LECTURE NOTES
ON

MOBILE COMPUTING

B. Tech CSE

6TH semester

Stay @ Home

Stay Safe

To Avoid and fight Corona Virus

Your Tutor/Mentor

CSE-306N	Mobile Computing					
Lecture	**Tutorial**	**Practical**	**Major Test**	**Minor Test**	**Total**	**Time**
3	**1**	**0**	**75**	**25**	**100**	**3 Hrs.**
Purpose	To impart knowledge of mobile and wireless computing systems and techniques.					
Course Outcomes(CO)						
CO1	Describe the concepts of mobile computing and cellular networks.					
CO2	Learn the basic concepts of wireless networks.					
CO3	Study of various issues of mobile computing and basics of cloud computing.					
CO4	Description and applications of Ad hoc networks.					

UNIT – I

Introduction, issues in mobile computing, overview of wireless telephony: cellular concept, Mobile computing Architecture, Design considerations for mobile computing, Mobile Computing through Internet, Making existing applications mobile enabled. GSM: air-interface, channel structure, location management: HLR-VLR, hierarchical, handoffs, channel allocation in Cellular systems, WCDMA, GPRS 3G, 4G.

UNIT – II

Wireless Networking, Wireless LAN Overview: MAC issues, IEEE 802.11, Blue Tooth, Wireless multiple access protocols, TCP over wireless, Wireless applications, data broadcasting, Mobile IP, WAP : Architecture, Traditional TCP, Classical TCP, improvements in WAP, WAP applications.

UNIT – III

Data management issues, data replication for mobile computers, adaptive clustering for mobile wireless networks, File system, Disconnected operations Mobile Agents computing, security and fault tolerance, transaction processing in mobile computing environment.

Cloud Architecture model, Types of Clouds: Public Private & Hybrid Clouds, Resource management and scheduling, Clustering, Data Processing in Cloud: Introduction to Map Reduce for Simplified data processing on Large clusters.

UNIT – IV

Ad hoc networks, localization, MAC issues, Routing protocols, global state routing (GSR), Destination sequenced distance vector routing (DSDV), Dynamic source routing (DSR), Ad Hoc on demand distance vector routing (AODV), Temporary ordered routing algorithm (TORA), QoS in Ad Hoc Networks, applications.

Text Books:

1. Rajkamal, Mobile Computing, 2/E Oxford University Press,2011.
2. J. Schiller, Mobile Communications, Addison Wesley
3. Yi Bing Lin, Wireless and Mobile Networks Architecture , John Wiley.

Reference Books

1. A. Mehrotra , GSM System Engineering.
2. M. V. D. Heijden, M. Taylor, Understanding WAP, Artech House.
3. Charles Perkins, Mobile IP, Addison Wesley.
4. Charles Perkins, Ad hoc Networks, Addison Wesley.
5. Judith Hurwitz, Robin Bllor, Marcia Kaufmann, Fern Halper, Cloud Computing for Dummies, 2009.

Wireless Network

Wireless networks are computer networks that are not connected by cables of any kind. The use of a wireless network enables enterprises to avoid the costly process of introducing cables into buildings or as a connection between different equipment locations. The basis of wireless systems are radio waves, an implementation that takes place at the physical level of network structure.

Wireless networks use radio waves to connect devices such as laptops to the Internet, the business network and applications. When laptops are connected to Wi-Fi hot spots in public places, the connection is established to that business's wireless network.

There are four main types of wireless networks:

- Wireless Local Area Network (LAN): Links two or more devices using a wireless distribution method, providing a connection through access points to the wider Internet.
- Wireless Metropolitan Area Networks (MAN): Connects several wireless LANs.
- Wireless Wide Area Network (WAN): Covers large areas such as neighboring towns and cities.
- Wireless Personal Area Network (PAN): Interconnects devices in a short span, generally within a person's reach.

WIRELESS MAC ISSUES

The three important issues are:

1. Half Duplex operation –> either send or receive but not both at a given time
2. Time varying channel
3. Burst channel errors

1. Half Duplex Operation

In wireless, it's difficult to receive data when the transmitter is sending the data, because: When node is transmitting, a large fraction of the signal energy leaks into the receiver path. The transmitted and received power levels can differ by orders of magnitude. The leakage signal typically has much higher power than the received signal -—Impossible to detect a received signal, while transmitting data‖. Collision detection is not possible, while sending data. As collision cannot be detected by the sender, all proposed protocols attempt to minimize the probability of collision - Focus on collision avoidance.

2. Time Varying Channel

Three mechanisms for radio signal propagation

· **Reflection** – occurs when a propagating wave impinges upon an object that has very large dimensions than the wavelength of the radio wave e.g. reflection occurs from the surface of the earth and from buildings and walls

· **Diffraction** – occurs when the radio path between the transmitter and the receiver is obstructed by a surface with sharp edges

· **Scattering** – occurs when the medium through which the wave travels consists of objects with

The received signal by a node is a superposition of time-shifted and attenuated versions of the transmitted signals the received signal varies with time .The time varying signals (time varying channel) phenomenon also known as multipath propagation. The rate of variation of channel is determined by the coherence time of the channel Coherence time is defined as time within which When a node's received signal strength

drops below a certain threshold the node is said to be in fade .Handshaking is widely used strategy to ensure the link quality is good enough for data communication. A successful handshake between a sender and a receiver (small message) indicates a good communication link.

3. Burst Channel Errors

As a consequence of time varying channel and varying signals strengths errors are introduced in the transmission (Very likely) for wire line networks the bit error rate (BER) is the probability of packet error is small .For wire line networks the errors are due to random For wireless networks the BER is as high. For wireless networks the errors are due to node being in fade as a result errors occur in a long burst. Packet loss due to burst errors - mitigation techniques

· Smaller packets

· Forward Error Correcting Codes

· Retransmissions (Acks)

Location Dependent Carrier Sensing

Location Dependent Carrier Sensing results in three types of nodes that protocols need to deal with

Hidden Nodes: Even if the medium is free near the transmitter, it may not be free near the

intended receiver

Exposed Nodes: Even if the medium is busy near the transmitter, it may be free near the intended receiver

Capture: Capture occurs when a receiver can cleanly receive a transmission from one of two simultaneous transmissions

Hidden Node/Terminal Problem

A hidden node is one that is within the range of the intended destination but out of range of sender Node B can communicate with A and C both A and C cannot hear each other When A transmits to B, C cannot detect the transmission using the carrier sense mechanism C falsely thinks that the channel is idle

Exposed Nodes
An exposed node is one that is within the range of the sender but out of range of destination .when a node‘s received signal strength drops below a certain threshold the node is said to be in fade .Handshaking is widely used strategy to ensure the link quality is good enough for data communication. A successful handshake between a sender and a receiver (small message) indicates a good communication link.

In theory C can therefore have a parallel transmission with any node that cannot hear the transmission from B, i.e. out of range of B. But C will not transmit to any node because its an exposed node. Exposed nodes waste bandwidth.

Capture
Capture is said to occur when a receiver can cleanly receive a transmission from one of two simultaneous transmissions both within its range Assume node A and D transmit simultaneously to B. The signal strength received from D is much higher than that from A, and

D‘s transmission can be decoded without errors in presence of transmissions from A.D has captured A. Capture is unfair because it gives preference to nodes that are closer to the receiver. It may improve protocol performance.

IEEE 802.11

Wireless LANs are those Local Area Networks that use high frequency radio waves instead of cables for connecting the devices in LAN. Users connected by WLANs can move around within the area of network coverage. Most WLANs are based upon the standard IEEE 802.11 or WiFi.

IEEE 802.11 Architecture

The components of an IEEE 802.11 architecture are as follows

1) Stations (STA) − Stations comprise all devices and equipments that are connected to the wireless LAN. A station can be of two types:

- **Wireless Access Pointz (WAP) –** WAPs or simply access points (AP) are generally wireless routers that form the base stations or access.
- **Client. –** Clients are workstations, computers, laptops, printers, smartphones, etc.

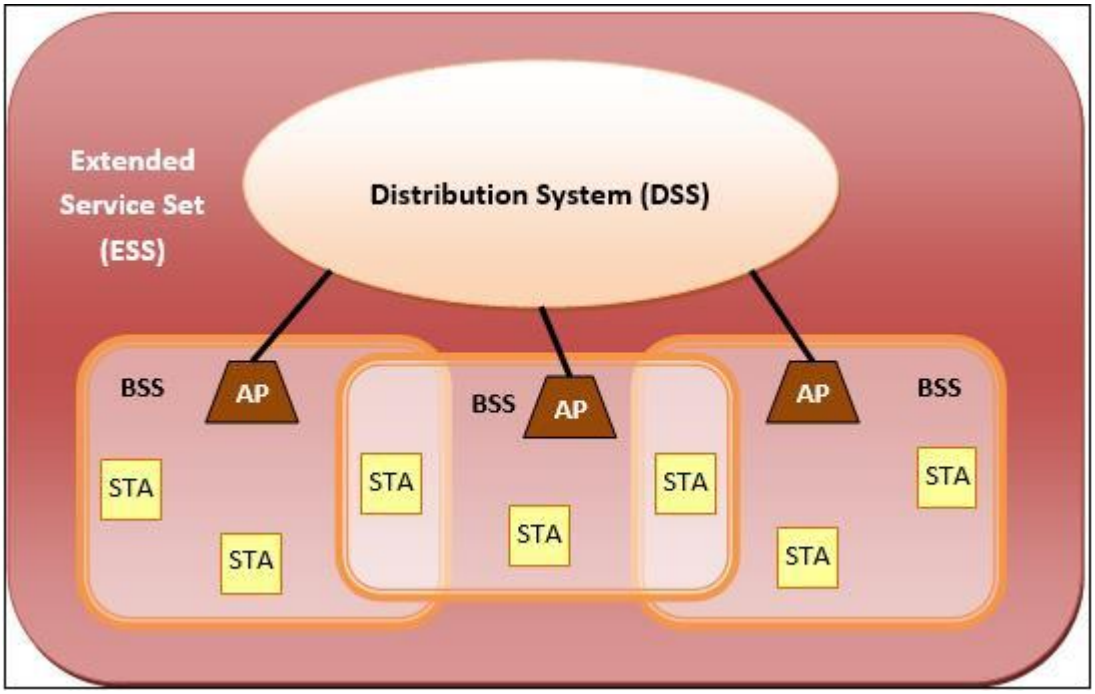
Each station has a wireless network interface controller.

2) Basic Service Set (BSS) −A basic service set is a group of stations communicating at physical layer level. BSS can be of two categories depending upon mode of operation:

- **Infrastructure BSS –** Here, the devices communicate with other devices through access points.
- **Independent BSS –** Here, the devices communicate in peer-to-peer basis in an ad hoc manner.

3) Extended Service Set (ESS) − It is a set of all connected BSS.

4) Distribution System (DS) − It connects access points in ESS.



Advantages of WLANs

- They provide clutter free homes, offices and other networked places.
- The LANs are scalable in nature, i.e. devices may be added or removed from the network at a greater ease than wired LANs.
- The system is portable within the network coverage and access to the network is not bounded by the length of the cables.

- Installation and setup is much easier than wired counterparts.
- The equipment and setup costs are reduced.

Disadvantages of WLANs

- Since radio waves are used for communications, the signals are noisier with more interference from nearby systems.
- Greater care is needed for encrypting information. Also, they are more prone to errors. So, they require greater bandwidth than the wired LANs.
- WLANs are slower than wired LANs.

IEEE 802.11

IEEE 802.11 is a set of media access control (MAC) and physical layer (PHY) specifications for implementing wireless local area network(WLAN) computer communication in the 900 MHz and 2.4, 3.6, 5, and 60 GHz frequency bands

The IEEE developed an international standard for WLANs. The 802.11 standard focuses on the bottom two layers of the OSI model, the physical layer (PHY) and data link layer (DLL).

The objective of the IEEE 802.11 standard was to define a medium access control (MAC) sublayer, MAC management protocols and services, and three PHYs for wireless connectivity of fixed, portable, and moving devices within a local area.

The three physical layers are an IR base band PHY, an FHSS radio in the 2.4 GHz band, and a DSSS radio in the 2.4 GHz.

IEEE 802.11 Architecture:

The architecture of the IEEE 802.11 WLAN is designed to support a network where most decision making is distributed to mobile stations. This type of architecture has several advantages. It is tolerant of faults in all of the WLAN equipment and eliminates possible bottlenecks a centralized architecture would introduce. The architecture is flexible and can easily support both small, transient networks and large, semipermanent or permanent networks. In addition, the architecture and protocols offer significant power saving and prolong the battery life of mobile equipment without losing network connectivity

Two network architectures are defined in the IEEE 802.11 standard:

- **Infrastructure network:** An infrastructure network is the network architecture for providing communication between wireless clients and wired network resources. The transition of data from the wireless to wired medium occurs via an AP. An AP and its associated wireless clients define the coverage area. Together all the devices form a basic service set (refer figure 1).
- **Point-to-point (ad-hoc) network:** An ad-hoc network is the architecture that is used to support mutual communication between wireless clients. Typically, an ad-hoc network is created spontaneously and does not support access to wired networks. An ad-hoc network does not require an AP.

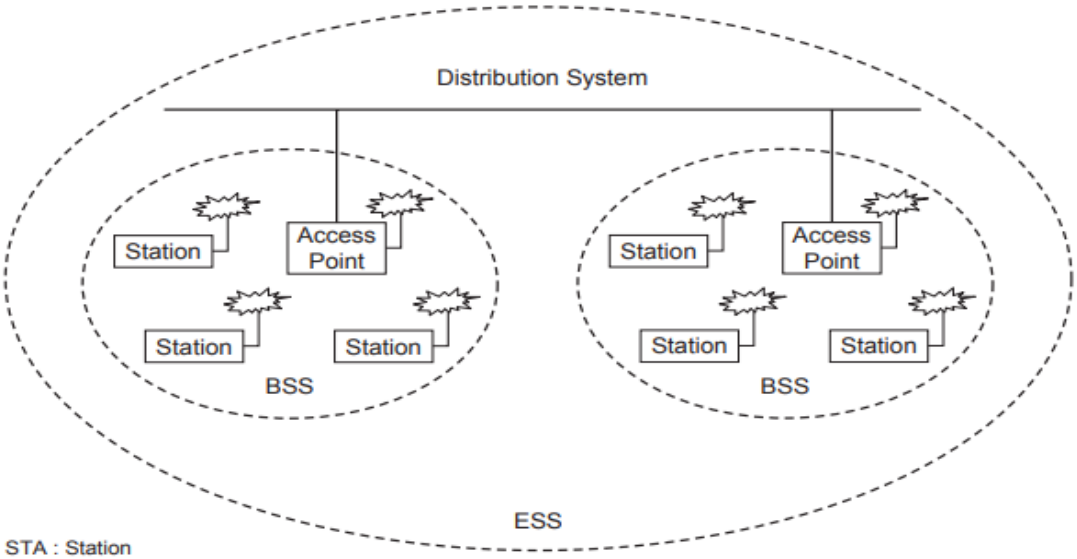


Fig1: BSS and ESS configuration of IEEE 802.11 WLAN

IEEE 802.11 supports three basic topologies for WLANs, the independent basic service set (IBSS), the basic service set, and the extended service set (ESS). The MAC layer supports implementations of IBSS, basic service set, and ESS configurations.

Independent basic service set: The IBSS configuration is referred to as an independent configuration or an ad-hoc network. An IBSS configuration is analogous to a peer-to-peer office network in which no single node is required to act as a server. IBSS WLANs include a number of nodes or wireless stations that communicate directly with one another on an ad-hoc, peer-to-peer basis. Generally, IBSS implementations cover a limited area and are not connected to any large network. An IBSS is typically a short-lived network, with a small number of stations, that is created for a particular purpose.

Basic service set: The basic service set configuration relies on an AP that acts as the logical server for a single WLAN cell or channel. Communications between station 1 and station 4 actually flow from station 1 to AP1 and then from AP1 to AP2 and then from AP2 to AP4 and finally AP4 to station 4 (refer to Figure 2). An AP performs a bridging function and connects multiple WLAN cells or channels, and connects WLAN cells to a wired enterprise LAN.

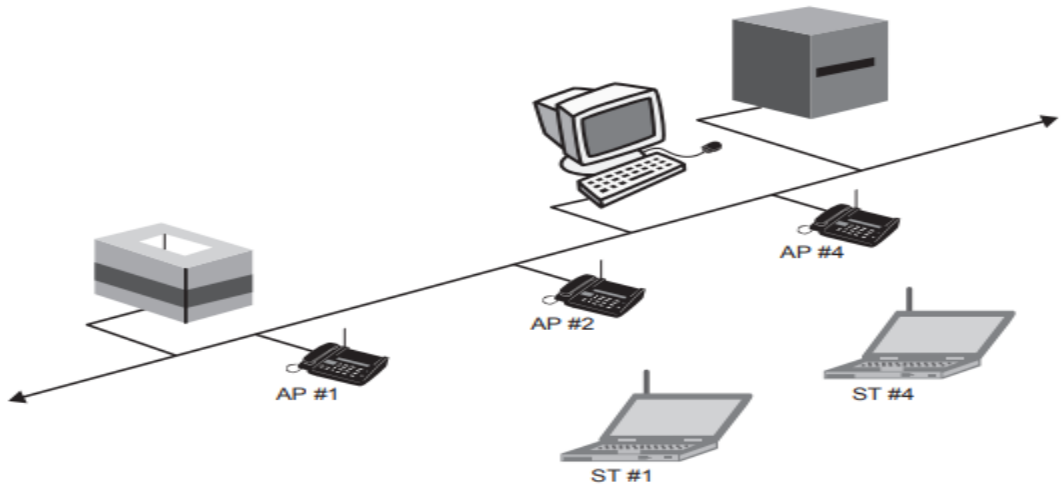


Fig.2 Access point-based topology

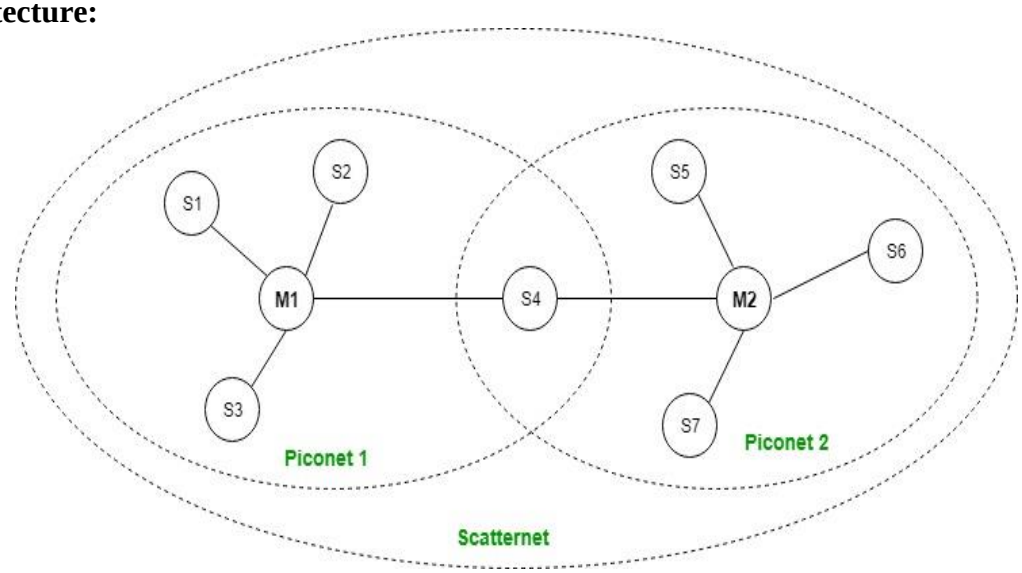
Extended service set: The ESS configuration consists of multiple basic service set cells that can be linked by either wired or wireless backbones called a distributed system. IEEE 802.11 supports ESS configurations in which multiple cells use the same channel, and configurations in which multiple cells use different channels to boost aggregate throughput. To network the equipment outside of the ESS, the ESS and all of its mobile stations appear

to be a single MAC layer network where all stations are physically stationary. Thus, the ESS hides the mobility of the mobile stations from everything outside the ESS (refer figure 1).

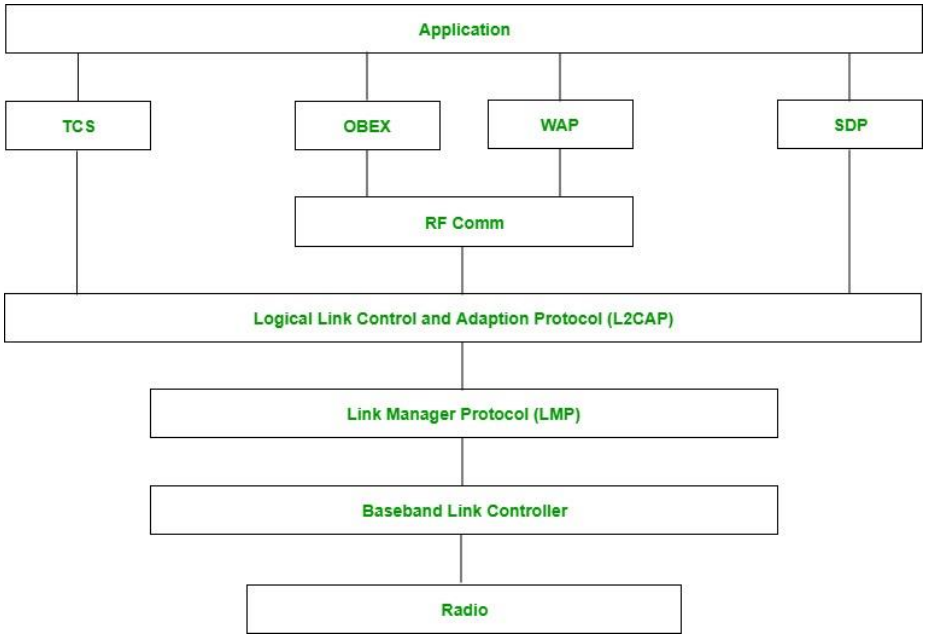
Bluetooth

It is a Wireless Personal Area Network (WPAN) technology and is used for exchanging data over smaller distances. This technology was invented by Ericson in 1994. It operates in the unlicensed, industrial, scientific and medical (ISM) band at 2.4 GHz to 2.485 GHz. Maximum devices that can be connected at the same time are 7. Bluetooth ranges upto 10 meters. It provides data rates upto 1 Mbps or 3 Mbps depending upon the version. The spreading technique which it uses is FHSS (Frequency hopping spread spectrum). A bluetooth network is called **piconet** and a collection of interconnected piconets is call **scatternet**.

Bluetooth Architecture:



Bluetooth protocol stack:



1. **Radio (RF) layer:**
   It performs modulation/demodulation of the data into RF signals. It defines the physical characteristics of bluetooth transceiver. It defines two types of physical link: connection-less and connection-oriented.

2. **Baseband Link layer:**
   It performs the connection establishment within a piconet.
3. **Link Manager protocol layer:**
   It performs the management of the already established links. It also includes authentication and encryption processes.
4. **Logical Link Control and Adaption protocol layer:**
   It is also known as the heart of the bluetooth protocol stack. It allows the communication between upper and lower layers of the bluetooth protocol stack. It packages the data packets received from upper layers into the form expected by lower layers. It also performs the segmentation and multiplexing.
5. **SDP layer:**
   It is short for Service Discovery Protocol. It allows to discover the services available on another bluetooth enabled device.
6. **RF comm layer:**
   It is short for Radio Frontend Component. It provides serial interface with WAP and OBEX.
7. **OBEX:**
   It is short for Object Exchange. It is a communication protocol to exchange objects between 2 devices.
8. **WAP:**
   It is short for Wireless Access Protocol. It is used for internet access.
9. **TCS:**
   It is short for Telephony Control Protocol. It provides telephony service.
10. **Application layer:**
    It enables the user to interact with the application.

Advantages:

- Low cost.
- Easy to use.
- It can also penetrate through walls.
- It creates an adhoc connection immediately without any wires.
- It is used for voice and data transfer.

Disadvantages:

- It can be hacked and hence, less secure.
- It has slow data transfer rate: 3 Mbps.
- It has small range: 10 meters.

Multiple Access Techniques

Multiple access schemes are used to allow many mobile users to share simultaneously a finite amount of radio spectrum.

In wireless communication systems, it is often desirable to allow the subscriber to send information simultaneously from the mobile station to the base station while receiving information from the base station to the mobile station.

A cellular system divides any given area into cells where a mobile unit in each cell communicates with a base station. The main aim in the cellular system design is to be able to **increase the capacity of the channel**, i.e., to handle as many calls as possible in a given bandwidth with a sufficient level of quality of service.

There are several different ways to allow access to the channel.

These includes mainly the following −

- Frequency division multiple-access (FDMA)
- Time division multiple-access (TDMA)
- Code division multiple-access (CDMA)
- Space division multiple access (SDMA)

Depending on how the available bandwidth is allocated to the users, these techniques can be classified as **narrowband** and **wideband** systems.

Narrowband Systems

Systems operating with channels substantially narrower than the coherence bandwidth are called as Narrow band systems. Narrow band TDMA allows users to use the same channel but allocates a unique time slot to each user on the channel, thus separating a small number of users in time on a single channel.

Wideband Systems

In wideband systems, the transmission bandwidth of a single channel is much larger than the coherence bandwidth of the channel. Thus, multipath fading doesn't greatly affect the received signal within a wideband channel, and frequency selective fades occur only in a small fraction of the signal bandwidth.

Frequency Division Multiple Access (FDMA)

FDMA is the basic technology for advanced mobile phone services. The features of FDMA are as follows.

- FDMA allots a different sub-band of frequency to each different user to access the network.
- If FDMA is not in use, the channel is left idle instead of allotting to the other users.
- FDMA is implemented in Narrowband systems and it is less complex than TDMA.
- Tight filtering is done here to reduce adjacent channel interference.
- The base station BS and mobile station MS, transmit and receive simultaneously and continuously in FDMA.

Time Division Multiple Access (TDMA)

In the cases where continuous transmission is not required, there TDMA is used instead of FDMA. The features of TDMA include the following.

- TDMA shares a single carrier frequency with several users where each users makes use of non-overlapping time slots.
- Data transmission in TDMA is not continuous, but occurs in bursts. Hence handsoff process is simpler.
- TDMA uses different time slots for transmission and reception thus duplexers are not required.
- TDMA has an advantage that is possible to allocate different numbers of time slots per frame to different users.
- Bandwidth can be supplied on demand to different users by concatenating or reassigning time slot based on priority.

Code Division Multiple Access (CDMA)

Code division multiple access technique is an example of multiple access where several transmitters use a single channel to send information simultaneously. Its features are as follows.

- In CDMA every user uses the full available spectrum instead of getting allotted by separate frequency.
- CDMA is much recommended for voice and data communications.
- While multiple codes occupy the same channel in CDMA, the users having same code can communicate with each other.
- CDMA offers more air-space capacity than TDMA.
- The hands-off between base stations is very well handled by CDMA.

Space Division Multiple Access (SDMA)

Space division multiple access or spatial division multiple access is a technique which is MIMO (multiple-input multiple-output) architecture and used mostly in wireless and satellite communication. It has the following features.

- All users can communicate at the same time using the same channel.

- SDMA is completely free from interference.
- A single satellite can communicate with more satellites receivers of the same frequency.
- The directional spot-beam antennas are used and hence the base station in SDMA, can track a moving user.
- Controls the radiated energy for each user in space.

Spread Spectrum Multiple Access

Spread spectrum multiple access (SSMA) uses signals which have a transmission bandwidth whose magnitude is greater than the minimum required RF bandwidth.

There are two main types of spread spectrum multiple access techniques −

- Frequency hopped spread spectrum (FHSS)
- Direct sequence spread spectrum (DSSS)

Frequency Hopped Spread Spectrum (FHSS)

This is a digital multiple access system in which the carrier frequencies of the individual users are varied in a pseudo random fashion within a wideband channel. The digital data is broken into uniform sized bursts which is then transmitted on different carrier frequencies.

Direct Sequence Spread Spectrum (DSSS)

This is the most commonly used technology for CDMA. In DS-SS, the message signal is multiplied by a Pseudo Random Noise Code. Each user is given his own code word which is orthogonal to the codes of other users and in order to detect the user, the receiver must know the code word used by the transmitter.

The combinational sequences called as **hybrid** are also used as another type of spread spectrum. **Time hopping** is also another type which is rarely mentioned.

Since many users can share the same spread spectrum bandwidth without interfering with one another, spread spectrum systems become **bandwidth efficient** in a multiple user environment.

The wireless channel is susceptible to a variety of transmission impediments such as **path loss, interference** and **blockage**. These factors restrict the range, data rate, and the reliability of the wireless transmission.

Types of Paths

The extent to which these factors affect the transmission depends upon the environmental conditions and the mobility of the transmitter and receiver. The path followed by the signals to get to the receiver, are two types, such as −

Direct-path

The transmitted signal, when reaches the receiver directly, can be termed as a **directpath** and the components presents that are present in the signal are called as **directpath components**.

Multi-path

The transmitted signal when reaches the receiver, through different directions undergoing different phenomenon, such a path is termed as **multi-path** and the components of the transmitted signal are called as **multi-path components**.

They are reflected, diffracted and scattered by the environment, and arrive at the receiver shifted in amplitude, frequency and phase with respect to the direct path component.

Characteristics of Wireless Channel

The most important characteristics of wireless channel are −

- Path loss
- Fading
- Interference
- Doppler shift

In the following sections, we will discuss these channel characteristics one by one.

Path Loss

Path loss can be expressed as the ratio of the power of the transmitted signal to the power of the same signal received by the receiver, on a given path. It is a function of the propagation distance.

- Estimation of path loss is very important for designing and deploying wireless communication networks
- Path loss is dependent on a number of factors such as the radio frequency used and the nature of the terrain.
- The free space propagation model is the simplest path loss model in which there is a direct-path signal between the transmitter and the receiver, with no atmosphere attenuation or multipath components.

In this model, the relationship between the transmitted power P_t and the received power P_r is given by

$$P_{r} = P_{t}G_{t}G_{r}(\frac{\lambda}{4\Pi d})^2$$

Where

- G_t is the transmitter antenna gain
- G_r is the receiver antenna gain
- **d** is the distance between the transmitter and receiver
- **λ** is the wavelength of the signal

Two-way model also called as two path models is widely used path loss model. The free space model described above assumes that there is only one single path from the transmitter to the receiver.

In reality, the signal reaches the receiver through multiple paths. The two path model tries to capture this phenomenon. The model assumes that the signal reaches the receiver through two paths, one a line-of-sight and the other the path through which the reflected wave is received.

According to the two-path model, the received power is given by

$$P_{r} = P_{t}G_{t}G_{r}(\frac{h_{t}h_{r}}{d^2})^2$$

Where

- p_t is the transmitted power
- G_t represent the antenna gain at the transmitter
- G_r represent the antenna gain at the receiver
- **d** is the distance between the transmitter and receiver
- h_t is the height of the transmitter
- h_r are the height of the receiver

Fading

Fading refers to the fluctuations in signal strength when received at the receiver. Fading can be classified in to two types −

- Fast fading/small scale fading and
- Slow fading/large scale fading

Fast fading refers to the rapid fluctuations in the amplitude, phase or multipath delays of the received signal, due to the interference between multiple versions of the same transmitted signal arriving at the receiver at slightly different times.

The time between the reception of the first version of the signal and the last echoed signal is called **delay spread**. The multipath propagation of the transmitted signal, which causes fast fading, is because of the three propagation mechanisms, namely −

- Reflection
- Diffraction
- Scattering

The multiple signal paths may sometimes add constructively or sometimes destructively at the receiver causing a variation in the power level of the received signal. The received single envelope of a fast fading signal is said to follow a **Rayleigh distribution** to see if there is no line-of-sight path between the transmitter and the receiver.

Slow Fading

The name Slow Fading itself implies that the signal fades away slowly. The features of slow fading are as given below.

- Slow fading occurs when objects that partially absorb the transmission lie between the transmitter and receiver.
- Slow fading is so called because the duration of the fade may last for multiple seconds or minutes.
- Slow fading may occur when the receiver is inside a building and the radio wave must pass through the walls of a building, or when the receiver is temporarily shielded from the transmitter by a building. The obstructing objects cause a random variation in the received signal power.
- Slow fading may cause the received signal power to vary, though the distance between the transmitter and receiver remains the same.
- Slow fading is also referred to as **shadow fading** since the objects that cause the fade, which may be large buildings or other structures, block the direct transmission path from the transmitter to the receiver.

Interference

Wireless transmissions have to counter interference from a wide variety of sources. Two main forms of interference are −

- Adjacent channel interference and
- Co-channel interference.

In Adjacent channel interference case, signals in nearby frequencies have components outside their allocated ranges, and these components may interfere with on-going transmission in the adjacent frequencies. It can be avoided by carefully introducing guard bands between the allocated frequency ranges.

Co-channel interference, sometimes also referred to as **narrow band interference**, is due to other nearby systems using the same transmission frequency.

Inter-symbol interference is another type of interference, where distortion in the received signal is caused by the temporal spreading and the consequent overlapping of individual pulses in the signal.

Adaptive equalization is a commonly used technique for combating inter symbol interference. It involves gathering the dispersed symbol energy into its original time interval. Complex digital processing algorithms are used in the equalization process.

The original TCP/IP protocol was defined as four software layers built upon the hardware. Today, however, TCP/IP is thought of as a five-layer model with the layers named similar to the ones in the OSI model.

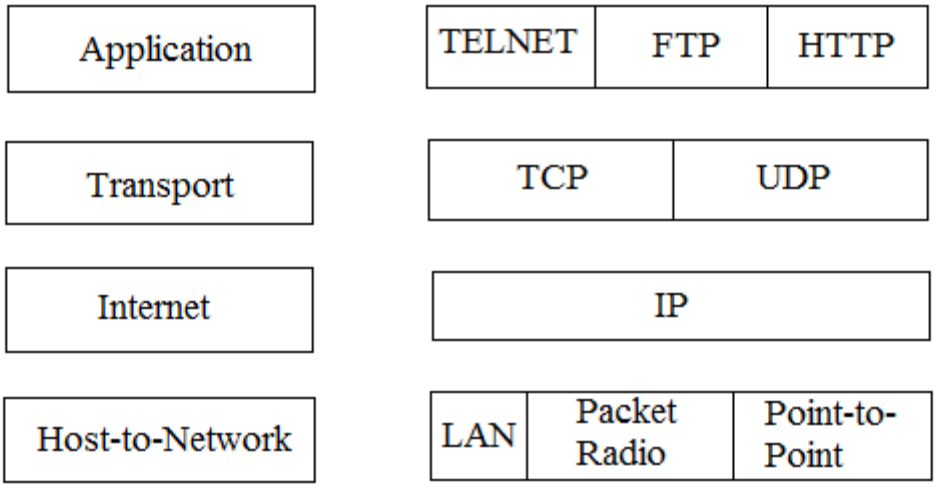
Comparison between OSI and TCP/IP Suite

When we compare the two models, we find that two layers, session and presentation, are missing from the TCP/IP protocol. The application layer in the suite is usually considered to be the combination of three layers in the OSI model.

The OSI model specifies which functions belong to each of its layers but the layers of the TCP/IP protocol suite contain relatively independent protocols that can be mixed and matched, depending on the needs of the system. The term hierarchical means that each upper level protocol is supported by one or more lower level protocols.

Layers in the TCP/IP Suite

The four layers of the TCP/IP model are the host-to-network layer, internet/network layer, transport layer and the application layer. The purpose of each layer in the TCP/IP protocol suite is detailed below.



The above image represents the layers of TCP/IP protocol suite.

Physical Layer

TCP/IP does not define any specific protocol for the physical layer. It supports all of the standard and proprietary protocols.

- At this level, the communication is between two hops or nodes, either a computer or router. The unit of communication is a **single bit**.
- When the connection is established between the two nodes, a stream of bits is flowing between them. The physical layer, however, treats each bit individually.

The responsibility of the physical layer, in addition to delivery of bits, matches with what mentioned for the physical layer of the OSI model, but it mostly depends on the underlying technologies that provide links.

Data Link Layer

TCP/IP does not define any specific protocol for the data link layer either. It supports all of the standard and proprietary protocols.

- At this level also, the communication is between two hops or nodes. The unit of communication however, is a packet called a **frame**.
- A **frame** is a packet that encapsulates the data received from the network layer with an added header and sometimes a trailer.
- The head, among other communication information, includes the source and destination of frame.
- The **destination address** is needed to define the right recipient of the frame because many nodes may have been connected to the link.
- The **source address** is needed for possible response or acknowledgment as may be required by some protocols.

LAN, Packet Radio and Point-to-Point protocols are supported in this layer

Network Layer

At the network layer, TCP/IP supports the Internet Protocol (IP). The Internet Protocol (IP) is the transmission mechanism used by the TCP/IP protocols.

- IP transports data in packets called **datagrams**, each of which is transported separately.
- Datagrams can travel along different routes and can arrive out of sequence or be duplicated.

IP does not keep track of the routes and has no facility for reordering datagrams once they arrive at their destination.

Transport Layer

There is a main difference between the transport layer and the network layer. Although all nodes in a network need to have the network layer, only the two end computers need to have the transport layer.

- The network layer is responsible for sending individual datagrams from computer A to computer B; the transport layer is responsible for delivering the whole message, which is called a **segment**, from A to B.
- A segment may consist of a few or tens of **datagrams**. The segments need to be broken into datagrams and each datagram has to be delivered to the network layer for transmission.
- Since the Internet defines a different route for each datagram, the datagrams may arrive out of order and may be lost.
- The transport layer at computer B needs to wait until all of these datagrams to arrive, assemble them and make a segment out of them.

Traditionally, the transport layer was represented in the TCP/IP suite by two protocols: **User Datagram Protocol (UDP)** and **Transmission Control Protocol (TCP)**.

A new protocol called **Stream Control Transmission Protocol (SCTP)** has been introduced in the last few years.

Application Layer

The application layer in TCP/IP is equivalent to the combined session, presentation, and application layers in the OSI model.

- The application layer allows a user to access the services of our private internet or the global Internet.
- Many protocols are defined at this layer to provide services such as electronic mail file transfer, accessing the World Wide Web, and so on.
- The protocols supported in this layer are **TELNET, FTP** and **HTTP**.
-

Applications of Wireless Communication

Following is a list of applications in wireless communication:

Vehicles

Many wireless communication systems and mobility aware applications are used for following purpose:

- Transmission of music, news, road conditions, weather reports, and other **broadcast information** are received via digital audio broadcasting (DAB) with 1.5Mbit/s.
- For **personal communication,** a universal mobile telecommunications system (UMTS) phone might be available offering voice and data connectivity with 384kbit/s.
- For **remote areas,** satellite communication can be used, while the current position of the car is determined via the GPS (Global Positioning System).
- A local ad-hoc network for the fast **exchange of information** (information such as distance between two vehicles, traffic information, road conditions) in emergency situations or to help each other keep a safe distance. Local ad-hoc network with vehicles close by to prevent guidance system, accidents, redundancy.
- Vehicle data from buses, trucks, trains and high speed train can be transmitted in advance for **maintenance.**
- In ad-hoc network, car can comprise personal digital assistants (PDA), laptops, or mobile phones connected with each other using the Bluetooth technology.

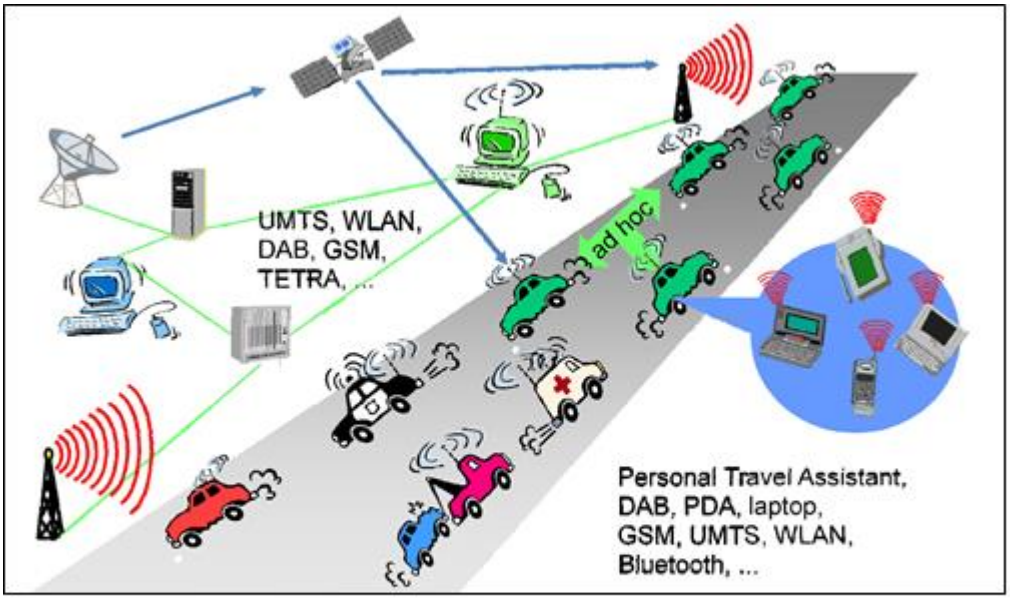


Fig: A Typical Application of Mobile Communication in Road Traffic

Emergency

Following services can be provided during emergencies:

- **Video communication:** Responders often need to share vital information. The transmission of real time situations of video could be necessary. A typical scenario includes the transmission of live video footage from a disaster area to the nearest fire department, to the police station or to the near NGOs etc.
- **Push To Talk (PTT):** PTT is a technology which allows half duplex communication between two users where switching from voice reception mode to the transmit mode takes place with the use of a dedicated momentary button. It is similar to walkie-talkie.
- **Audio/Voice Communication:** This communication service provides full duplex audio channels unlike PTT. Public safety communication requires novel full duplex speech transmission services for emergency response.
- **Real Time Text Messaging (RTT):** Text messaging (RTT) is an effective and quick solution for sending alerts in case of emergencies. Types of text messaging can be email, SMS and instant message.

Business

Travelling Salesman

- Directly access to customer files stored in a central location.
- Consistent databases for all agents
- Mobile office
- To enable the company to keep track of all the activities of their travelling employees.

In Office

- **Wi-Fi** wireless technology saves businesses or companies a considerable amount of money on installations costs.

- There is no need to physically setup wires throughout an office building, warehouse or store.
- **Bluetooth** is also a wireless technology especially used for short range that acts as a complement to Wi-Fi. It is used to transfer data between computers or cellphones.

Transportation Industries

- In transportation industries, GPS technology is used to find efficient routes and tracking vehicles.

Replacement of Wired Network

- Wireless network can also be used to replace wired network. Due to economic reasons it is often impossible to wire remote sensors for weather forecasts, earthquake detection, or to provide environmental information, wireless connections via satellite, can help in this situation.
- Tradeshows need a highly dynamic infrastructure, since cabling takes a long time and frequently proves to be too inflexible.
- Many computers fairs use WLANs as a replacement for cabling.
- Other cases for wireless networks are computers, sensors, or information displays in historical buildings, where excess cabling may destroy valuable walls or floors.

Location dependent service

It is important for an application to know something about the location because the user might need location information for further activities. Several services that might depend on the actual location can be described below:

- **Follow-on Services:**
- **Location aware services:** To know about what services (e.g. fax, printer, server, phone, printer etc.) exist in the local environment.
- **Privacy:** We can set the privacy like who should get knowledge about the location.
- **Information Services:** We can know about the special offers in the supermarket. Nearest hotel, rooms, cabs etc.

Infotainment: (Entertainment and Education)

- Wireless networks can provide information at any appropriate location.
- Outdoor internet access.
- You may choose a seat for movie, pay via electronic cash, and send this information to a service provider.
- Ad-hoc network is used for multiuser games and entertainment.

Mobile and Wireless devices

Even though many mobile and wireless devices are available, there will be many more devices in the future. There is no precise classification of such devices, by sizes, shape, weight, or computing power. The following list of given examples of mobile and wireless devices graded by increasing performance (CPU, memory, display, input devices, etc.)

Sensor: Wireless device is represented by a sensor transmitting state information. 1 example could be a switch, sensing the office door. If the door is closed, the switch transmits this information to the mobile

phone inside the office which will not accept incoming calls without user interaction; the semantics of a closed door is applied to phone calls.

Embedded Controller: Many applications already contain a simple or sometimes more complex controller. Keyboards, mouse, headsets, washing machines, coffee machines, hair dryers and TV sets are just some examples.

Pager: As a very simple receiver, a pager can only display short text messages, has a tiny display, and cannot send any messages.

Personal Digital Assistant: PDAs typically accompany a user and offer simple versions of office software (calendar, notepad, mail). The typically input device is a pen, with built-in character recognition translating handwriting into characters. Web browsers and many other packages are available for these devices.

Pocket computer: The next steps towards full computers are pocket computers offering tiny keyboards, color displays, and simple versions of programs found on desktop computers (text processing, spreadsheets etc.)

Notebook/laptop: Laptops offer more or less the same performance as standard desktop computers; they use the same software - the only technical difference being size, weight, and the ability to run on a battery. If operated mainly via a sensitive display (touch sensitive or electromagnetic), the device are also known as notepads or tablet PCs.

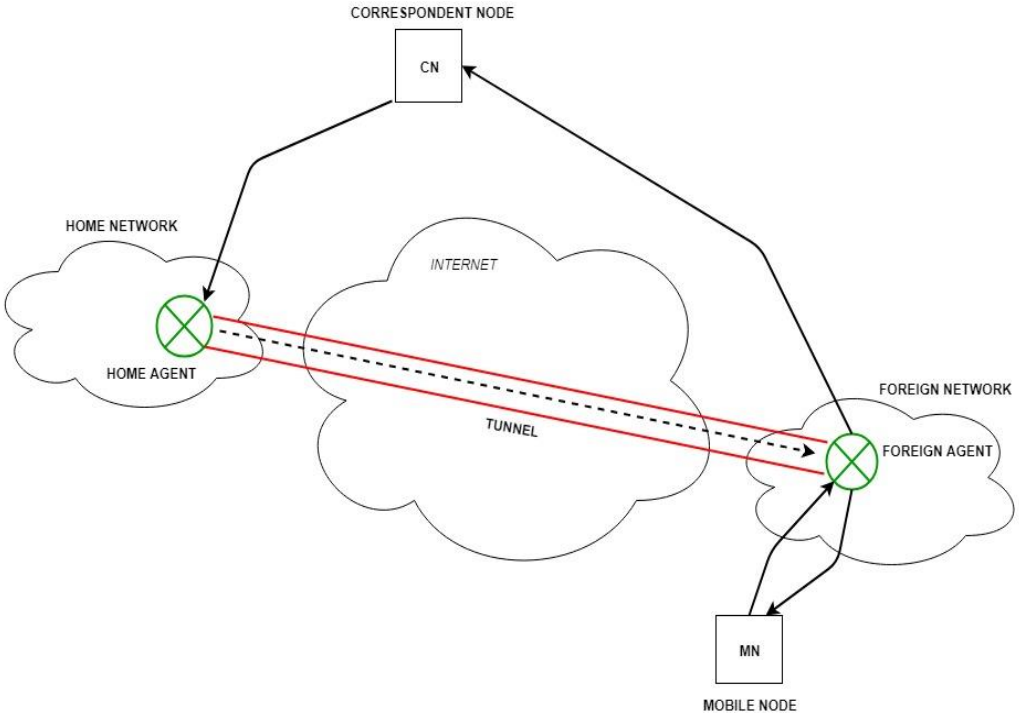
Datacasting (data broadcasting) is the broadcasting of data over a wide area via radio waves. It most often refers to supplemental information sent by television stations along with digital terrestrial television, but may also be applied to digital signals on analog TV or radio. It generally does not apply to data which is inherent to the medium, such as PSIP data which defines virtual channels for DTT or direct broadcast satellite systems; or to things like cable modem or satellite modem, which use a completely separate channel for data.

Mobile Internet Protocol (or Mobile IP)

Mobile IP is a communication protocol (created by extending Internet Protocol, IP) that allows the users to move from one network to another with the same IP address. It ensures that the communication will continue without user's sessions or connections being dropped.

Terminologies:

- **Mobile Node (MN):**
  It is the hand-held communication device that the user caries e.g. Cell phone.
- **Home Network:**
  It is a network to which the mobile node originally belongs to as per its assigned IP address (home address).
- **Home Agent (HA):**
  It is a router in home network to which the mobile node was originally connected
- **Home Address:**
  It is the permanent IP address assigned to the mobile node (within its home network).
- **Foreign Network:**
  It is the current network to which the mobile node is visiting (away from its home network).
- **Foreign Agent (FA):**
  It is a router in foreign network to which mobile node is currently connected. The packets from the home agent are sent to the foreign agent which delivers it to the mobile node.
- **Correspondent Node (CN):**
  It is a device on the internet communicating to the mobile node.
- **Care of Address (COA):**
  It is the temporary address used by a mobile node while it is moving away from its home network.



Working:

Correspondent node sends the data to the mobile node. Data packets contains correspondent node's address (Source) and home address (Destination). Packets reaches to the home agent. But now mobile node is not in the home network, it has moved into the foreign network. Foreign agent sends the care-of-address to the home agent to which all the packets should be sent. Now, a tunnel will be established between the home agent and the foreign agent by the process of tunneling.

Tunneling establishes a virtual pipe for the packets available between a tunnel entry and an endpoint. It is the process of sending a packet via a tunnel and it is achieved by a mechanism called encapsulation.

Now, home agent encapsulates the data packets into new packets in which the source address is the home address and destination is the care-of-address and sends it through the tunnel to the foreign agent. Foreign agent, on other side of the tunnel receives the data packets, decapsulates them and sends them to the mobile node. Mobile node in response to the data packets received, sends a reply in response to foreign agent. Foreign agent directly sends the reply to the correspondent node.

Key Mechanisms in Mobile IP:

1. **Agent Discovery:**
   Agents advertise their presence by periodically broadcasting their agent advertisement messages. The mobile node receiving the agent advertisement messages observes whether the message is from its own home agent and determines whether it is in the home network or foreign network.
2. **Agent Registration:**
   Mobile node after discovering the foreign agent, sends registration request (RREQ) to the foreign agent. Foreign agent in turn, sends the registration request to the home agent with the care-of-address. Home agent sends registration reply (RREP) to the foreign agent. Then it forwards the registration reply to the mobile node and completes the process of registration.
3. **Tunneling:**
   It establishes a virtual pipe for the packets available between a tunnel entry and an endpoint. It is the process of sending a packet via a tunnel and it is achieved by a mechanism called encapsulation. It takes place to forward an IP datagram from the home agent to the care-of-address. Whenever home agent receives a packet from correspondent node, it encapsulates the packet with source address as home address and destination as care-of-address.

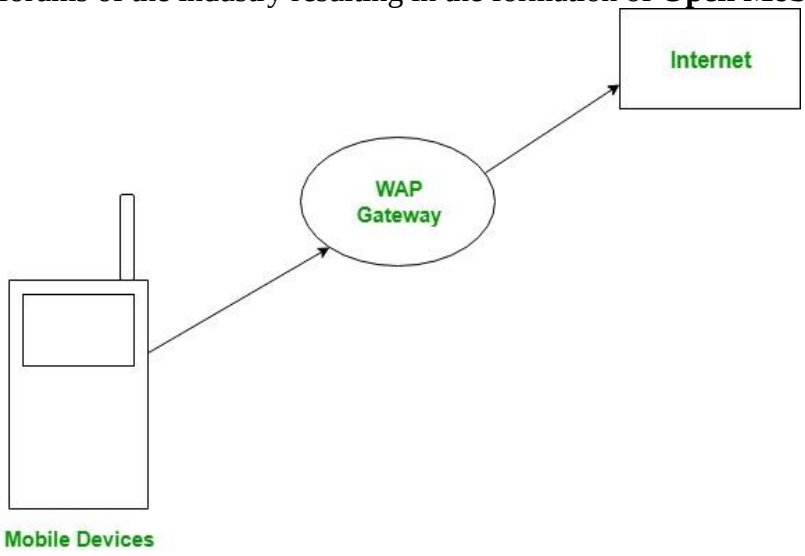
Route Optimization in Mobile IP:

The route optimization adds a conceptual data structure, the binding cache, to the correspondent node. The binding

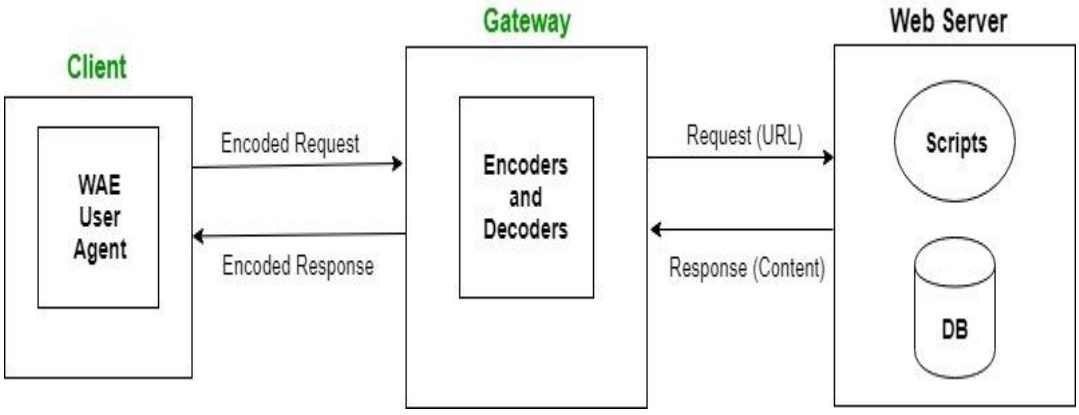
cache contains bindings for mobile node's home address and its current care-of-address. Every time the home agent receives a IP datagram that is destined to a mobile node currently away from the home network, it sends a binding update to the correspondent node to update the information in the correspondent node's binding cache. After this the correspondent node can directly tunnel packets to the mobile node.

Wireless Application Protocol

WAP stands for **Wireless Application Protocol**. It is a protocol designed for micro-browsers and it enables the access of internet in the mobile devices. It uses the mark-up language WML (Wireless Markup Language and not HTML), WML is defined as XML 1.0 application. It enables creating web applications for mobile devices. In 1998, *WAP Forum* was founded by Ericson, Motorola, Nokia and Unwired Planet whose aim was to standardize the various wireless technologies via protocols.
WAP protocol was resulted by the joint efforts of the various members of WAP Forum. In 2002, WAP forum was merged with various other forums of the industry resulting in the formation of **Open Mobile Alliance (OMA)**.

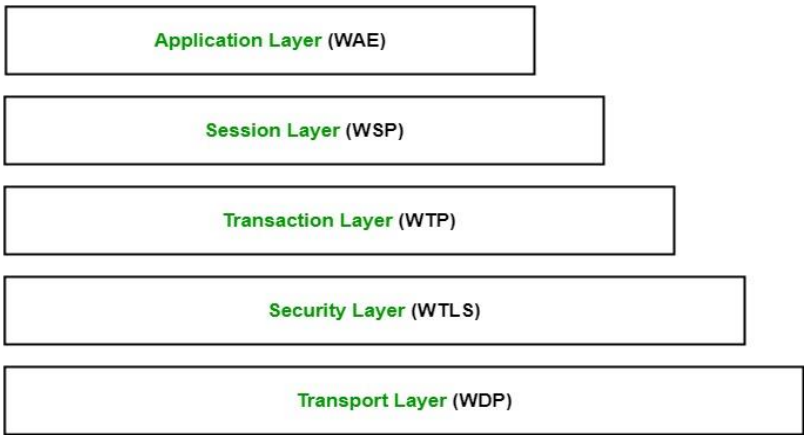


WAP Model:
The user opens the mini-browser in a mobile device. He selects a website that he wants to view. The mobile device sends the URL encoded request via network to a WAP gateway using WAP protocol.



The WAP gateway translates this WAP request into a conventional HTTP URL request and sends it over the internet. The request reaches to a specified Web server and it processes the request just as it would have processed any other request and sends the response back to the mobile device through WAP gateway in WML file which can be seen in the micro-browser.

WAP Protocol stack:



1. **Application Layer:**
   This layer contains the *Wireless Application Environment (WAE)*. It contains mobile device specifications and content development programming languages like WML.
2. **Session Layer:**
   This layer contains *Wireless Session Protocol (WSP)*. It provides fast connection suspension and reconnection.
3. **Transaction Layer:**
   This layer contains *Wireless Transaction Protocol (WTP)*. It runs on top of UDP (User Datagram Protocol) and is a part of TCP/IP and offers transaction support.
4. **Security Layer:**
   This layer contains *Wireless Transaction Layer Security (WTLS)*. It offers data integrity, privacy and authentication.
5. **Transport Layer:**
   This layer contains *Wireless Datagram Protocol*. It presents consistent data format to higher layers of WAP protocol stack.

Traditional TCP

Transmission Control Protocol (TCP) is the transport layer protocol that serves as an interface between client and server. The TCP/IP protocol is used to transfer the data packets between transport layer and network layer. Transport protocol is mainly designed for fixed end systems and fixed, wired networks. In simple terms, the traditional TCP is defined as a wired network while classical TCP uses wireless approach. Mainly TCP is designed for fixed networks and fixed, wired networks.

The main research activities in TCP are as listed below.

1. Congestion control:

During data transmission from sender to receiver, sometimes the data packet may be lost. It is not because of hardware or software problem. Whenever the packet loss is confirmed, the probable reason might be the temporary overload at some point in the transmission path. This temporary overload is otherwise called as Congestion.

Congestion is caused often even when the network is designed perfectly. The transmission speed of receiver may not be equal to the transmission speed of the sender. if the capacity of the sender is more than the capacity of output link, then the packet buffer of a router is filled and the router cannot forward the packets fast enough. The only thing the router can do in this situation is to drop some packets.

The receiver sense the packet loss but does not send message regarding packet loss to the sender. Instead, the receiver starts to send acknowledgement for all the received packets and the sender soon identifies the missing acknowledgement. The sender now notices that a packet is lost and slows down the transmission process. By this, the congestion is reduced. This feature of TCP is one of the reason for its demand even today.

2. Slow start:

The behavior TCP shows after the detection of congestion is called as slow start. The sender always calculates a congestion window for a receiver. At first the sender sends a packet and waits for the acknowledgement. Once the acknowledgement is back it doubles the packet size and sends two packets. After receiving two

acknowledgements, one for each packet, the sender again doubles the packet size and this process continues. This is called Exponential growth.
It is dangerous to double the congestion window each time because the steps might become too large. The exponential growth stops at congestion threshold. As it reaches congestion threshold, the increase in transmission rate becomes linear (i.e., the increase is only by 1). Linear increase continues until the sender notices gap between the acknowledgments. In this case, the sender sets the size of congestion window to half of its congestion threshold and the process continues.

3. Fast re-transmission:
In TCP, two things lead to a reduction of the congestion threshold. One of those is sender receiving continuous acknowledgements for the single packet. By this it can convey either of two things. One such thing is that the receiver received all the packets up to the acknowledged one and the other thing is the gap is due to packet loss. Now the sender immediately re-transmit the missing packet before the given time expires. This is called as Fast re-transmission.

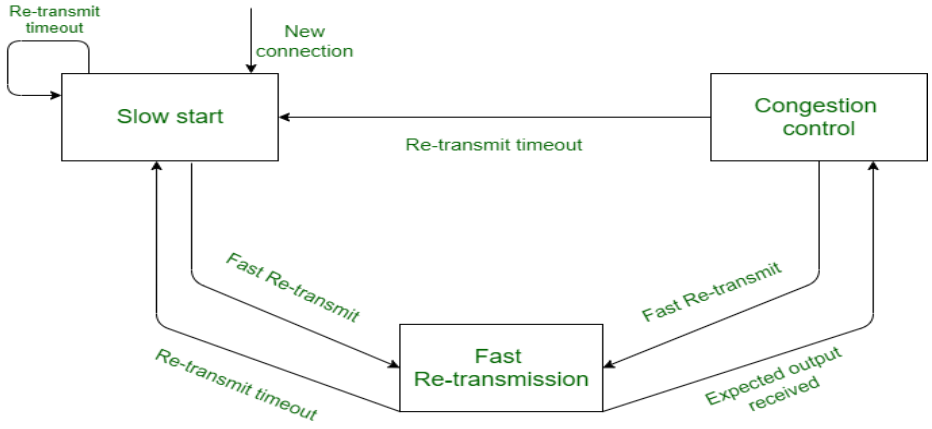


Figure: Traditional TCP

Example:
Assume that few packets of data are being transferred from sender to receiver, and the speed of sender is 2 Mbps and the speed of receiver is 1 Mbps respectively. Now the packets that are being transferred from sender sender to receiver makes a traffic jam inside the network. Due to this the network may drop some of the packets. When these packets are lost, the receiver sends the acknowledgement to the sender and the sender identifies the missing acknowledgement. This process is called as congestion control.
Now the slowstart mechanism takes up the plan. The sender slows down the packet transfer and then the traffic is slightly reduces. After sometime it puts a request to fast re-transmission through which the missing packets can be sent again as fast as possible. After all these mechanisms, the process of next packet begins.

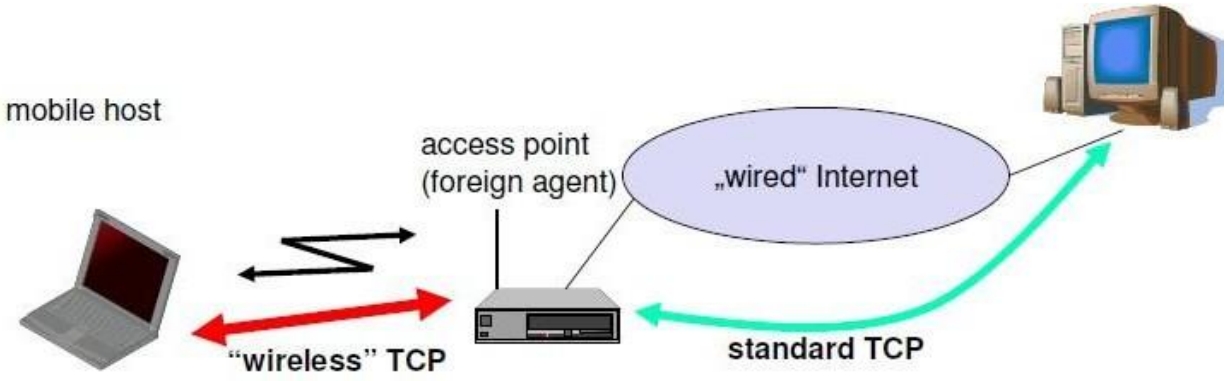
Problems with Traditional TCP in wireless environments

- Slow Start mechanism in fixed networks decreases the efficiency of TCP if used with mobile receivers orsenders.
- Error rates on wireless links are orders of magnitude higher compared to fixed fiber or copper links. This makes compensation for packet loss by TCP quite difficult.
- Mobility itself can cause packet loss. There are many situations where a soft handover from one access point to another is not possible for a mobile end-system.
- Standard TCP reacts with slow start if acknowledgements are missing, which does not help in the case of transmission errors over wireless links and which does not really help during handover. This behavior results in a severe performance degradation of an unchanged TCP if used together with wireless links or mobile nodes

Classical TCP Improvements

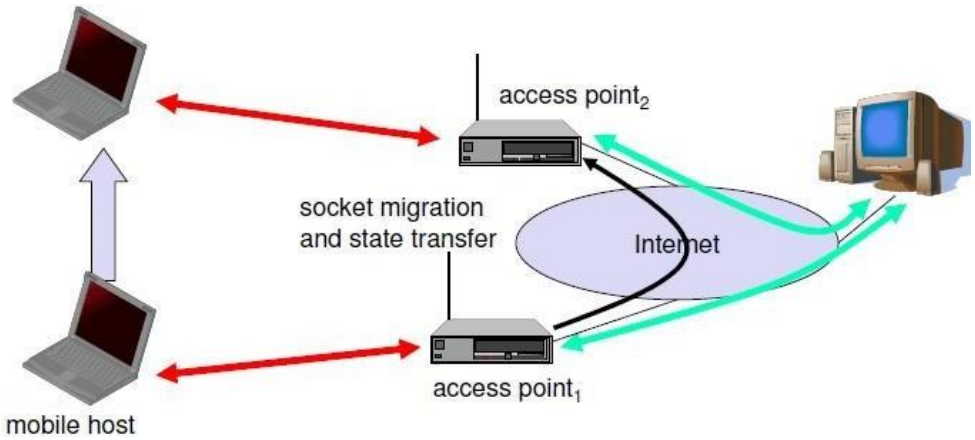
Indirect TCP (I-TCP)

Indirect TCP segments a TCP connection into a fixed part and a wireless part. The following figure shows an example with a mobile host connected via a wireless link and an access point to the 'wired' internet where the correspondent hostresides.



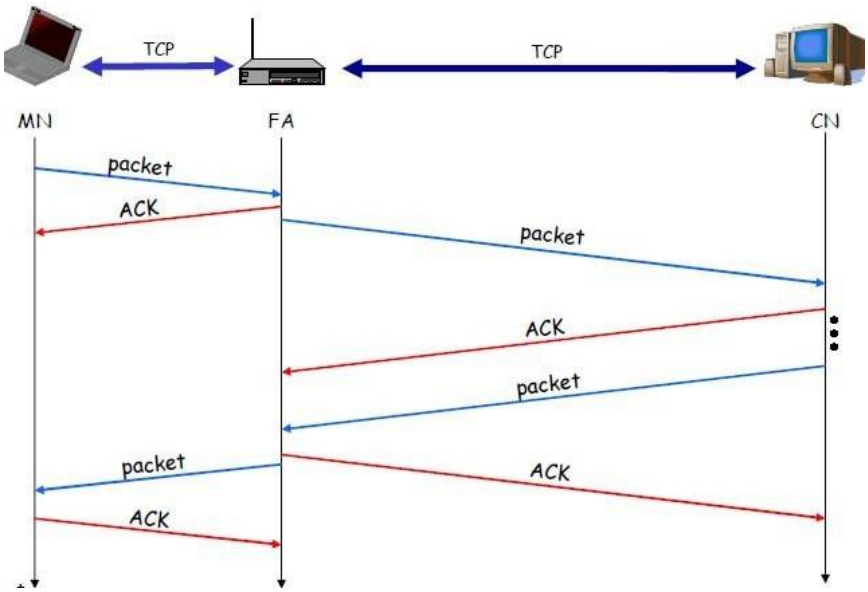
Standard TCP is used between the fixed computer and the access point. No computer in the internet recognizes any changes to TCP. Instead of the mobile host, the access point now terminates the standard TCP connection, acting as a proxy. This means that the access point is now seen as the mobile host for the fixed host and as the fixed host for the mobile host. Between the access point and the mobile host, a special TCP, adapted to wireless links, is used. However, changing TCP for the wireless link is not a requirement. A suitable place for segmenting the connection is at the foreign agent as it not only controls the mobility of the mobile host anyway and can also hand over the connection to the next foreign agent when the mobile host moves on.

The foreign agent acts as a proxy and relays all data in both directions. If CH (correspondent host) sends a packet to the MH, the FA acknowledges it and forwards it to the MH. MH acknowledges on successful reception, but this is only used by the FA. If a packet is lost on the wireless link, CH doesn't observe it and FA tries to retransmit it locally to maintain reliable data transport. If the MH sends a packet, the FA acknowledges it and forwards it to CH. If the packet is lost on the wireless link, the mobile hosts notice this much faster due to the lower round trip time and can directly retransmit the packet. Packet loss in the wired network is now handled by the foreign agent.



Socket and state migration after handover of a mobile host

During handover, the buffered packets, as well as the system state (packet sequence number, acknowledgements, ports, etc), must migrate to the new agent. No new connection may be established for the mobile host, and the correspondent host must not see any changes in connection state. Packet delivery in I-TCP is shown below:



~~Advantages of I-TCP~~

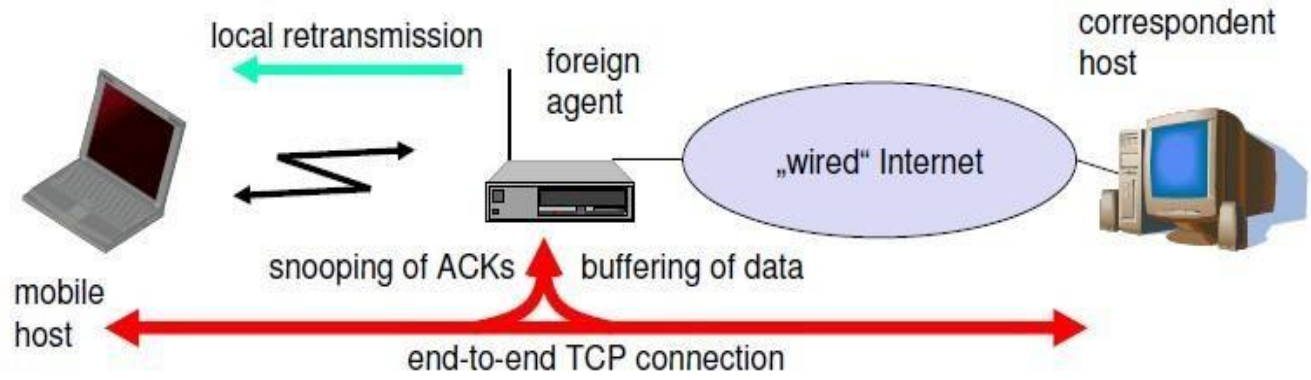
- No changes in the fixed network necessary, no changes for the hosts (TCP protocol) necessary, all current optimizations to TCP still work
- Simple to control, mobile TCP is used only for one hop between, e.g., a foreign agent and mobile host
  1. transmission errors on the wireless link do not propagate into the fixed network
  2. therefore, a very fast retransmission of packets is possible, the short delay on the mobile hop sknown
- It is always dangerous to introduce new mechanisms in a huge network without knowing exactly how they behave.
  - New optimizations can be tested at the last hop, without jeopardizing the stability of the Internet.
- It is easy to use different protocols for wired and wireless networks.

~~Disadvantages of I-TCP~~

- Loss of end-to-end semantics:- an acknowledgement to a sender no longer means that a receiver really has received a packet, foreign agents might crash.
- Higher latency possible:- due to buffering of data within the foreign agent and forwarding to a new foreign agent
- Security issue:- The foreign agent must be a trusted entity

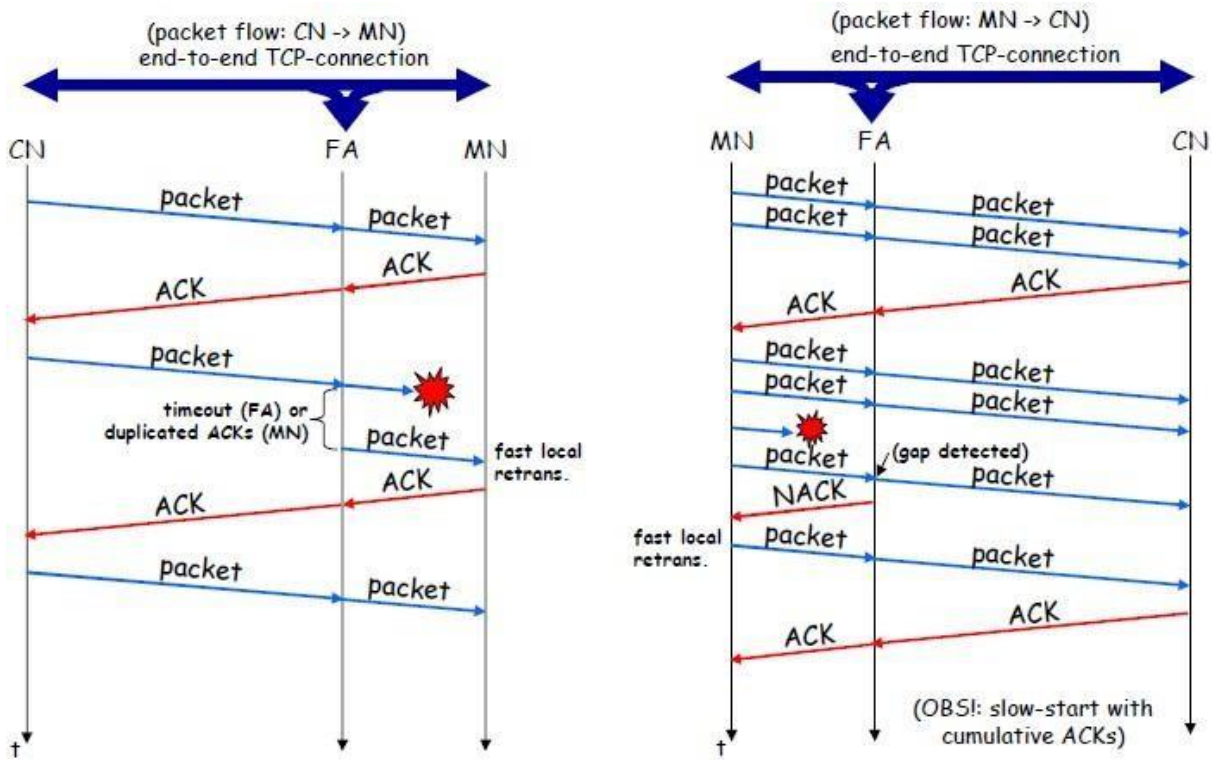
Snooping TCP

The main drawback of I-TCP is the segmentation of the single TCP connection into two TCP connections, which loses the original end-to-end TCP semantic. A new enhancement, which leaves the TCP connection intact and is completely transparent, is Snooping TCP. The main function is to buffer data close to the mobile host to perform fast local retransmission in case of packet loss.



Snooping TCP as a transparent TCP extension

Here, the foreign agent buffers all packets with **destination mobile host** and additionally 'snoops' the packet flow in both directions to recognize acknowledgements. The foreign agent buffers every packet until it receives an acknowledgement from the mobile host. If the FA does not receive an acknowledgement from the mobile host within a certain amount of time, either the packet or the acknowledgement has been lost. Alternatively, the foreign



agent could receive a duplicate ACK which also shows the loss of a packet. Now, the FA retransmits the packet directly from the buffer thus performing a faster retransmission compared to the CH. For transparency, the FA does not acknowledge data to the CH, which would violate end-to-end semantic in case of a FA failure. The foreign agent can filter the duplicate acknowledgements to avoid unnecessary retransmissions of data from the correspondent host. If the foreign agent now crashes, the time-out of the correspondent host still works and triggers a retransmission. The foreign agent may discard duplicates of packets already retransmitted locally and acknowledged by the mobile host. This avoids unnecessary traffic on the wireless link.For data transfer from the mobile host with **destination correspondent host**, the FA snoops into the packet stream to detect gaps in the sequence numbers of TCP. As soon as the foreign agent detects a missing packet, it returns a negative acknowledgement (NACK) to the mobile host. The mobile host can now retransmit the missing packet immediately. Reordering of packets is done automatically at the correspondent host by TCP.

Snooping TCP: Packet delivery

~~Advantages of snooping TCP~~:

- The end-to-end TCP semantic is preserved.
- Most of the enhancements are done in the foreign agent itself which keeps correspondent host unchanged.
- Handover of state is not required as soon as the mobile host moves to another foreign agent. Even though packets are present in the buffer, time out at the CH occurs and the packets are transmitted to the new COA.
- No problem arises if the new foreign agent uses the enhancement or not. If not, the approach automatically falls back to the standard solution.

Disadvantages of snooping TCP

- Snooping TCP does not isolate the behavior of the wireless link as well as I-TCP. Transmission errors may propagate till CH.
- Using negative acknowledgements between the foreign agent and the mobile host assumes additional mechanisms on the mobile host. This approach is no longer transparent for arbitrary mobile hosts.
- Snooping and buffering data may be useless if certain encryption schemes are applied end- to-end between the correspondent host and mobile host. If encryption is used above the transport layer, (eg. SSL/TLS), snooping TCP can be used.

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 but forwards the ACKs from the MH.
- If the MH is disconnected, it avoids useless retransmissions, slow starts or breaking connections 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 a valid 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 resume operation at exactly the same point where it had been forced to stop. For TCP time simply does not 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.

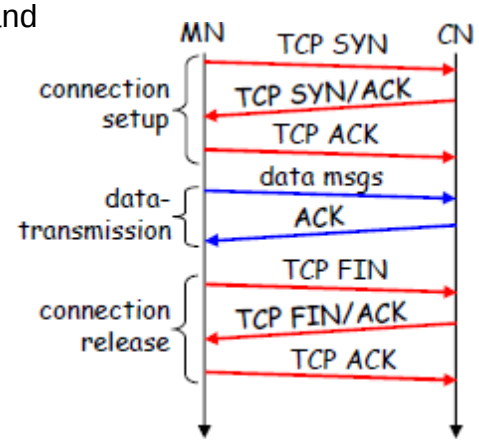
Selective retransmission

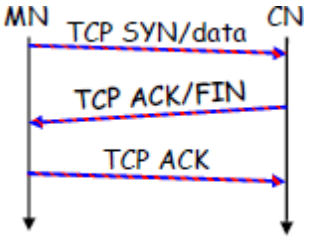
A very useful extension of TCP is the use of selective retransmission. TCP acknowledgements are cumulative, i.e., they acknowledge in-order receipt of packets up to a certain packet. A single acknowledgement confirms reception of all packets upto a certain packet. If a single packet is lost, the sender 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.

Classical Enhancements to TCP for mobility: A comparison

Approach	Mechanism	Advantages	Disadvantages
Indirect TCP	splits TCP connection into two connections	isolation of wireless link, simple	loss of TCP semantics, higher latency at handover
Snooping TCP	"snoops" data and acknowledgements, local retransmission	transparent for end-to-end connection, MAC integration possible	problematic with encryption, bad isolation of wireless link
M-TCP	splits TCP connection, chokes sender via window size	Maintains end-to-end semantics, handles long term and frequent disconnections	Bad isolation of wireless link, processing overhead due to bandwidth management
Fast retransmit/ fast recovery	avoids slow-start after roaming	simple and efficient	mixed layers, not transparent
Transmission/ time-out freezing	freezes TCP state at disconnect, resumes after reconnection	independent of content or encryption, works for longer interrupts	changes in TCP required, MAC dependant
Selective retransmission	retransmit only lost data	very efficient	slightly more complex receiver software, more buffer needed
Transaction oriented TCP	combine connection setup/release and data transmission	Efficient for certain applications	changes in TCP required, not transparent