WTA) environment provides a means to create telephony services using WAP.

Hardware and Software Requirement

At minimum developing WAP applications requires a web server and a WAP simulator. Using simulator software while developing a WAP application is convenient as all the required software can be installed on the development PC.

Although, software simulators are good in their own right, no WAP application should go into production without testing it with actual hardware. The following list gives a quick overview of the necessary hardware and software to test and develop WAP applications −

- A web server with connection to the Internet
- A WML to develop WAP application
- A WAP simulator to test WAP application
- A WAP gateway
- A WAP phone for final testing.

Microsoft IIS or Apache on Windows or Linux can be used as the web server and Nokia WAP Toolkit version 2.0 as the WinWAP simulator.

Please have look at WAP - Useful Resources to find out all the above components.

Configure Web Server for WAP

In the WAP architecture, the web server communicates with the WAP gateway, accepting HTTP requests and returning WML code to the gateway. The HTTP protocol mandates that each reply must include something called a Multi-Purpose Internet Mail Extensions (MIME) type.

In normal web applications, this MIME type is set to text/html, designating normal HTML code. Images on the other hand could be specified as image/gif or image/jpeg for instance. With this content type specification, the web browser knows the data type that the web server returns.

The topmost layer in the WAP architecture is made up of WAE (Wireless Application Environment), which consists of WML and WML scripting language.

WML scripting language is used to design applications that are sent over wireless devices such as mobile phones. This language takes care of the small screen and the low bandwidth of transmission. WML is an application of XML, which is defined in a document-type definition.

WML pages are called decks. They are constructed as a set of cards, related to each other with links. When a WML page is accessed from a mobile phone, all the cards in the page are downloaded from the WAP server to mobile phone showing the content.

WML commands and syntaxes are used to show content and to navigate between the cards. Developers can use these commands to declare variables, format text, and show images on the mobile phone.

WAP Program Structure

A WML program is typically divided into two parts − the **document prolog** and the **body**. Consider the following code −

```
<?xml version="1.0"?>
<!DOCTYPE wml PUBLIC "-//WAPFORUM//DTD WML 1.2//EN"
"http://www.wapforum.org/DTD/wml12.dtd">
<wml>
  <card>

  ...
  </card>
  ...more cards...
</wml>
```

The first line of this text says that this is an XML document and the version is 1.0. The second line selects the document type and gives the URL of the **document type definition** (DTD). This DTD gives the full XML definition of WML. The DTD referenced is defined in WAP 1.1, but this header changes with the versions of the WML. The header must be copied exactly so that the tool kits automatically generate this prolog.

The body is enclosed within a <wml>...</wml> tag pair as shown above. The body of a WML document can consist of one or more of the following −

- Deck
- Card
- Content to be shown
- Navigation instructions

MESSAGING SERVICES

Mobile communication industry uses a number of acronyms for the various technologies that are being developed.

Popular mobile communication technologies

- Short Message Service (SMS)
- Multimedia Messaging Service (MMS)
- Global Positioning System (GPS)
- Smart cards

Short Message Service (SMS)

Short Message Service (SMS) is a text messaging service in mobile communication systems that allows exchanging short text messages. SMS is the most widely used data application by mobile phone users. The GSM standard allows to send a message containing upto 160 characters. When a message is sent, it reaches a Short Message Service Center (SMSC), which provides a 'store and forward' mechanism. SMSC attempts to send messages to the recipients. If a recipient is not reachable, the SMSC waits and then retries later. Some SMSC's also provide a 'forward and forget' option where transmission is tried only once and if it fails, the message is not sent again. SMS messages are exchanged using the protocol called SS7 (Signalling System No # 7).

Multimedia Messaging Service (MMS)

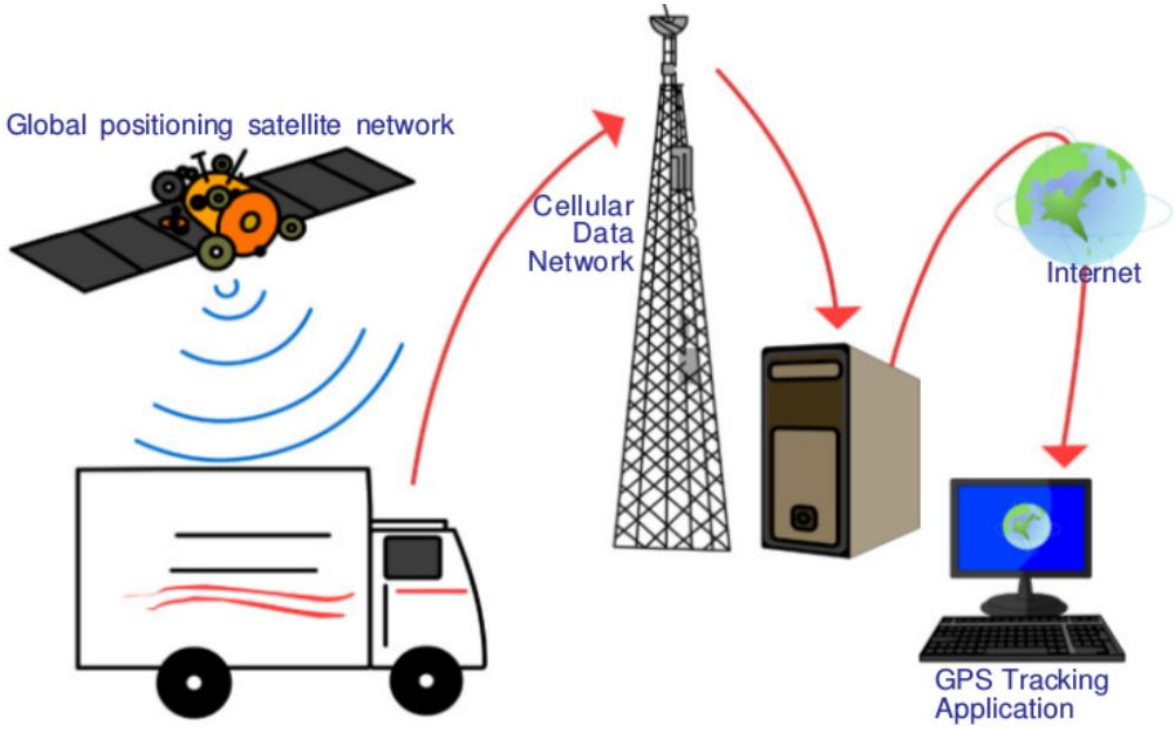
Multimedia Messaging Service (MMS) is a standard way to send and receive messages that consists of multimedia content using mobile phones. It is an extension of the capability of SMS to send text messages. Unlike SMS, MMS does not specify a maximum size for a multimedia message. MMS supports contents such as text, graphics, music, video clips and more. An MMS server is responsible for storing and handling incoming and outgoing messages. Associated with the MMS server is the MMS proxy relay, which is responsible for transferring messages between different messaging systems.

Global Positioning System (GPS)

The Global Positioning System (GPS) is a satellite based navigation system that is used to locate a geographical position anywhere on earth, using its longitude and latitude. GPS is designed and operated by the U.S. Department of Defence and it consists of satellites, control and monitoring stations, and receivers.

The basis of the GPS is a group of satellites that are continuously orbiting the earth. These satellites transmit radio signals that contain their exact location, time,and other information. The radio signals from the satellites, which are monitored and corrected by control stations, are picked up by the GPS receiver. GPS receivers take information transmitted from the satellites to calculate a user's exact location on earth. A GPS receiver needs only three satellites to plot a 2D position, which will not be very accurate. Ideally, four or more satellites are needed to plot a 3D position, which is much more accurate.

GPS is used for vehicle fleet tracking by transporting companies to track the movement of their trucks. Figure 12.19 depicts the working of a truck tracking application using GPS. Vehicle navigation systems will direct the driver to his or her destination through the best route. In commercial aviation GPS is used for aircraft navigation. GPS is also used in oil exploration, farming, atmospheric studies, etc. GPS receivers are now integrated in many mobile phones for implementing various tracking applications.



-
-
  - Generations in mobile communication
  - Mobile communication services
    - Short Message Service (SMS)
    - Multimedia Messaging Service (MMS)
    - Global Positioning System (GPS)
    - **Smart cards**

- Mobile operating system

« Previous Next »

Smart cards

Let us recollect about smart cards and smart card readers that we discussed in Chapter 3. A smart card is a plastic card embedded with a computer chip / memory that stores and transacts data. The advantages of using smart cards is that it is secure (data is protected), intelligent (it can store and process data) and that it is convenient (it is easy to carry). That is why businesses and other organisations use smart cards for authentication and storing data.A model of smart card issued by Government of India for RSBY scheme is shown in Figure

In mobile communication the smart card technology is used in Subscriber Identity Modules (SIM) for GSM phone systems (refer Figure 12.21). The smart card is inserted or integrated into the mobile handset. The card stores personal subscriber information and preferences. SIM cards help to identify a subscriber, roam across networks and provide security to value added services like Internet browsing, mobile commerce, mobile banking, etc. Smart cards also work as credit cards, ATM cards, fuel cards, authorization cards for television receiver, high-security identification

USEFUL LINKS

https://www.tutorialspoint.com/mobile_computing/index.htm
https://studentsfocus.com/cs8601-mc-notes-mobile-computing-notes-csc-6th-sem/
https://lecturenotes.in/notes/10134-notes-for-mobile-computing-mc-by-annapurna-mishra
https://cseexamhacks.files.wordpress.com/2017/01/mobile-computing.pdf