Mobile Architecture and Programming
UNIT 1

Gauri Rao

- History, Types, Benefits, Application, Evolution, Security Concern regarding Mobile Computing, Different Propagation Modes, Wireless Architecture and its types, needs of mobile user, Mobile Development Importance, Survey of mobile based application development.

History

- Mobile Computing refers a technology that allows transmission of data, voice and video via a computer or any other wireless enabled device. It is free from having a connection with a fixed physical link. It facilitates the users to move from one physical location to another during communication. Mobile Computing is a technology that provides an environment that enables users to transmit data from one device to another device without the use of any physical link or cables.

- In other words, you can say that mobile computing allows transmission of data, voice and video via a computer or any other wireless-enabled device without being connected to a fixed physical link. In this technology, data transmission is done wirelessly with the help of wireless devices such as mobiles, laptops etc. This is only because of Mobile Computing technology that you can access and transmit data from any remote locations without being present there physically. Mobile computing technology provides a vast coverage diameter for communication. It is one of the fastest and most reliable sectors of the computing technology field.

The concept of Mobile Computing can be divided into three parts:

- Mobile Communication
- Mobile Hardware
- Mobile Software

A communication device can exhibit one of the following characterstics:

- Fixed and Wired
- Fixed and Wireless
- Mobile and Wired
- Mobile and Wireless

1. Fixed and Wired: In Fixed and Wired configuration, the devices are fixed at a position, and they are connected through a physical link to communicate with other devices.

For Example, Desktop Computer.

2. Fixed and Wireless: In Fixed and Wireless configuration, the devices are fixed at a position, and they are connected through a wireless link to make communication with other devices.

For Example, Communication Towers, WiFi router

3. Mobile and Wired: In Mobile and Wired configuration, some devices are wired, and some are mobile. They altogether make communication with other devices.

For Example, Laptops. User carry laptop which is connected to company network via modem

4. Mobile and Wireless: In Mobile and Wireless configuration, the devices can communicate with each other irrespective of their position. They can also connect to any network without the use of any wired device.

For Example, GSM, WiFi Dongle

Mobile Hardware

Mobile hardware consists of mobile devices or device components that can be used to receive or access the service of mobility. Examples of mobile hardware can be smartphones, laptops, portable PCs, tablet PCs, Personal Digital Assistants, etc. These devices are inbuilt with a receptor medium that can send and receive signals. These devices are capable of operating in full-duplex. It means they can send and receive signals at the same time. They don't have to wait until one device has finished communicating for the other device to initiate communications.

Mobile Software

Mobile software is a program that runs on mobile hardware. This is designed to deal capably with the characteristics and requirements of mobile applications. This is the operating system for the appliance of mobile devices. In other words, you can say it the heart of the mobile systems. This is an essential component that operates the mobile device.

Applications of Mobile Computing

- Web or Internet access.
- Global Position System (GPS).
- Emergency services.
- Entertainment services.
- Educational services.

Emergencies

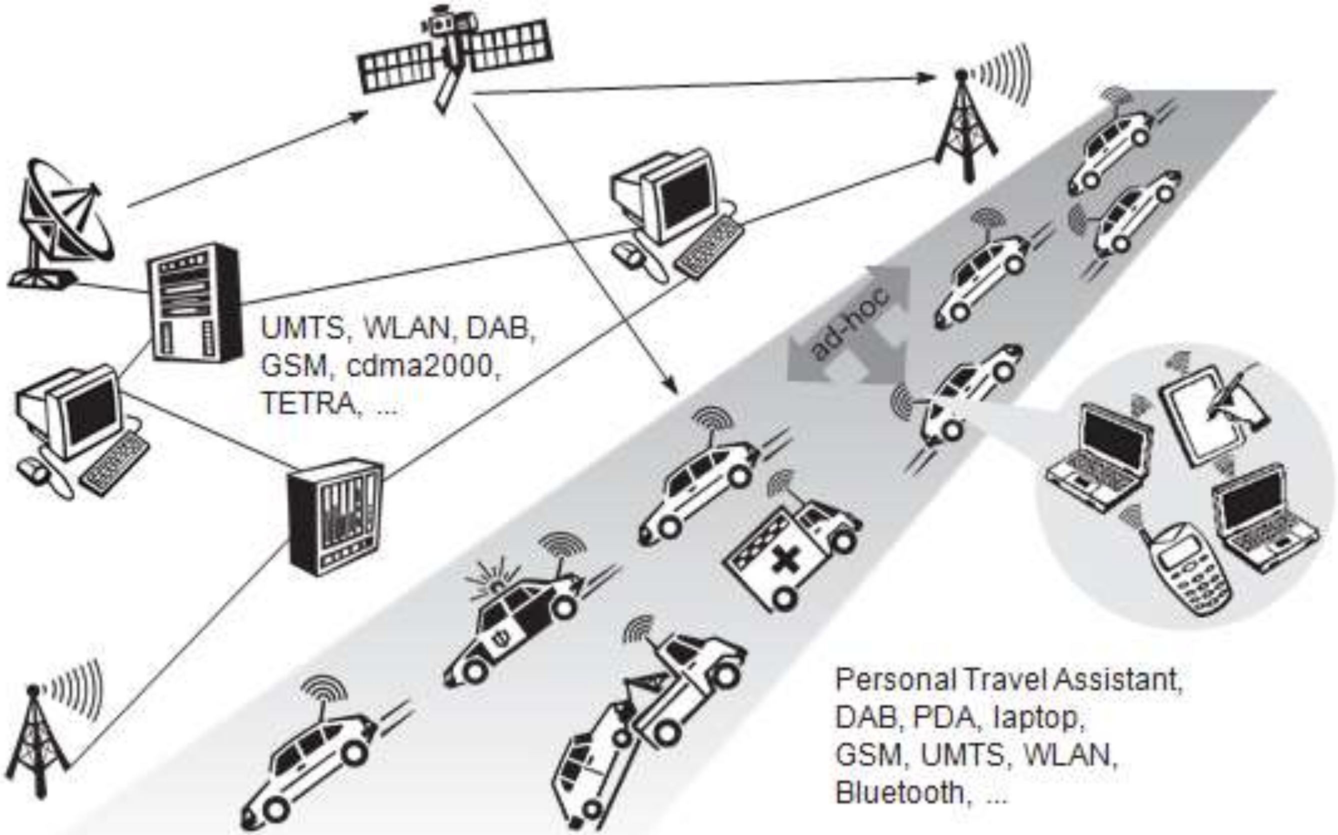
Just imagine the possibilities of an ambulance with a high-quality wireless connection to a hospital. Vital information about injured persons can be sent to the hospital from the scene of the accident. All the necessary steps for this particular type of accident can be prepared and specialists can be 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. The breakdown of all cabling not only implies the failure of the standard wired telephone system, but also the crash of all mobile phone systems requiring base stations!

Vehicles

Today's cars already comprise some, but tomorrow's cars will comprise many wireless communication systems and mobility aware applications. 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. For remote areas, satellite communication can be used, while 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. Cars with this technology are already available.

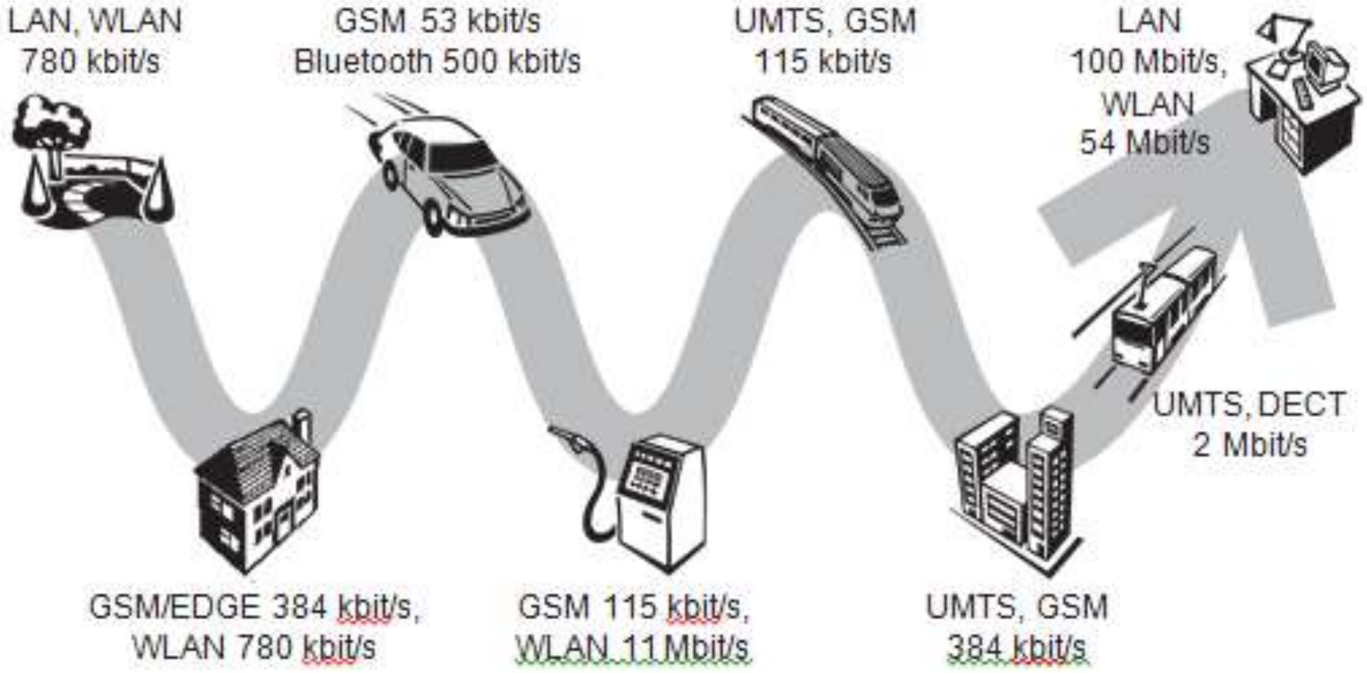
In the future, cars will also inform other cars about accidents via the ad-hoc network to help them slow down in time, even before a driver can recognize an accident. Buses, trucks, and trains are already transmit- ting maintenance and logistic information to their home base, which helps to improve organization (fleet management), and saves time and money.


UMTS, WLAN, DAB,
GSM, cdma2000,
TETRA, ...
ad-hoc
Personal Travel Assistant,
DAB, PDA, laptop,
GSM, UMTS, WLAN,
Bluetooth, ...

Business

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 data- bases 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.

Mobile and wireless services- best connected



At home, the laptop connects via a WLAN or LAN and DSL to the Internet. Leaving home requires a handover to another technology, e.g., to an enhanced version of GSM, as soon as the WLAN coverage ends. Due to interference and other factors, data rates drop while cruising at higher speed. Gas stations may offer WLAN hot spots as well as gas. Trains already offer support for wireless connectivity. Several more handovers to different technologies might be necessary before reaching the office.

Internet everywhere?

- Not without wireless networks! Imagine a travel guide for a city. Static information might be loaded via CD-ROM, DVD, or even at home via the Internet. But 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. You may choose a seat, pay via electronic cash, and send this information to a service provider

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.

1. Sensor: A very simple wireless device is represented by a sensor transmitting state information. One example could be a switch sensing the office door. If the door is closed, the switch transmits this 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.

2. Embedded controllers: Many appliances already contain a simple or some- times more complex controller. Keyboards, mice, headsets, washing machines, coffee machines, hair dryers and TV sets are just some examples. Why not have the hair dryer as a simple mobile and wireless device (from a communication point of view) that is able to communicate with the mobile phone? Then the dryer would switch off as soon as the phone starts ringing that would be a nice application!

3. Pager: As a very simple receiver, a pager can only display short text messages, has a tiny display, and cannot send any messages. Pagers can even be integrated into watches. The tremendous success of mobile phones, has made the pager virtually redundant in many countries. Short messages have replaced paging. The situation is somewhat different for emergency services where it may be necessary to page a larger number of users reliably within short time.

4. Mobile phones: The traditional mobile phone only had a simple black and white text display and could send/receive voice or short messages. Today, mobile phones migrate more and more toward PDAs. Mobile phones with full color graphic display, touch screen, and Internet browser are easily available.

5. Personal digital assistant: PDAs typically accompany a user and offer simple versions of office software (calendar, note-pad, mail). The typical input device is a pen, with built-in character recognition translating hand-writing into characters. Web browsers and many other software packages are available for these devices.

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

7. Notebook/laptop: Finally, 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 devices are also known as notepads or tablet PCs.

Evolution /Generations

1G

- This is the first generation of wireless telephone technology, mobile telecommunications, which was launched in Japan by NTT in 1979.
- The main technological development in this generation that distinguished the First Generation mobile phones from the previous generation was the use of multiple cell sites, and the ability to transfer calls from one site to the next site as the user travelled between cells during a conversation.
- It uses analog signals.
- It allows the voice calls in one country.

Disadvantages

- Poor quality of voice
- Poor life of Battery
- Size of phone was very large
- No security
- Capacity was limited
- Poor handoff reliability

2G

- This is the second generation of mobile telecommunication was launched in Finland in 1991.
- It was based on GSM standard.
- It enables data transmission like as text messaging (SMS - Short Message Service), transfer or photos or pictures (MMS ? Multimedia Messaging Service), but not videos.
- The later versions of this generation, which were called 2.5G using GPRS (General Packet Radio Service) and 2.75G using EDGE (Enhanced data rates for GSM Evolution) networks.
- It provides better quality and capacity.

Disadvantages

- Unable to handle complex data such as Video
- Requires strong digital signals

3G

- 3G is the third generation was introduced in early 2000s.
- The transmission of data was increased up to 2Mbits/s, which allows you to sending or receiving large email messages.
- The main difference between 3G and 2G is the use of packet switching rather than circuit switching for data transmission.
- Faster communication
- High speed web or more security
- Video conferencing
- 3D gaming
- TV streaming, Mobile TV, phone calls etc. are the features of 3G.

Disadvantages

- Costly
- Requirement of high bandwidth
- Expensive 3G phones
- Size of cell phones was very large.

- 4G is the fourth generation of mobile telecommunication which was appeared in 2010.
- It was based on LTE (Long Term Evolution) and LTE advanced standards.
- Offer a range of communication services like video calling, real time language translation and video voice mail.
- It was capable of providing 100 Mbps to 1Gbps speed.
- High QoS (Quality of Service) and High security.

The basic term used to describe 4G technology is MAGIC. Where :

- M - Mobile multimedia
- A - Anytime anywhere
- G - Global mobility support
- I - Integrated wireless solution
- C - Customized personal service

Disadvantages

- Uses more battery
- Difficult to implement
- Expensive equipment are required

5G

- It is referred to fifth generation wireless connection which will be probably implemented by 2020, or even some years earlier.
- Machine to machine communication can be possible in 5G.
- 5G will be able to performs Internet of Things (IoT) for smart home and smart city, connected cars etc.
- This generation will be based on lower cost, low battery consumption and lower latency than 4G equipment.
- There will be much faster transmission rate of data to the previous versions. Thus the speed of 5G will be 1Gbit/s.

Survey of mobile based application development.

The following are the key drivers of mobile apps:

- Innovation in mobile space such as proliferation of smart phones, higher bandwidths offered by 3G (Third generation) and 4G (Fourth generation) technologies are coupled with higher capacity storage technologies with higher speed chips would keep powering mobile devices.
- Consumer behaviour: Customers are more used to mobile devices and is easy for them to access information on the move.
- Personalized content delivery: Enterprise can leverage the location and sensors to offer more contextualized, relevant and personalized content, offers and advertisements.

- Mobile ecosystem: An explosive growth in Mobile Applications stores such as Apple store, Google Play store, Windows marketplace store was coupled with availability of games, utilities and other apps.
- Social Networking: With the popularity of web 2.0 and social media technologies such as Facebook, Twitter users are increasingly using the location based features in the social media platforms.

Advantages of Mobile Computing Technology

- Enhanced Productivity

We can use mobile devices in various companies, which can reduce the time and cost for clients and themselves and enhance the productivity of the company.

- Location Flexibility

This technology facilitates users to work efficiently and effectively from whichever location they want to do their tasks. So, a user can work without being in a fixed position. This facility makes them able to carry out numerous tasks at the same time and also benefitted the company.

- Saves Time

The location flexibility facility of mobile computing makes it time-saving. It cuts down the time consumed or wasted while traveling from different locations or to the office and back. It facilitates users to access all the essential documents and files over a secure channel and work on their computers. It has also reduced many unnecessary incurred expenses.

- Support Cloud Computing

By using mobile Computing technology, you can save your documents on an online server and access them anytime and anywhere when you have an internet connection. You can access these files on several mobiles simultaneously.

- Entertainment

Nowadays, mobile devices can be used as an entertainment source. They provide a lot of entertainment facilities to their users.

- Besides the above advantages, it provides some other facilities such as Device Mobility, Simple Framework, easy and simple infrastructure etc.

Issues occurred in Mobile Computing.

Mobile computing technology provides vast features from mobility to portability and from cloud to productivity. But, along with these advantages, you can face specific issues while using mobile computing technology.

1.Costly due to Wireless Medium

The Mobile computing technology mainly focuses on wireless infrastructure, so the cost of implementation is always high. It also faces issues like efficiency, delays and security, which we have to consider in project establishment.

2. Issue due to Device Mobility

The device mobility is one of the most significant advantages of mobile computing technology. But, it is one of its major issues too. To obtain the device mobility feature of mobile computing technology, we have to install the highest standards' types of equipment. So, whenever the mobile device changes its environment, we have to restructure its configuration environment.

We have to configure the device mobility feature according to the location, environment and surroundings of a mobile device regularly.

3. Security Issues in Mobile Computing

This is undoubtedly the biggest and one of the most discussed issues we face in mobile computing technology. It arises due to the shared medium ability of mobile computing. Security Issues

Mobile VPNs are not very safe to connect, and there is always a chance of security concerns.

The most significant security issues are:

- Physical Security or Data Security
- System Security or Network Security

These issues can be resolved by using some common tactics. These issues are:

- Using VPN technology
- Using Cryptography & Network Security in your project
- Use of Firewall technology in the project

4. Poor Quality of Connectivity

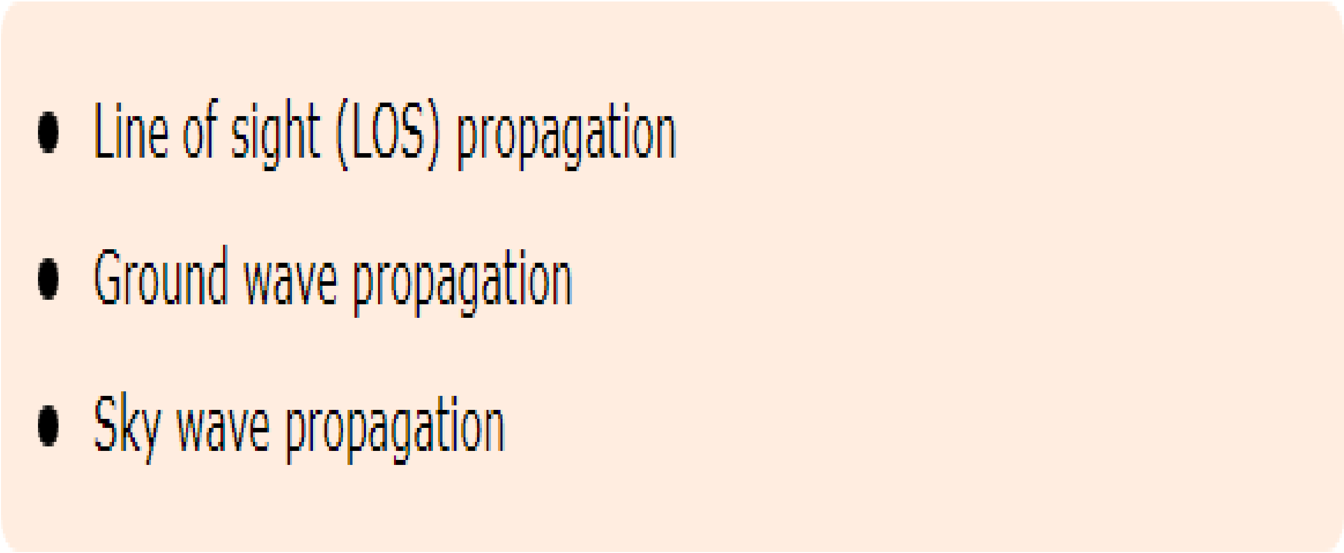
This is one of the biggest disadvantages because if you are not near any of these connection providers, your access to the internet may be minimal.

5. High on Power Consumption

These devices run on batteries that do not tend to long-lasting. So, if in a situation where there is no source of power for charging, then that will be a failure.

- Besides the above, there are also some disadvantages such as low data transmission rates, High data losses, Frequent network issues etc.

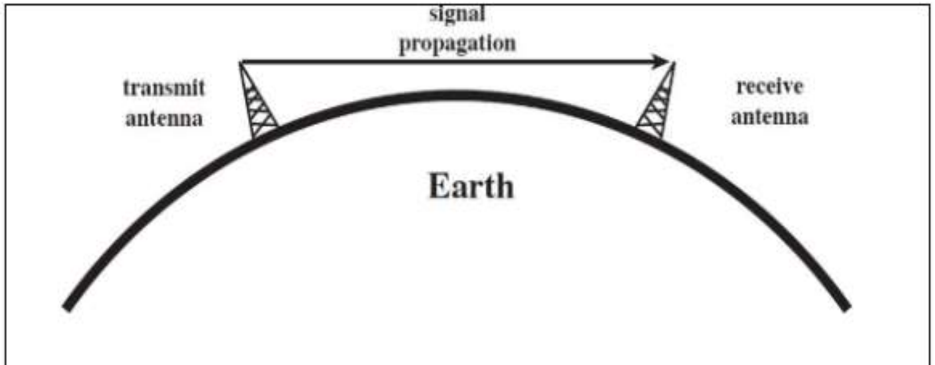
Propagation modes



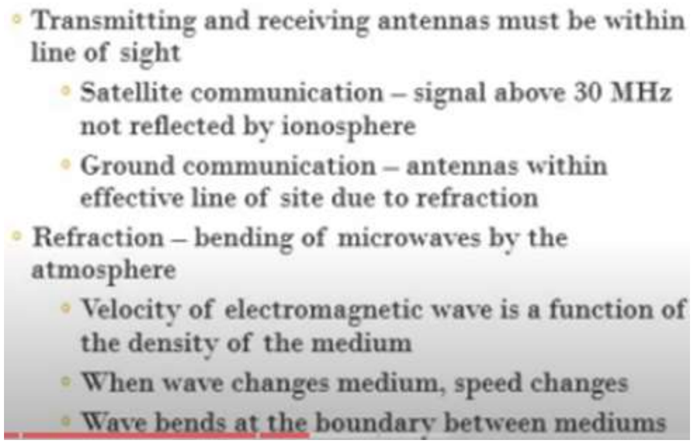
- Line of sight (LOS) propagation
- Ground wave propagation
- Sky wave propagation

1. Line of Sight (LOS) Propagation

- Among the modes of propagation, this line-of-sight propagation is the one, which we commonly notice. In the line-of-sight communication, as the name implies, the wave travels a minimum distance of sight. Which means it travels to the distance up to which a naked eye can see. Now what happens after that? We need to employ an amplifier cum transmitter here to amplify the signal and transmit again. The line-of-sight propagation will not be smooth if there occurs any obstacle in its transmission path. As the signal can travel only to lesser distances in this mode, this transmission is used for infrared or microwave transmissions.

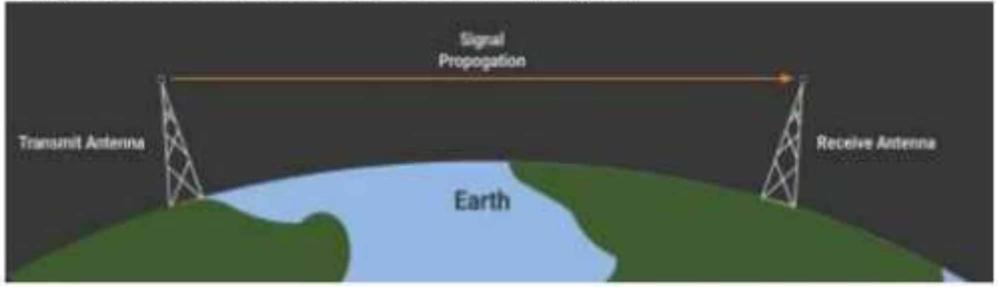


Line-of-sight (LOS) propagation (above 30 MHz)



- Transmitting and receiving antennas must be within line of sight
  - Satellite communication – signal above 30 MHz not reflected by ionosphere
  - Ground communication – antennas within effective line of site due to refraction
- Refraction – bending of microwaves by the atmosphere
  - Velocity of electromagnetic wave is a function of the density of the medium
  - When wave changes medium, speed changes
  - Wave bends at the boundary between mediums

- In the line of sight propagation the electromagnetic wave with **very high frequency** is transmitted in the **“straight-line”** direction from the transmitter antenna to the receiver antenna.
- In this way, the electromagnetic wave does not get affected by the curvature of the earth because the transmitter and receiver antenna are either tall or close enough to maintain the line of sight.



2. Ground wave propagation

- Ground wave propagation of the wave follows the contour of earth. Such a wave is called as direct wave. The wave sometimes bends due to the Earth’s magnetic field and gets reflected to the receiver. Such a wave can be termed as reflected wave. The wave when propagates through the Earth’s atmosphere is known as ground wave. The direct wave and reflected wave together contribute the signal at the receiver station. When the wave finally reaches the receiver, the lags are cancelled out. In addition, the signal is filtered to avoid distortion and amplified for clear output.

- In the ground propagation, the electromagnetic wave is propagated close to the atmosphere i.e. between ionosphere and earth surface.
- These electromagnetic waves are the **low-frequency waves** transmitted in all direction and
- The wave follows the **curvature of the earth**.
- More the power of the signal longer the distance it travels.

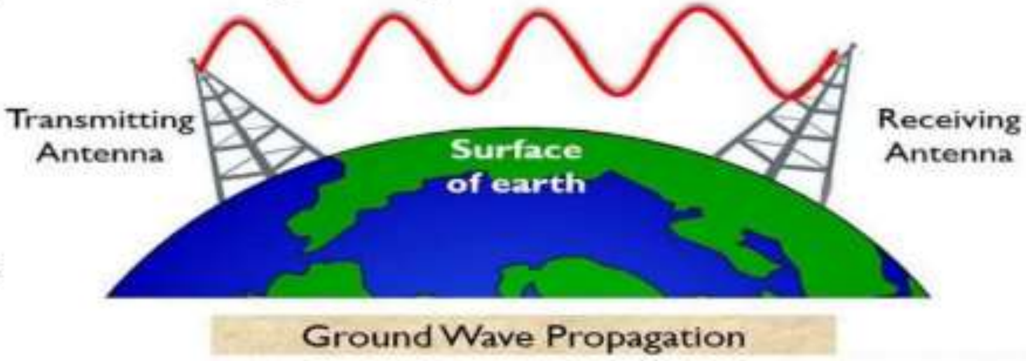


Figure 2: Ground Propagation

Advantages of Ground Wave Propagation

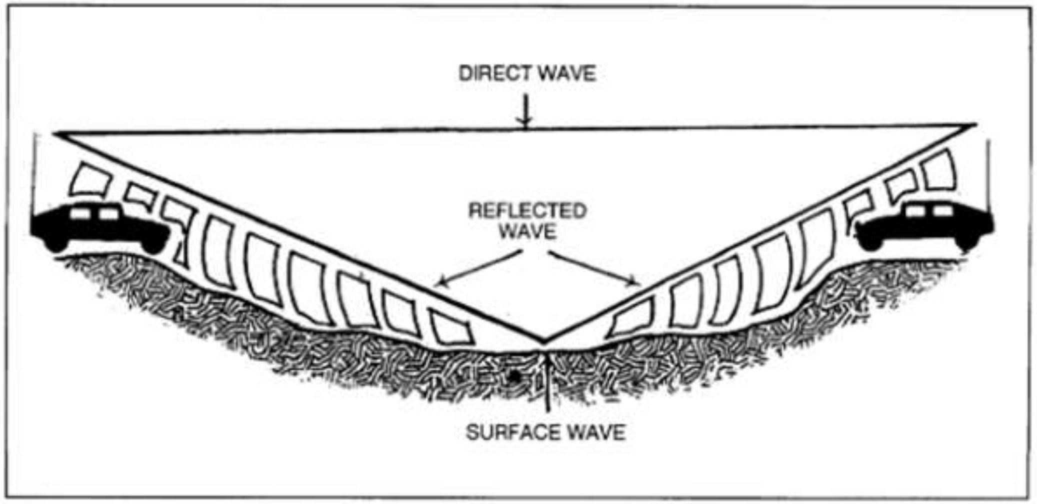
- The factor of **large wavelength** supports efficient bending thereby causing **less attenuation**.
- The signal loss due to **atmospheric conditions** is less.

Disadvantages of Ground Wave Propagation

- The **distance** between the transmitting and receiving end must not be very large.
- The operational **frequency range** is limited to up to **2MHz**.

Applications

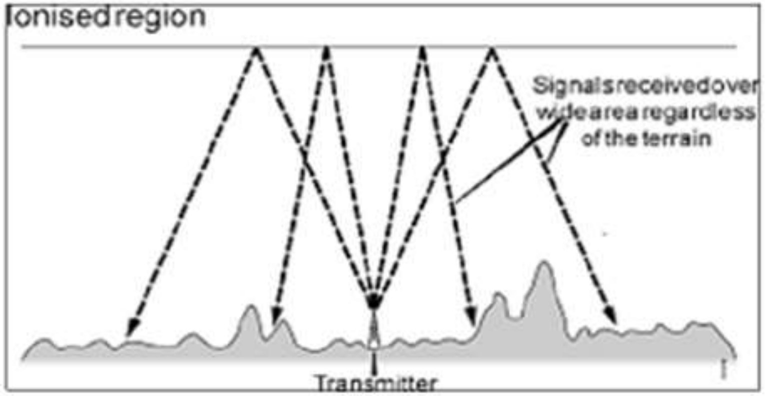
- Television signal broadcasting, target detection for military purposes, radio signal transmission, and in all such applications those require a distance of operation in the local range.

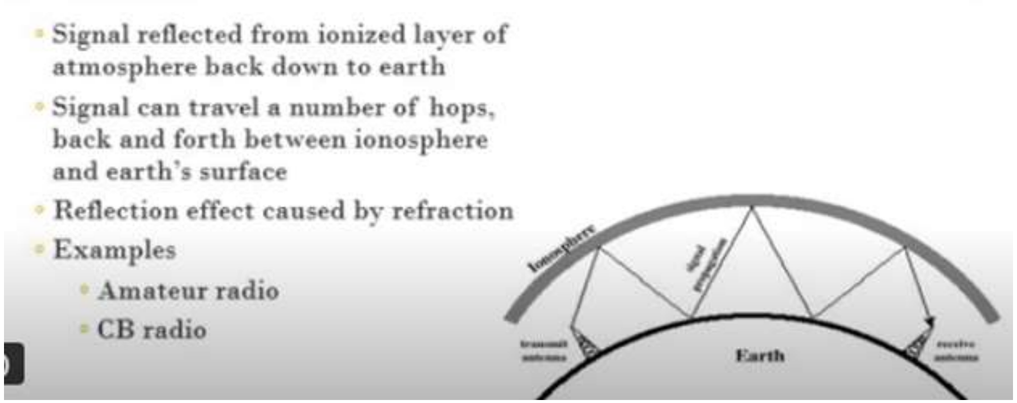


- Follows contour of the earth
- Can Propagate considerable distances
- Frequencies up to 2 MHz
- Example
- AM radio

3. Sky Wave Propagation

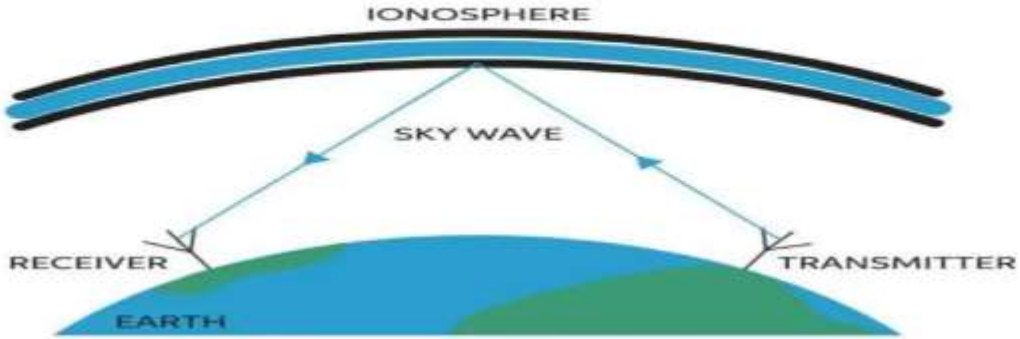
- Sky wave propagation is preferred when the wave has to travel a longer distance. Here the wave is projected onto the sky and it is again reflected back onto the earth.
- Here the waves are shown to be transmitted from one place and where it is received by many receivers. Hence, it is an example of broadcasting.
- The waves, which are transmitted from the transmitter antenna, are reflected from the ionosphere. It consists of several layers of charged particles ranging in altitude from 30- 250 miles above the surface of the earth. Such a travel of the wave from transmitter to the ionosphere and from there to the receiver on Earth is known as Sky Wave Propagation. Ionosphere is the ionized layer around the Earth's atmosphere, which is suitable for sky wave propagation.





- Signal reflected from ionized layer of atmosphere back down to earth
- Signal can travel a number of hops, back and forth between ionosphere and earth's surface
- Reflection effect caused by refraction
- Examples
  - Amateur radio
  - CB radio

- In sky propagation, the electromagnetic wave with a **higher frequency** is radiated up in the ionosphere from the source transmitter antenna which gets reflected back to the earth surface.
- In sky propagation, the electromagnetic wave travel a **longer distance** and it can be received thousands of Kilometres away from the transmitter.



Advantages of sky wave propagation

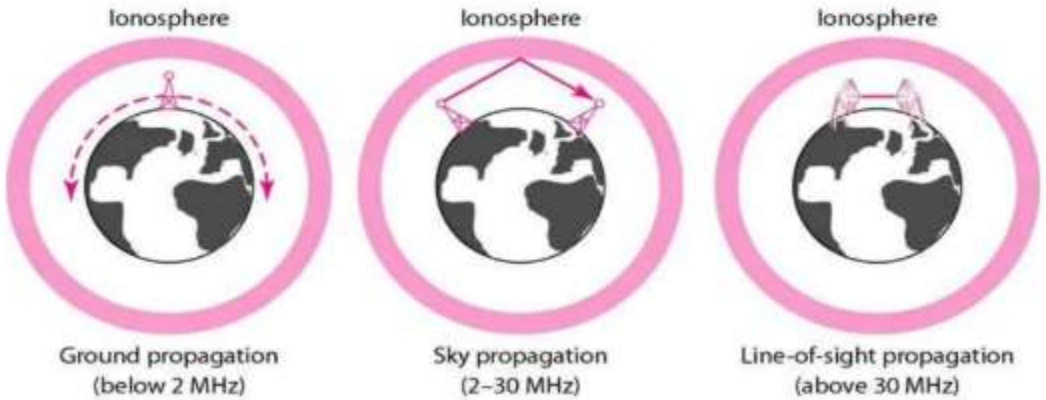
- The **frequency range** of operation is considerably **high**.
- Attenuation due to atmospheric conditions is less.
- It supports **large distance** propagation.

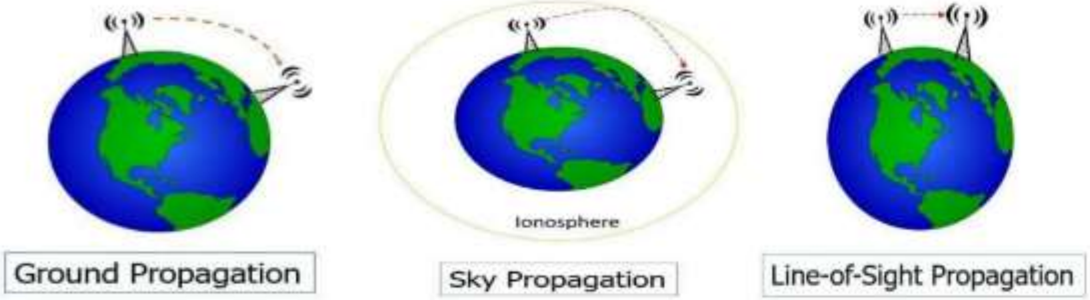
Disadvantages of sky wave propagation

- Long-distance propagation requires **large-sized antennas**.
- Due to the presence of the ionosphere, near and far during night and day respectively there **exist variation in signal transmission in day and night.**

Applications

- Sky wave propagation is widely used in **mobile and satellite communications** as it needs suitable atmospheric conditions.


Ionosphere
Ionosphere
Ionosphere
Ground propagation
(below 2 MHz)
Sky propagation
(2–30 MHz)
Line-of-sight propagation
(above 30 MHz)


Ionosphere
Ground Propagation
Sky Propagation
Line-of-Sight Propagation

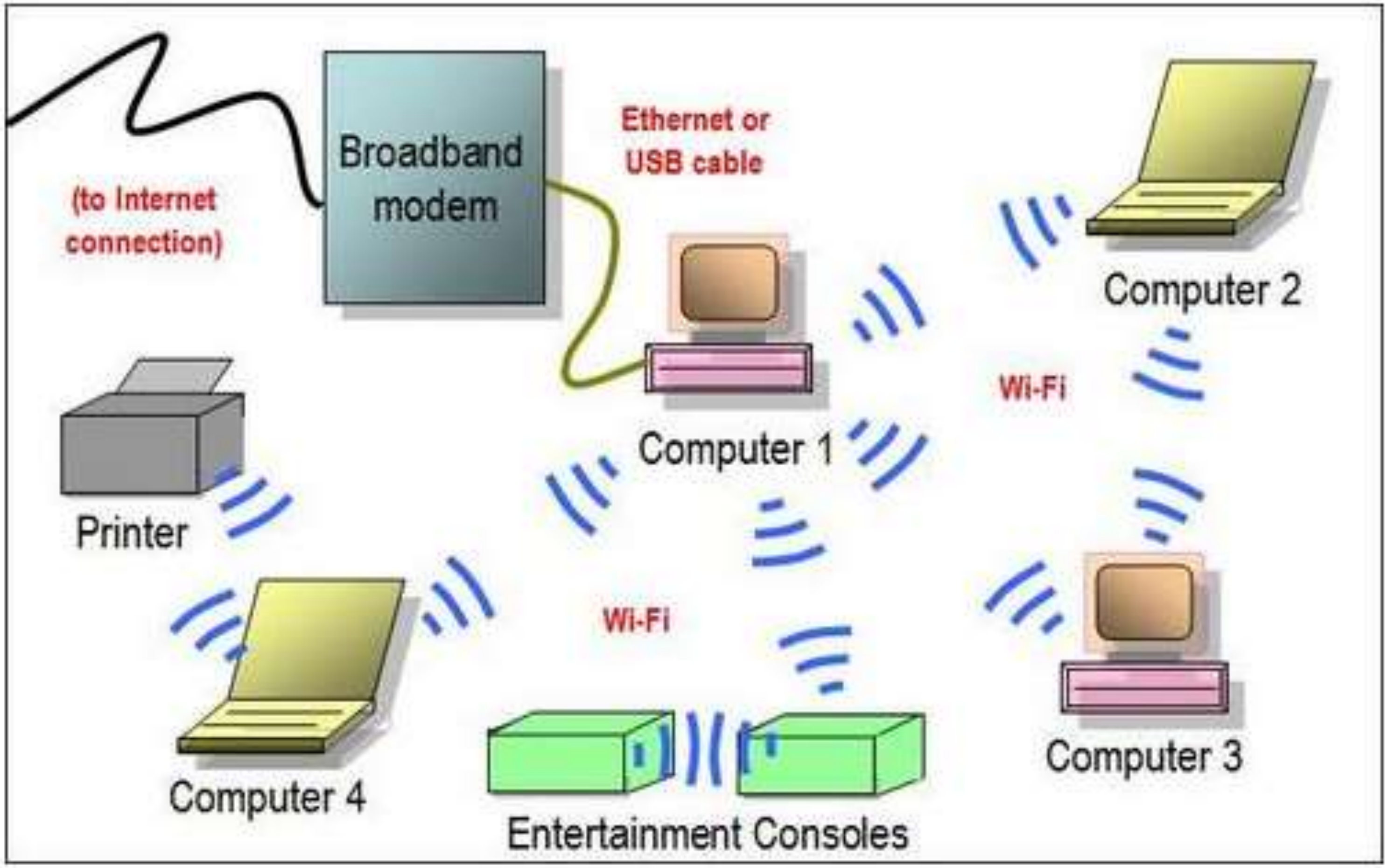
Wireless Architecture and types

Gauri Rao

- Fixed and Wireless Networks are both used in Mobile computing. Fixed networks commonly operate on radio transmission to connect established, wired communications systems.
- The differences between Fixed and Wireless networks can be distinguished as that the wireless networks do not require any cables to make a physical connection with the device. It is easily assessable because it is a shared medium.
- On the other hand, in the case of fixed networks, a physical configuration of devices is mandatory to perform data transmission. In this medium, you have to connect every new device separately and physically to the network.

Wireless Networks	Fixed Networks
There is no requirement of any physical configuration in the wireless network.	In Fixed Networks, a physical configuration is required in any condition.
The data loss rate is high in Wireless Networks.	In Fixed Networks, a perfect link is established between the devices, so; the data loss rate is very low.
In Wireless Networks, the data transmission rate is comparatively low, so it provides less speed.	In Fixed Networks, the rate of data transmission is high, so it provides high speed.
Latency is high in Wireless Networks, which finally results in more delay.	There is no issue of latency in Fixed Networks because there is a perfect connection established between the devices that provide less delay.
The Wireless Networks may be hacked; that's why the security is always low in this type of network.	Fixed Networks connections are highly secured.

Standalone architecture (Ad hoc mode)



- By using ad hoc mode, all devices in the wireless network are directly communicating with each other in peer to peer communication mode. No access point (routers/switches) is required for communication between devices.

- For setting up ad hoc mode, we need to manually configure the wireless adaptors of all devices to be at ad hoc mode instead of infrastructure mode, and all adaptors must use the same channel name and same SSID for making the connection active.

- Ad hoc mode is most suitable for small group of devices and all of these devices must be physically present in close proximity with each other. The performance of network suffers while the number of devices grows. Disconnections of random device may occur frequently and also, ad hoc mode can be a tough job for network administrator to manage the network. Ad hoc mode has another limitation is that, ad hoc mode networks cannot bridge to wired local area network and also cannot access internet if without the installation of special gateways.

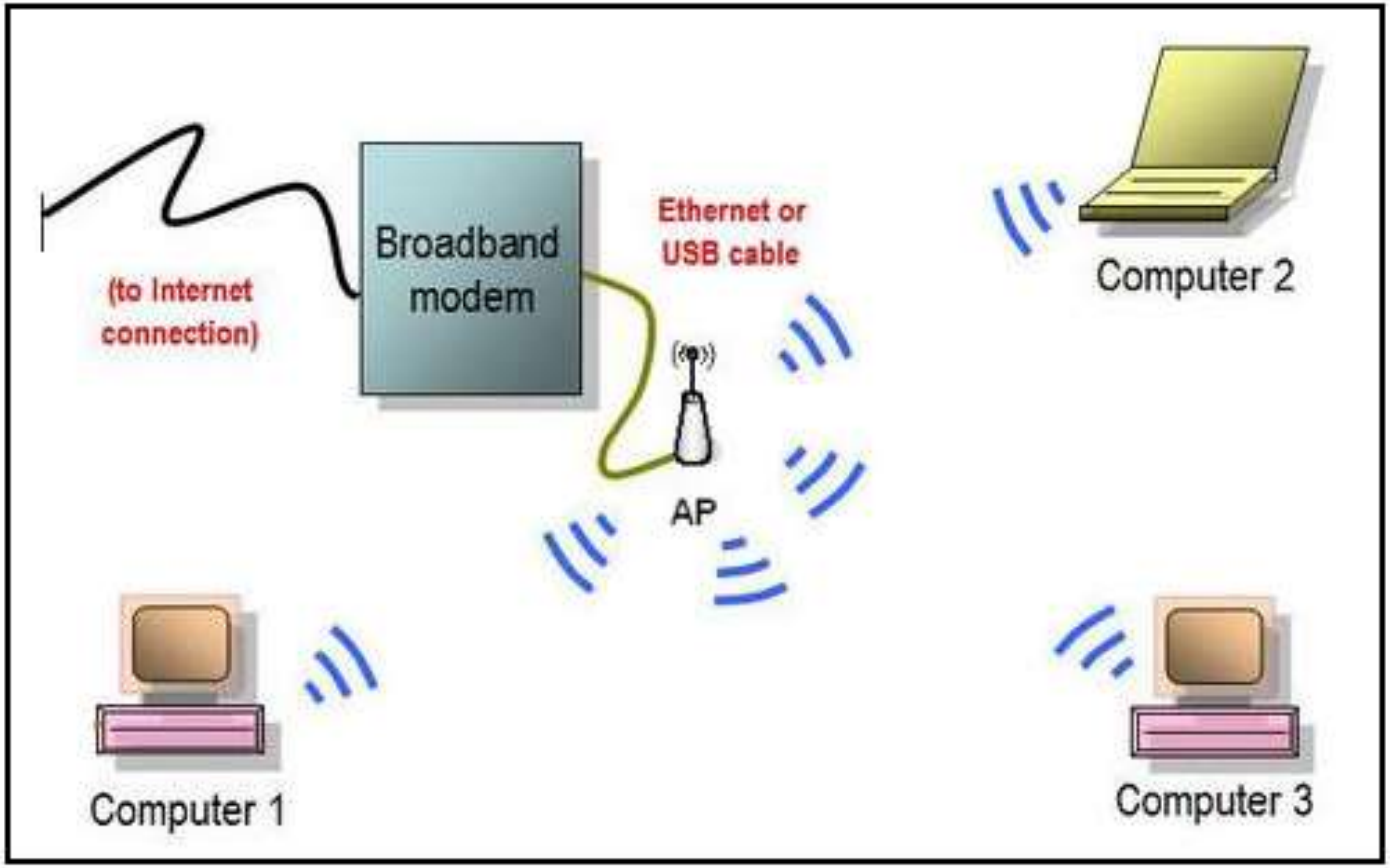
- However, Ad hoc mode works fine in small environment. Because ad hoc mode does not need any extra access point (routers/switches), therefore it reduces the cost. Ad hoc can be very useful as a backup option for time being if network based on centrally coordinated wireless network (infrastructure mode) and access points are malfunctioning.

- An ad hoc mode uses the integrated functionality of each adaptor to enable wireless services and security authentication.

The characteristics of an Ad hoc wireless network :

- All access points in the network operate independently and has own configuration file.
- Access point is responsible for the encryption and decryption.
- The network configuration is static and does not respond to changing network conditions

Centrally Coordinated Architecture (Infrastructure mode)



- All devices are connected to wireless network with the help of Access Point (AP). Wireless APs are usually routers or switches which are connected to internet by broadband modem.

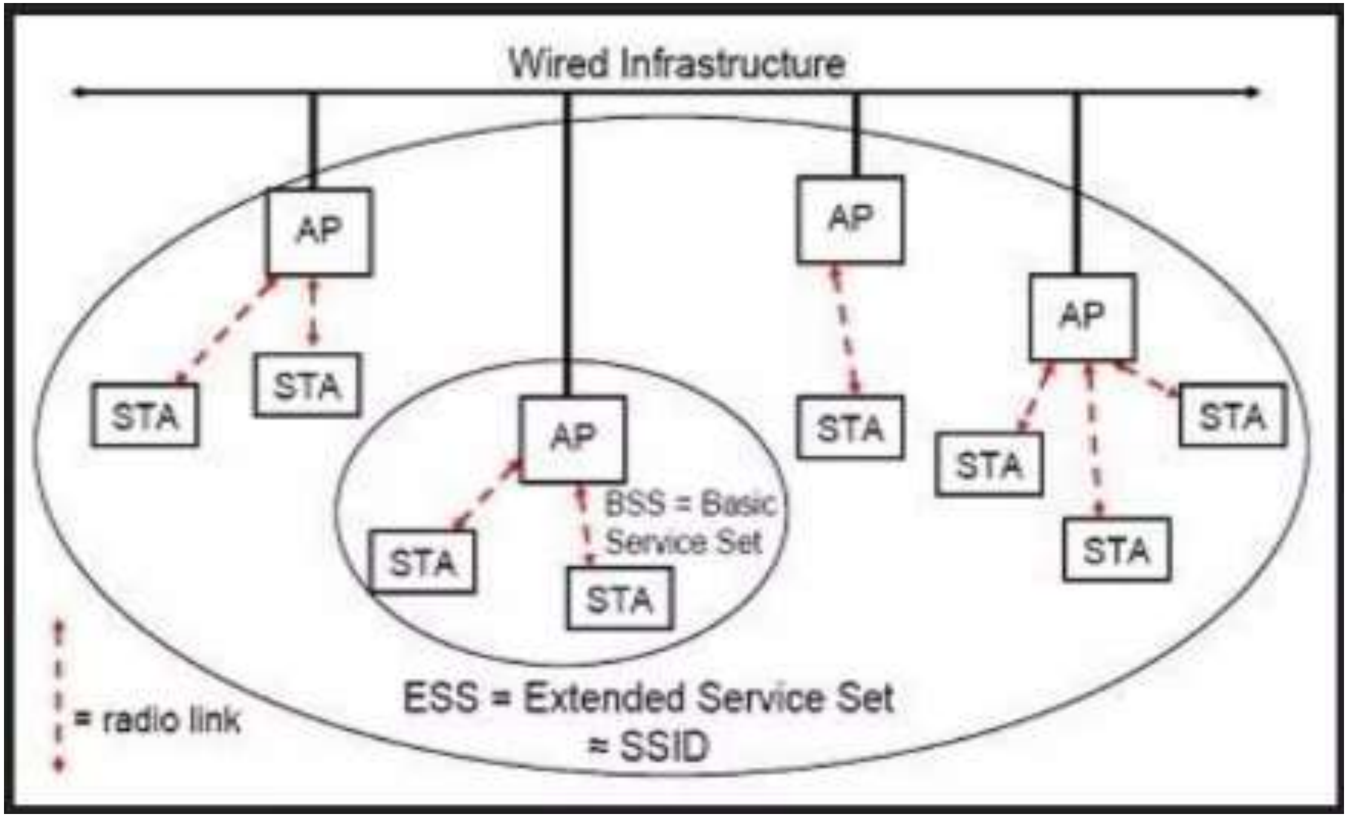
- Infrastructure mode deployments are more suitable for larger organizations or facility. This kind of deployment helps to simplify network management, and allows the facility to address operational concerns. And resiliency is also assured while more users can get connected to the network subsequently.

- The infrastructure mode provides improved security, ease of management, and much more scalability and stability. However, the infrastructure mode incurs extra cost in deploying access points such as routers or switches.

Characteristics of an infrastructure mode wireless network :

- The wireless centralized controller coordinates the activity of access point.
- The controller is able to monitor and control the wireless network by automatically reconfiguring the access point parameters in order to maintain the health of the network.
- The wireless network can be easily expanded or reduced by adding or removing access points and the network can be reconfigured by the controller based on the changes in RF footprint.
- Tasks such as user authentication, fault tolerance, control of configuration, policy enforcement and expansion of network are done by the wireless network controller.
- Redundant access points can be deployed in separate locations to maintain control in the event of an access point or switch failure.

Wireless Network Architecture



types of wireless networks

- **wireless LAN**
- **wireless MAN**
- **wireless PAN**
- **wireless WAN**

1. Wireless LAN

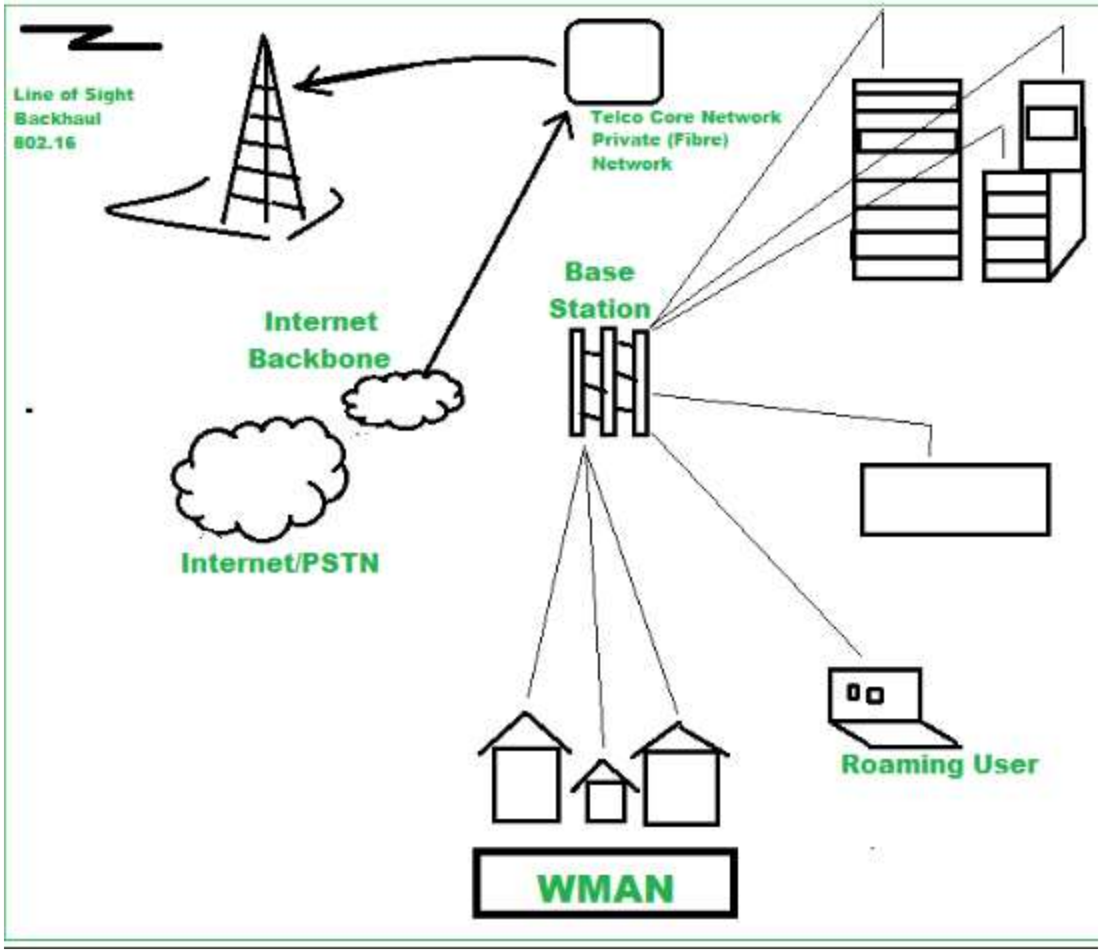
- Wireless LAN (WLAN) technology provides internet access within a building or a limited outdoor area. First used within offices and homes, WLAN technology is now also used in stores and restaurants. The use of home networks greatly increased as the COVID-19 pandemic forced office workers, students, teachers and others to work and study from home.

- Most home network designs are simple. A modem connects to the cable or fiber from a local service provider. A wireless router is connected to the modem and receives the signal from the modem. The router also serves as the wireless access point (AP), which then broadcasts using a wireless protocol, such as the 802.11 standards.

- Office networks are more complicated. APs are usually mounted on the ceiling, with each broadcasting a wireless signal to the surrounding area. Multiple APs are required in large offices, each connecting to the office backbone network via a wired connection to a switch. APs coordinate support for users walking through the office area and hand off support to maintain open, connected sessions from AP to AP.

2. Wireless MAN

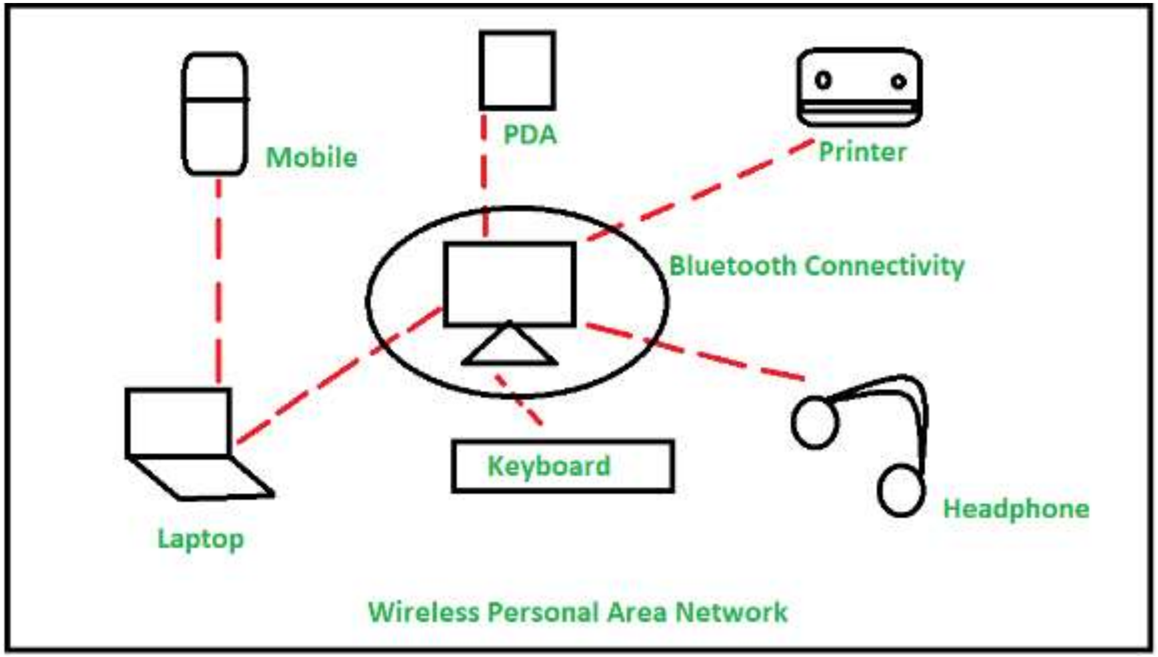
- Wireless metropolitan area networks have been installed in cities worldwide to provide access for people outside an office or home network. These networks cover a wider area than office or home networks, but the principles are the same.
- APs are located on the sides of buildings or on telephone poles throughout the covered area. APs are connected to the internet via a wired network and broadcast a wireless signal throughout the area. Users connect to their desired destination by connecting to the nearest AP, which forwards the connection through its internet connection.


Line of Sight
Backhaul
802.16
Telco Core Network
Private (Fibre)
Network
Base
Station
Internet
Backbone
Internet/PSTN
Roaming User
WMAN

3. Wireless PAN

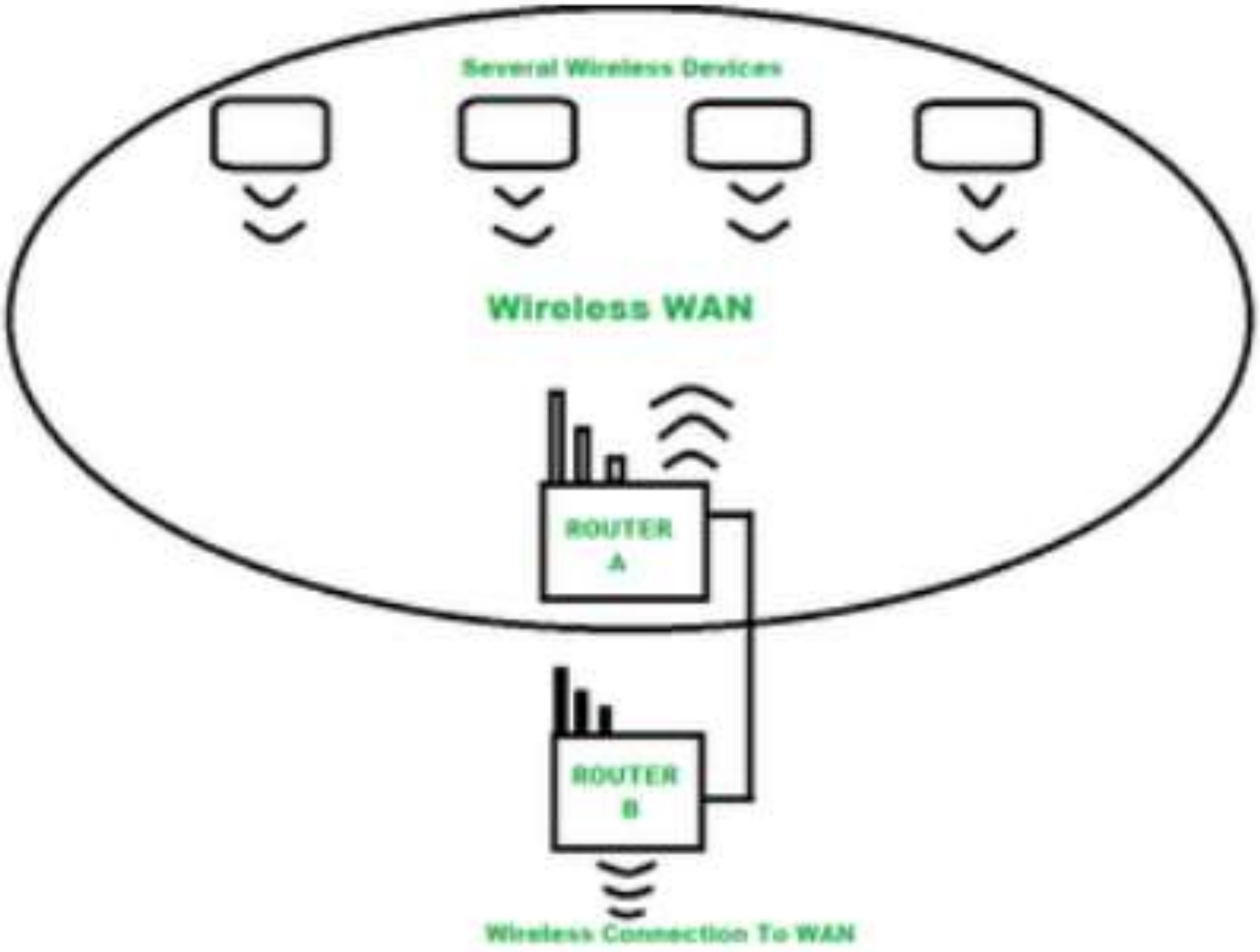
- Wireless personal area networks cover a very limited area -- typically a maximum of 100 meters for most applications -- using protocols like Bluetooth and Zigbee.
- Bluetooth enables hands-free phone calls, connects a phone to earpieces or transmits signals between smart devices.
- Zigbee connects stations along an IoT network. Infrared technology is limited to line of sight, such as connecting TV remotes to televisions.

- Wireless developers have constantly improved technology by discovering new ways to transmit signals to users. These advances enable higher data rates and increasing range for each of these wireless technologies.

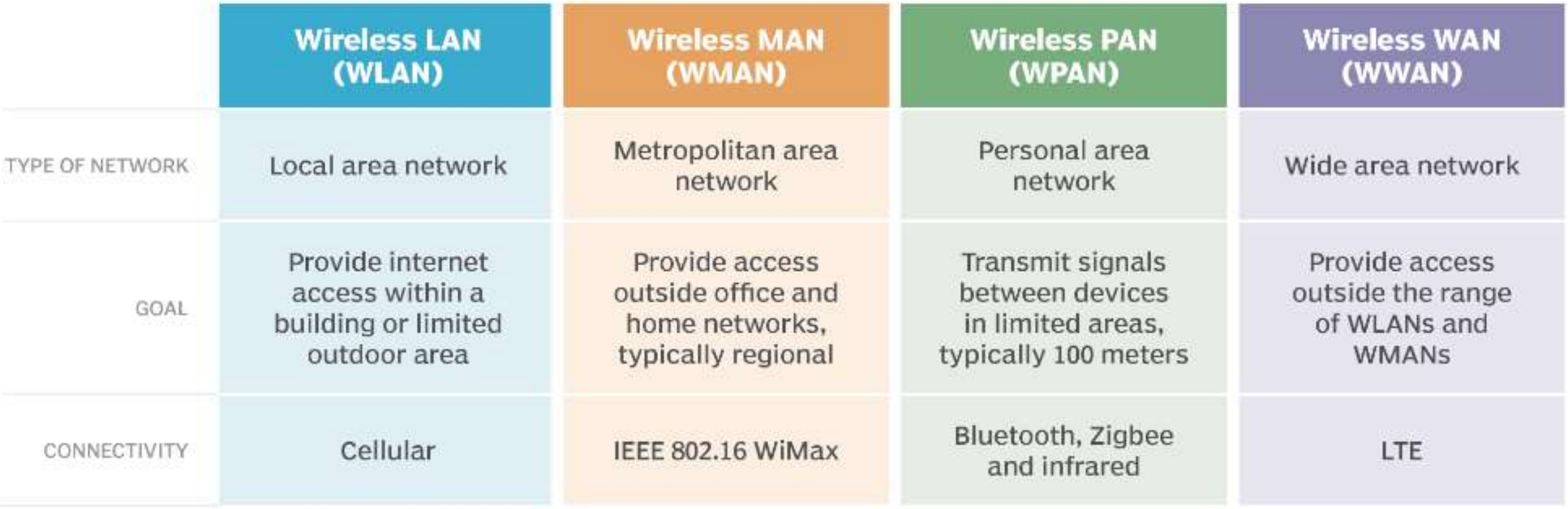

Mobile
PDA
Printer
Bluetooth Connectivity
Laptop
Keyboard
Headphone
Wireless Personal Area Network

4. Wireless WAN

- Wireless WANs use cellular technology to provide access outside the range of a wireless LAN or metropolitan network. These networks enable users to make phone calls to others.
- WANs can support either speech or data transfer using the same technology. Users can also connect to the internet to access websites or server-based applications.
- Cell towers are located nearly everywhere within the U.S. and most other countries. A user connection is routed to the nearest cell tower, which, in turn, is connected either to the wired internet or to another tower connected to wired internet.


Several Wireless Devices
Wireless WAN
ROUTER A
ROUTER B
Wireless Connection To WAN

Types of wireless networks



	Wireless LAN (WLAN)	Wireless MAN (WMAN)	Wireless PAN (WPAN)	Wireless WAN (WWAN)
TYPE OF NETWORK	Local area network	Metropolitan area network	Personal area network	Wide area network
GOAL	Provide internet access within a building or limited outdoor area	Provide access outside office and home networks, typically regional	Transmit signals between devices in limited areas, typically 100 meters	Provide access outside the range of WLANs and WMANs
CONNECTIVITY	Cellular	IEEE 802.16 WiMax	Bluetooth, Zigbee and infrared	LTE

UNIT-II GSM

GAURI RAO

- GSM stands for Global System for Mobile Communication. GSM is an open and digital cellular technology used for mobile communication. It uses 4 different frequency bands 850 MHz, 900 MHz, 1800 MHz, and 1900 MHz. It uses the combination of FDMA and TDMA.

- The goal of GSM was to produce a mobile phone system that allows users to roam through Europe and provide voice services compatible to ISDN and PSTN.
- GSM permits integration of different voice and data services and interworking with existing networks

GSM services

1. Bearer Services
2. Tele services
3. Supplementary services

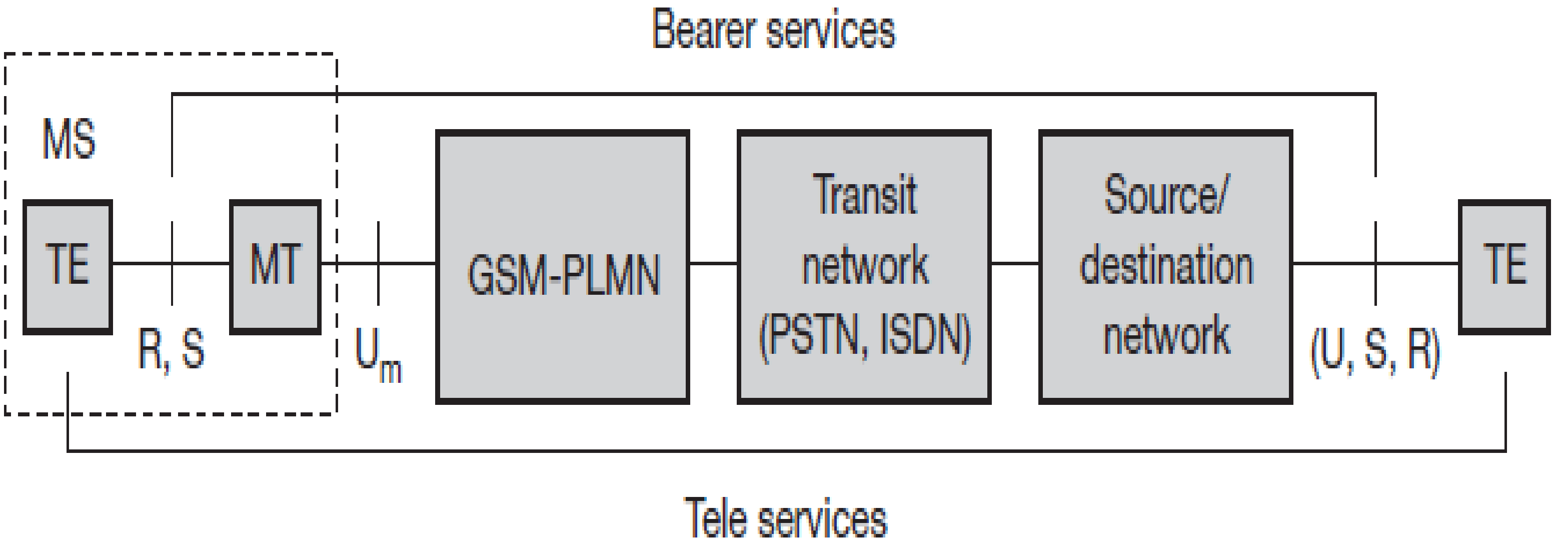

Bearer services
MS
TE
R, S
MT
U_m
GSM-PLMN
Transit network (PSTN, ISDN)
Source/ destination network
(U, S, R)
TE
Tele services

Figure shows a reference model for GSM services. A **mobile station MS** is connected to the **GSM public land mobile network (PLMN)** via the U_m interface. (GSM-PLMN is the infrastructure needed for the GSM network.) This network is connected to transit networks, e.g., **integrated services digital network (ISDN)** or traditional **public switched telephone network (PSTN)**. There might be an additional network, the source/destination network, before another **terminal** TE is connected. **Bearer services** now comprise all services that enable the transparent transmission of data between the interfaces to the network, i.e., S in case of the mobile station, and a similar interface for the other terminal (e.g., S_0 for ISDN terminals). Interfaces like U, S, and R in case of ISDN have not been defined for all networks, so it depends on the specific network which interface is used as a reference for the transparent transmission of data. In the classical GSM model, bearer services are connection-oriented and circuit- or packet-switched. These services only need the lower three layers of the ISO/OSI reference model.

Bearer Services

- Comprises all services that enable the transparent transmission of data between interfaces and network.
- Connection oriented
- Circuit/packet switching
- Uses lower 3 layers of OSI model.
- Permits Transparent /Non transparent, Synchronous/Asynchronous data transmission

- Bearer services/ data services: GSM specifies different mechanism for data transmission, The original GSM allowing for data rates of up to 9600 bits/s. Bearer services permit transparent or non transparent data transmission.
- Transparent bearer services: Transparent bearer services only use the physical layer to transmit data. Data transmission has a constant delay at throughput if no transmission error occurs.
- Non-transparent bearer services: Non-transparent bearer services use protocols of layer two and three two three to implement error correction and flow control.(data link layer and network layer).

Bearer services

Transparent

- Use the functions of physical layer
- Use forward error correction code
- Do not try to recover lost data due to say handover

Non Transparent

- Use 2nd and 3rd layer of OSI model
- Error correction and flow control
- Uses RLP for retransmission of erroneous data

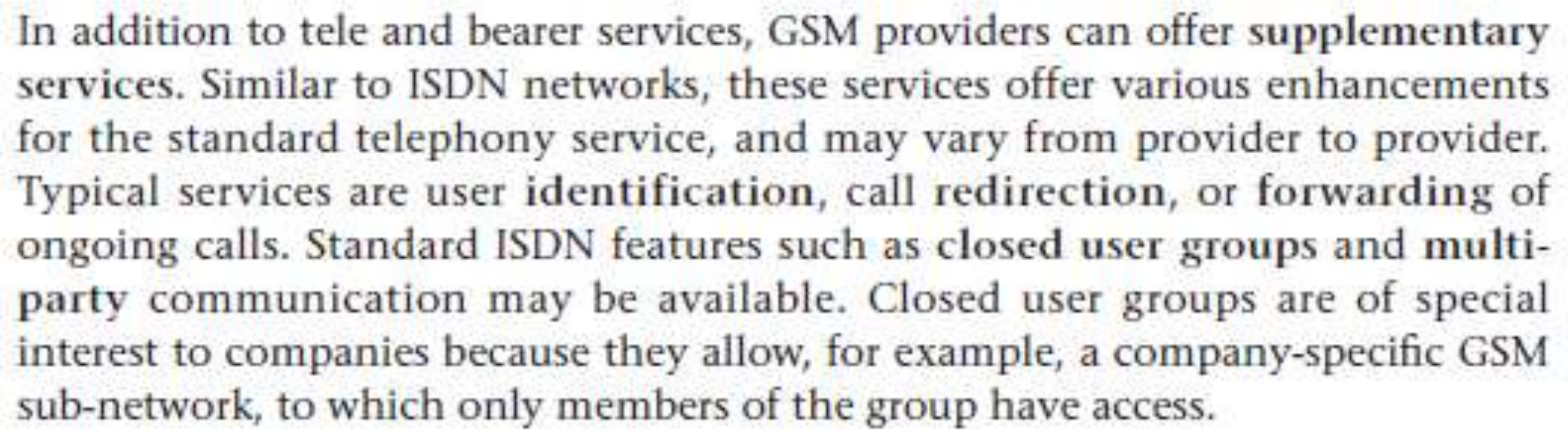
Tele services

- Voice oriented services
- SMS/MMS
- Emergency number
- Uses all 7 layers of OSI model
- Application specific

Tele services: Tele services are nothing but we use now as Video calls.

- Video text and face emoji.
- Short text message(SMS).

Supplementary

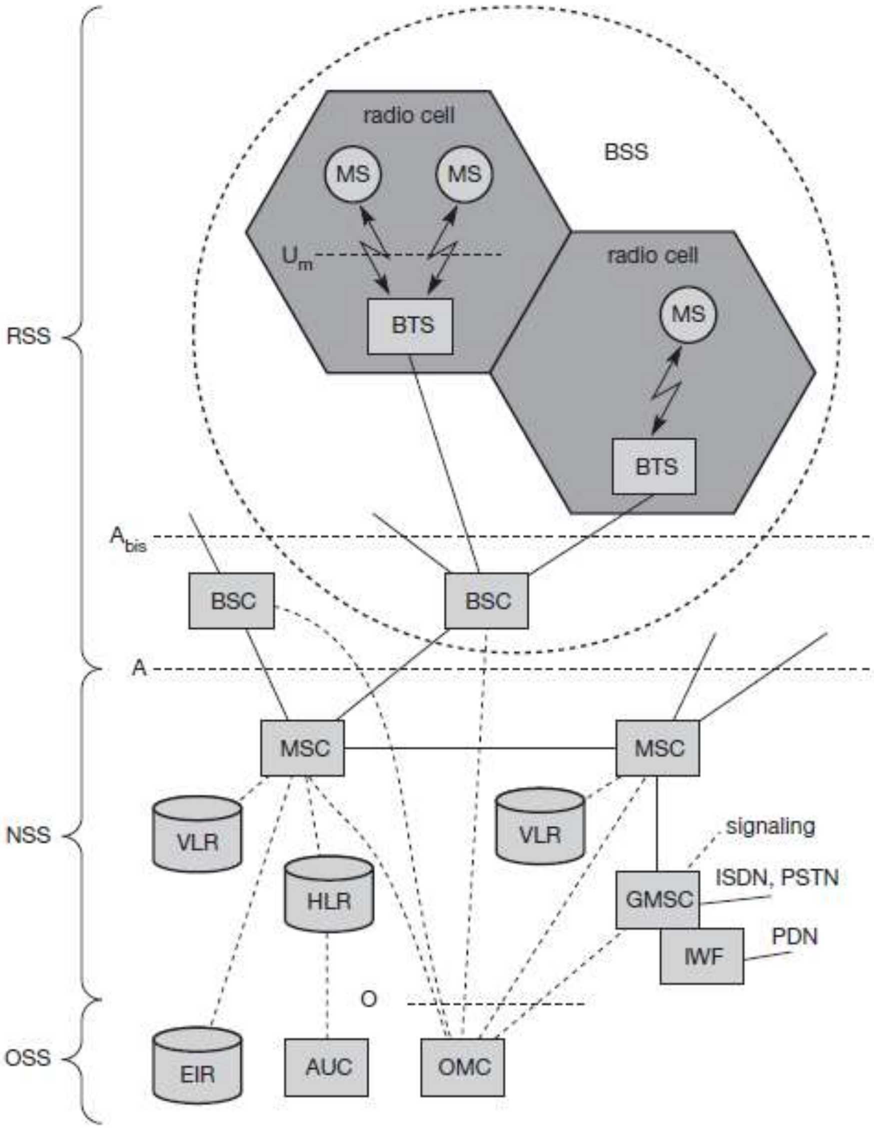


In addition to tele and bearer services, GSM providers can offer **supplementary services**. Similar to ISDN networks, these services offer various enhancements for the standard telephony service, and may vary from provider to provider. Typical services are user **identification**, call **redirection**, or **forwarding** of ongoing calls. Standard ISDN features such as **closed user groups** and **multi-party** communication may be available. Closed user groups are of special interest to companies because they allow, for example, a company-specific GSM sub-network, to which only members of the group have access.

Supplementary services: Supplementary services it means advanced services.

- Conference calls.
- Call waiting.
- Call forwarding.

GSM Architecture



The Architecture of GSM

BSS : BSS stands for Base Station Subsystem. BSS handles traffic and signaling between a mobile phone and the network switching subsystem. BSS having two components BTS and BSC.

NSS : NSS stands for Network and Switching Subsystem. NSS is the core network of GSM. That carried out call and mobility management functions for mobile phone present in network. NSS have different components like VLR, HLR and EIR.

OSS : OSS stands for Operating Subsystem. OSS is a functional entity which the network operator monitor and control the system. OMC is the part of OSS. Purpose of OSS is to offer the customer cost-effective support for all GSM related maintenance services.

- MS : MS stands for Mobile System. MS comprises user equipment and software needed for communication with a mobile network. Mobile Station (MS) = Mobile Equipment(ME) + Subscriber Identity Module (SIM). Now, these mobile stations are connected to tower and that tower connected with BTS through TRX. TRX is a transceiver which comprises transmitter and receiver. Transceiver has two performance of sending and receiving.

- BTS : BTS stands for Base Transceiver Station which facilitates wireless communication between user equipment and a network. Every tower has BTS.

- BSC : BSC stands for Base Station Controller. BSC has multiple BTS. You can consider the BSC as a local exchange of your area which has multiple towers and multiple towers have BTS.

- MSC : MSC stands for Mobile Switching Center. MSC is associated with communication switching functions such as call setup, call release and routing. Call tracing, call forwarding all functions are performed at the MSC level. MSC is having further components like VLR, HLR, AUC, EIR and PSTN.
- VLR : VLR stands for Visitor Location Register. VLR is a database which contains the exact location of all mobile subscribers currently present in the service area of MSC. If you are going from one state to another state then your entry is marked into the database of VLR.
- HLR : HLR stands for Home Location Register. HLR is a database containing pertinent data regarding subscribers authorized to use a GSM network. If you purchase SIM card from in the HLR. HLR is like a home which contains all data like your ID proof, which plan you are taking, which caller tune you are using etc.
- AUC : AUC stands for Authentication Center. AUC authenticates the mobile subscriber that wants to connect in the network.

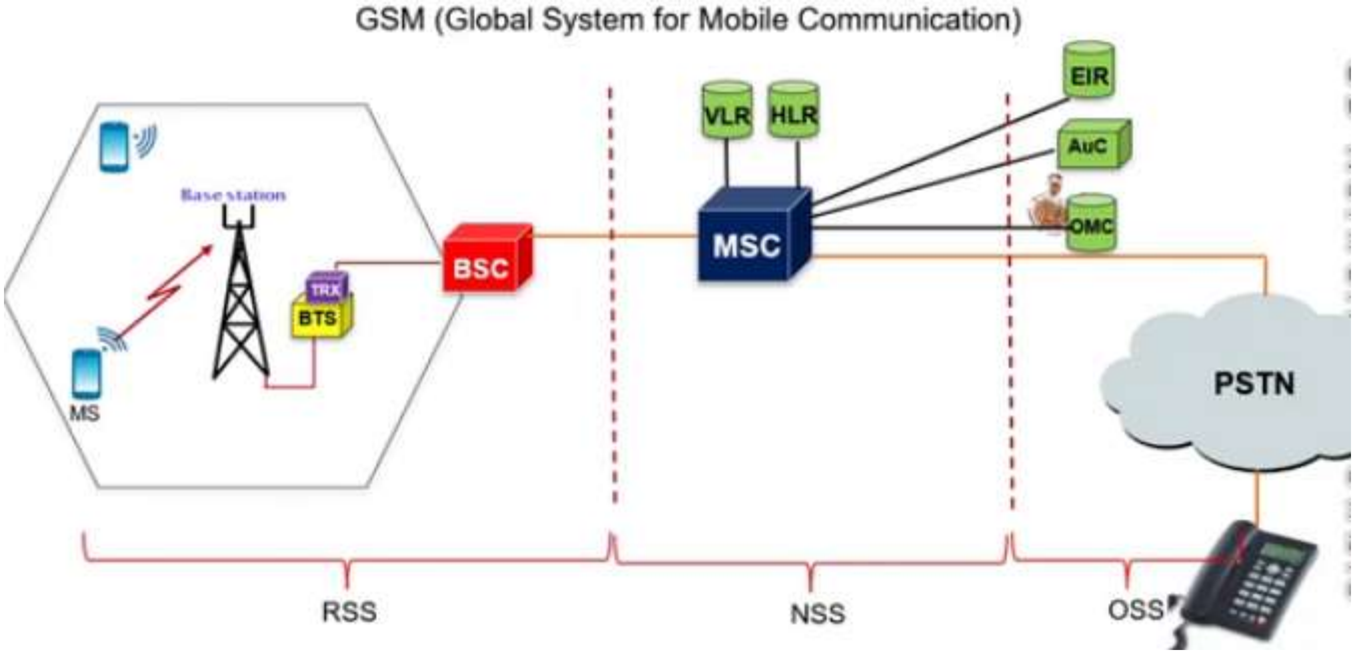
- EIR : EIR stands for Equipment Identity Register. EIR is a database that keeps the record of all allowed or banned in the network. If you are banned in the network then you can’t enter the network, and you can’t make the calls.
- PSTN : PSTN stands for Public Switched Telephone Network. PSTN connects with MSC. PSTN originally a network of fixed line analog telephone systems. Now almost entirely digital in its core network and includes mobile and other networks as well as fixed telephones. The earlier landline phones which places at our home is nothing but PSTN.
- OMC : OMC stands for Operation Maintenance Center. OMC monitor and maintain the performance of each MS, BSC and MSC within a GSM system.

Interfaces

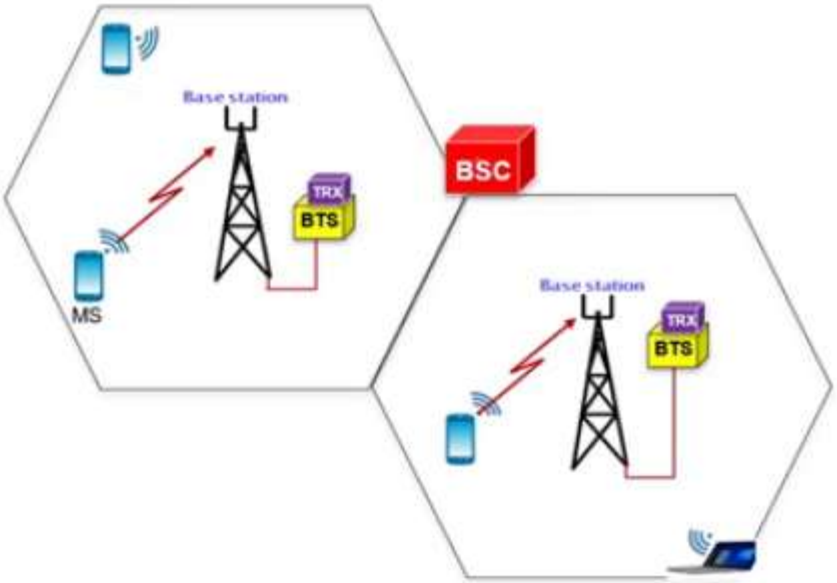
Three subsystem BSS, NSS and OSS are connected with each other via some interfaces. Total three interfaces are there:

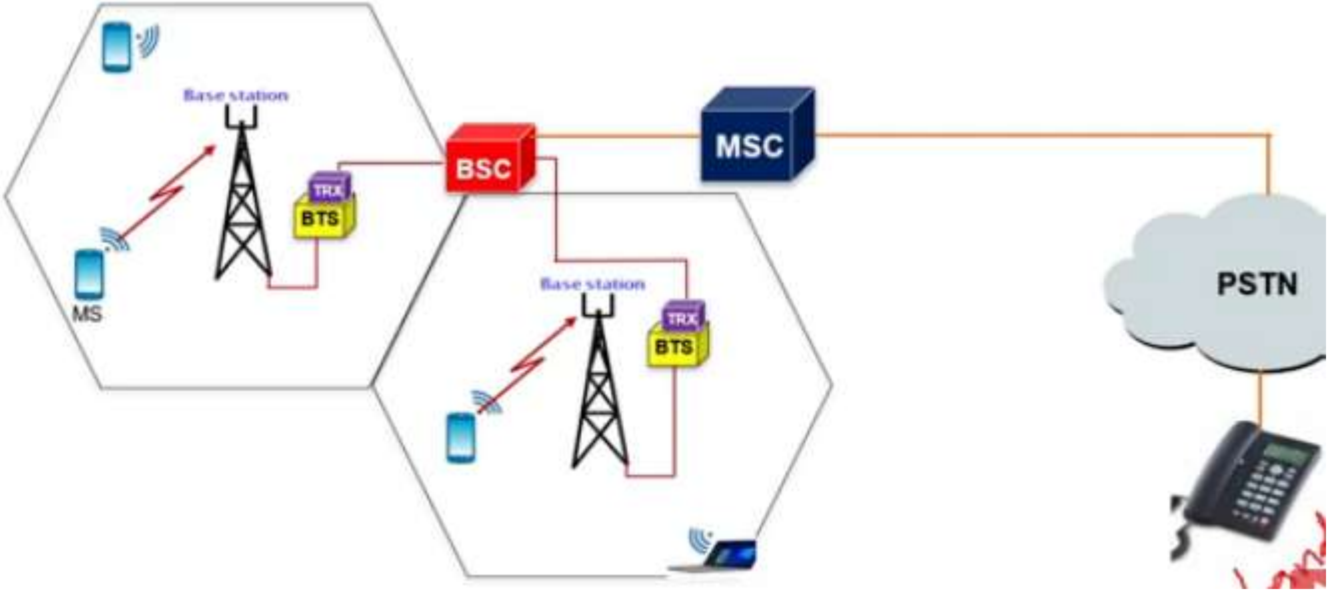
- Air Interface : Air interface is also known as UM interface. Interface between MS and BTS is called as UM interface because it is mobile analog to the U interface of ISDN.
- Abis Interface : It is a BSS internal interface linking with BTS and BSC.
- A interface : It provides communication between BSS and MSC.

GSM

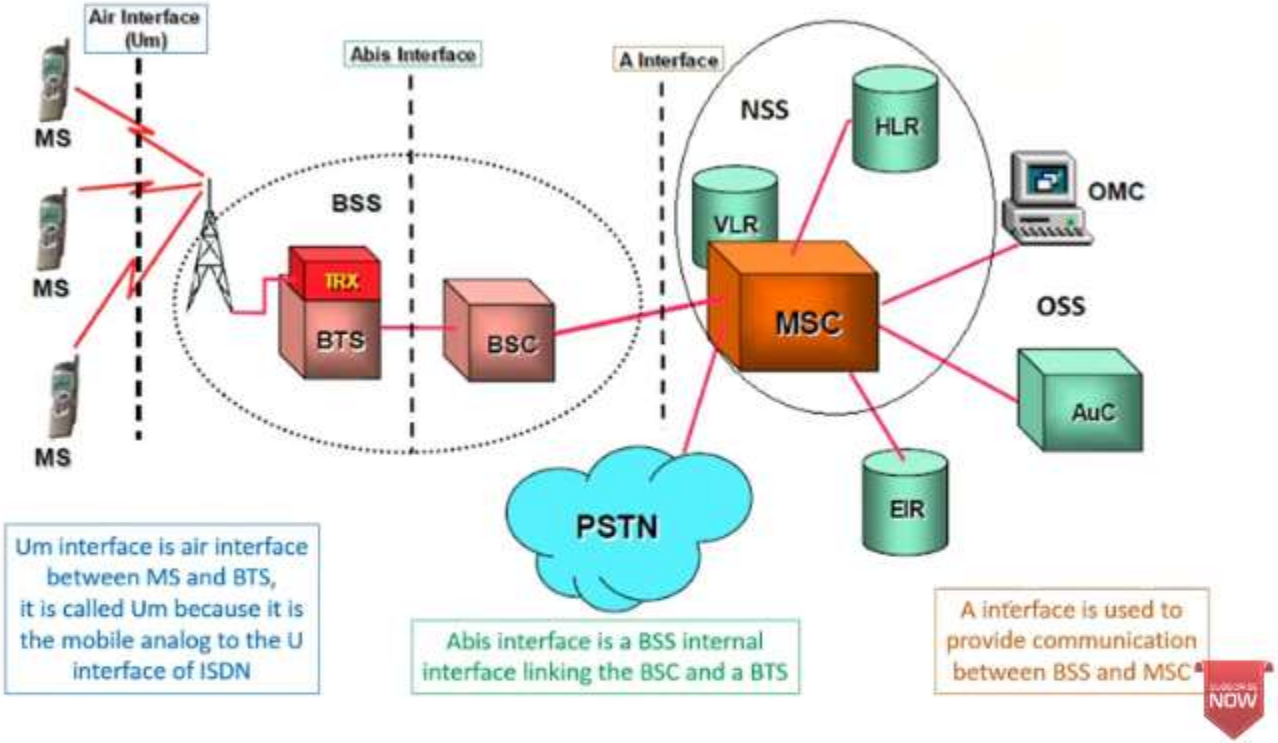

GSM (Global System for Mobile Communication)
Base station
MS
TRX
BTS
BSC
VLR
HLR
MSC
EIR
AuC
OMC
PSTN
RSS
NSS
OSS

Radio Sub system




Base station
MS
TRX
BTS
BSC
MSC
Base station
TRX
BTS
PSTN

Interface

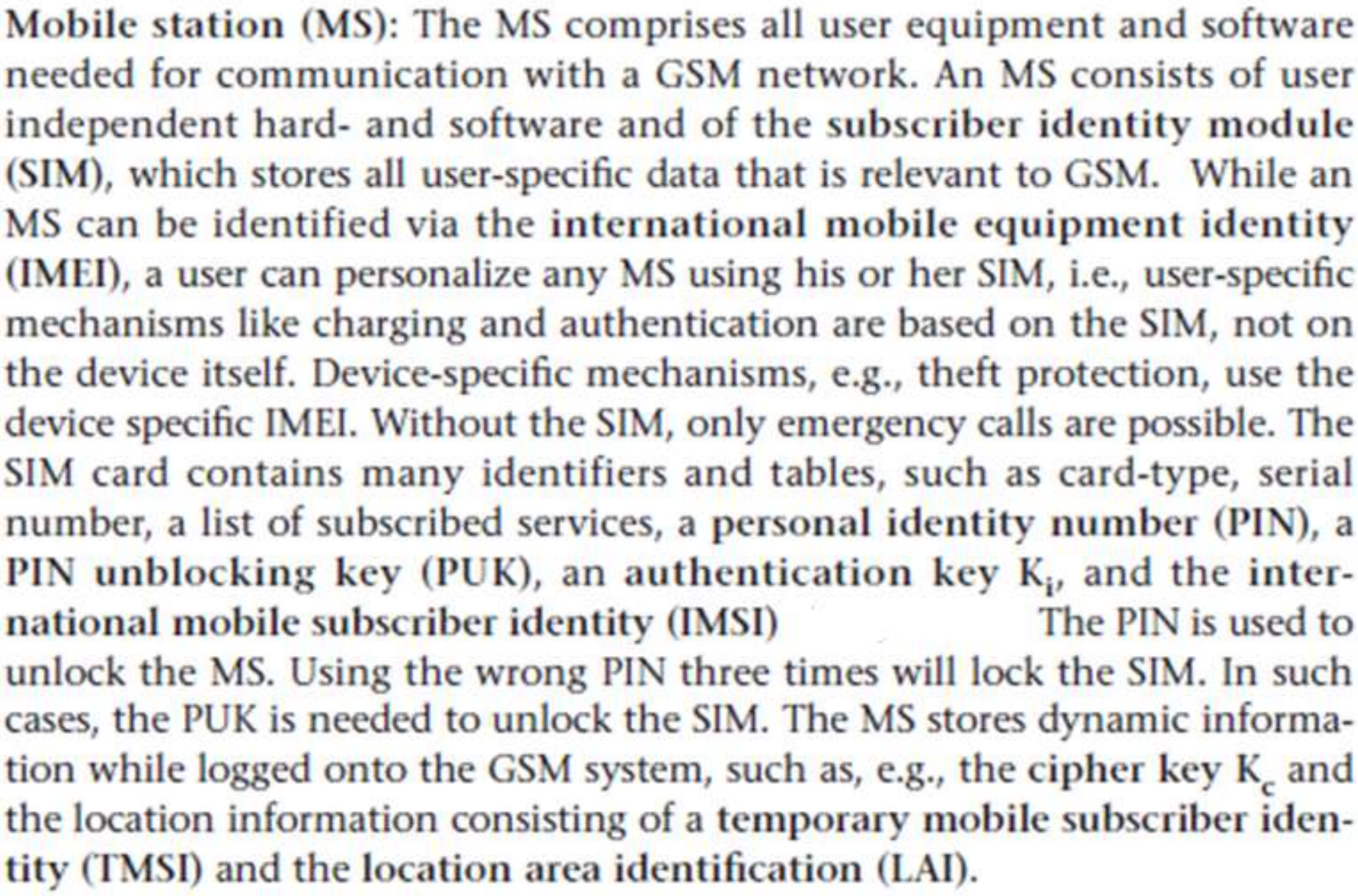


Um interface is air interface between MS and BTS, it is called Um because it is the mobile analog to the U interface of ISDN

Abis interface is a BSS internal interface linking the BSC and a BTS

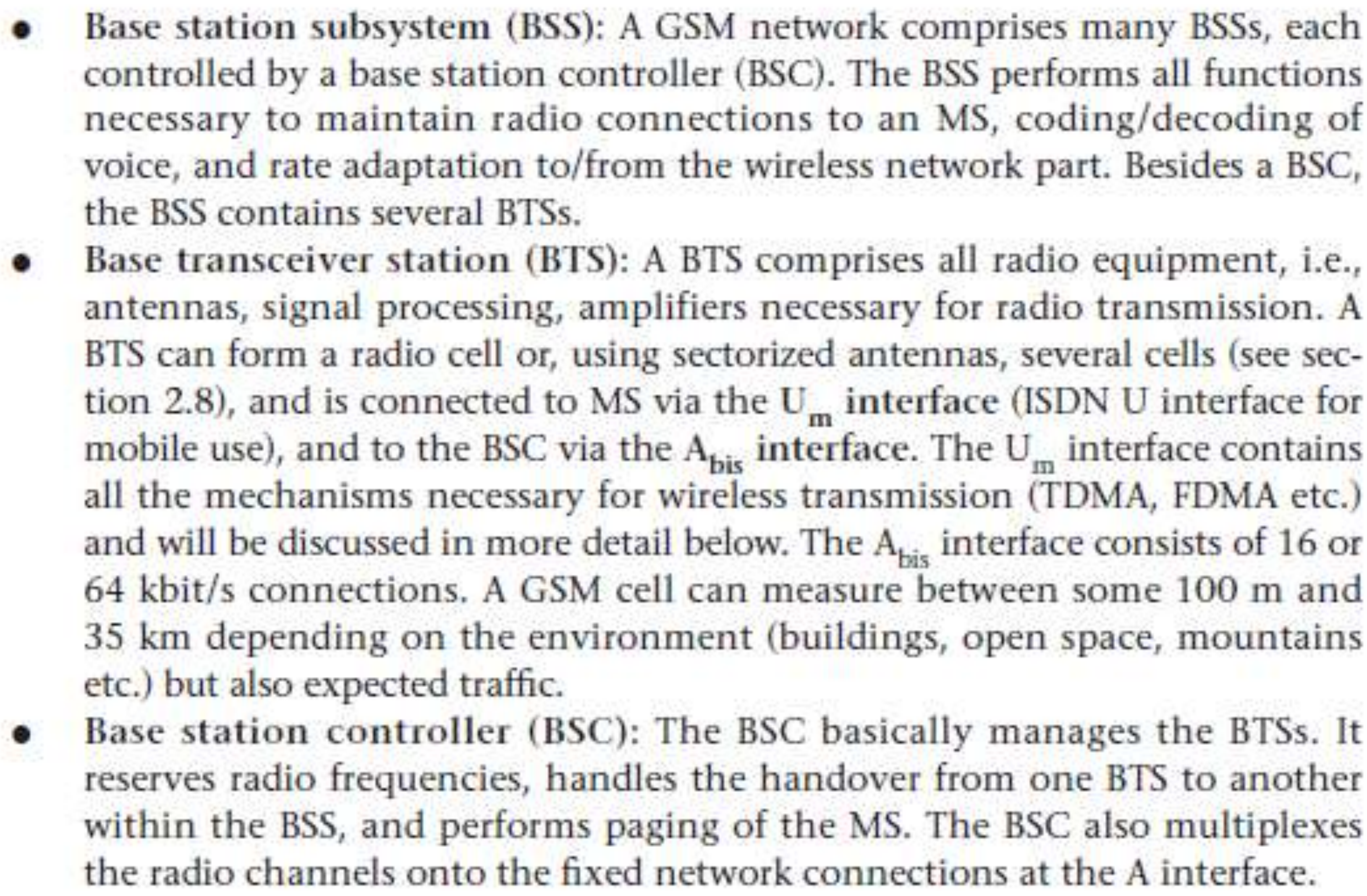
A interface is used to provide communication between BSS and MSC

1. RSS



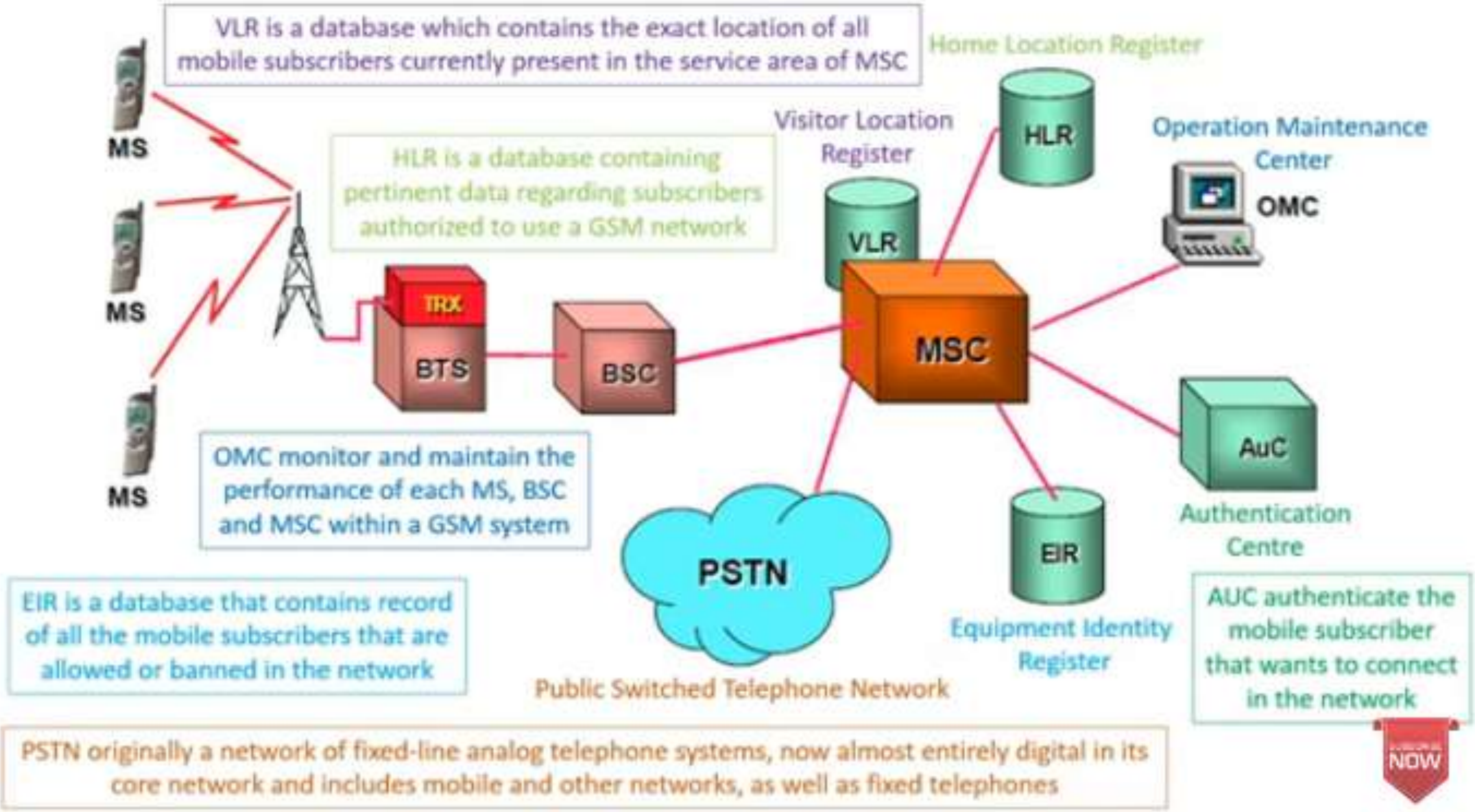
Mobile station (MS): The MS comprises all user equipment and software needed for communication with a GSM network. An MS consists of user independent hard- and software and of the **subscriber identity module** (SIM), which stores all user-specific data that is relevant to GSM. While an MS can be identified via the **international mobile equipment identity** (IMEI), a user can personalize any MS using his or her SIM, i.e., user-specific mechanisms like charging and authentication are based on the SIM, not on the device itself. Device-specific mechanisms, e.g., theft protection, use the device specific IMEI. Without the SIM, only emergency calls are possible. The SIM card contains many identifiers and tables, such as card-type, serial number, a list of subscribed services, a **personal identity number** (PIN), a **PIN unblocking key** (PUK), an **authentication key** K_i, and the **international mobile subscriber identity** (IMSI) The PIN is used to unlock the MS. Using the wrong PIN three times will lock the SIM. In such cases, the PUK is needed to unlock the SIM. The MS stores dynamic information while logged onto the GSM system, such as, e.g., the **cipher key** K_c and the location information consisting of a **temporary mobile subscriber identity** (TMSI) and the **location area identification** (LAI).

RSS



- **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 transceiver station (BTS):** A BTS comprises all radio equipment, i.e., antennas, signal processing, amplifiers necessary for radio transmission. A BTS can form a radio cell or, using sectorized antennas, several cells (see section 2.8), and is connected to MS via the U_m **interface** (ISDN U interface for mobile use), and to the BSC via the A_{bis} **interface**. The U_m interface contains all the mechanisms necessary for wireless transmission (TDMA, FDMA etc.) and will be discussed in more detail below. The A_{bis} interface consists of 16 or 64 kbit/s connections. A GSM cell can measure between some 100 m and 35 km depending on the environment (buildings, open space, mountains etc.) but also expected traffic.
- **Base station controller (BSC):** The BSC basically manages the BTSs. It reserves radio frequencies, handles the handover from one BTS to another within the BSS, and performs paging of the MS. The BSC also multiplexes the radio channels onto the fixed network connections at the A interface.

NSS



VLR is a database which contains the exact location of all mobile subscribers currently present in the service area of MSC

HLR is a database containing pertinent data regarding subscribers authorized to use a GSM network

OMC monitor and maintain the performance of each MS, BSC and MSC within a GSM system

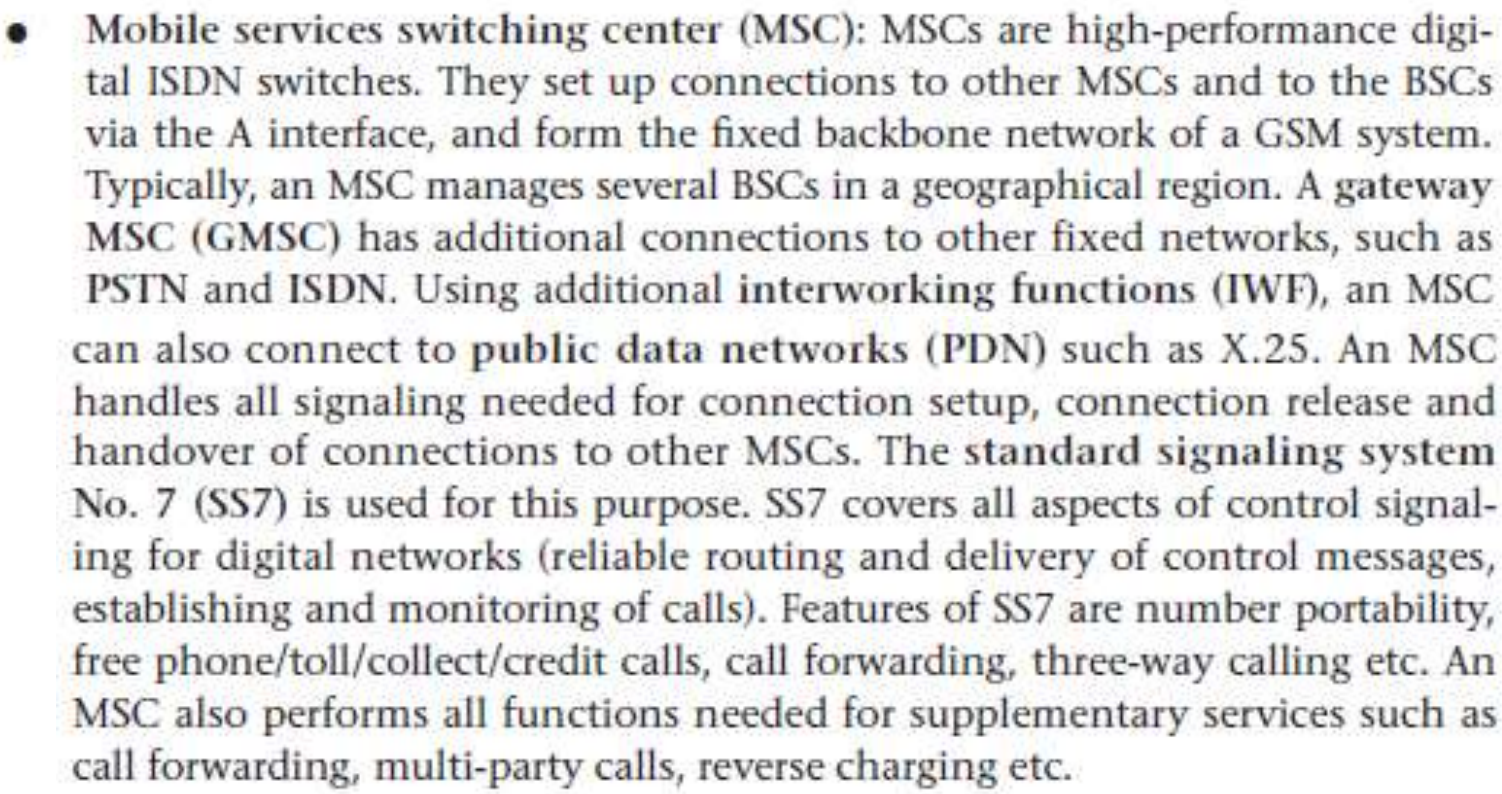
EIR is a database that contains record of all the mobile subscribers that are allowed or banned in the network

AUC authenticate the mobile subscriber that wants to connect in the network

PSTN originally a network of fixed-line analog telephone systems, now almost entirely digital in its core network and includes mobile and other networks, as well as fixed telephones

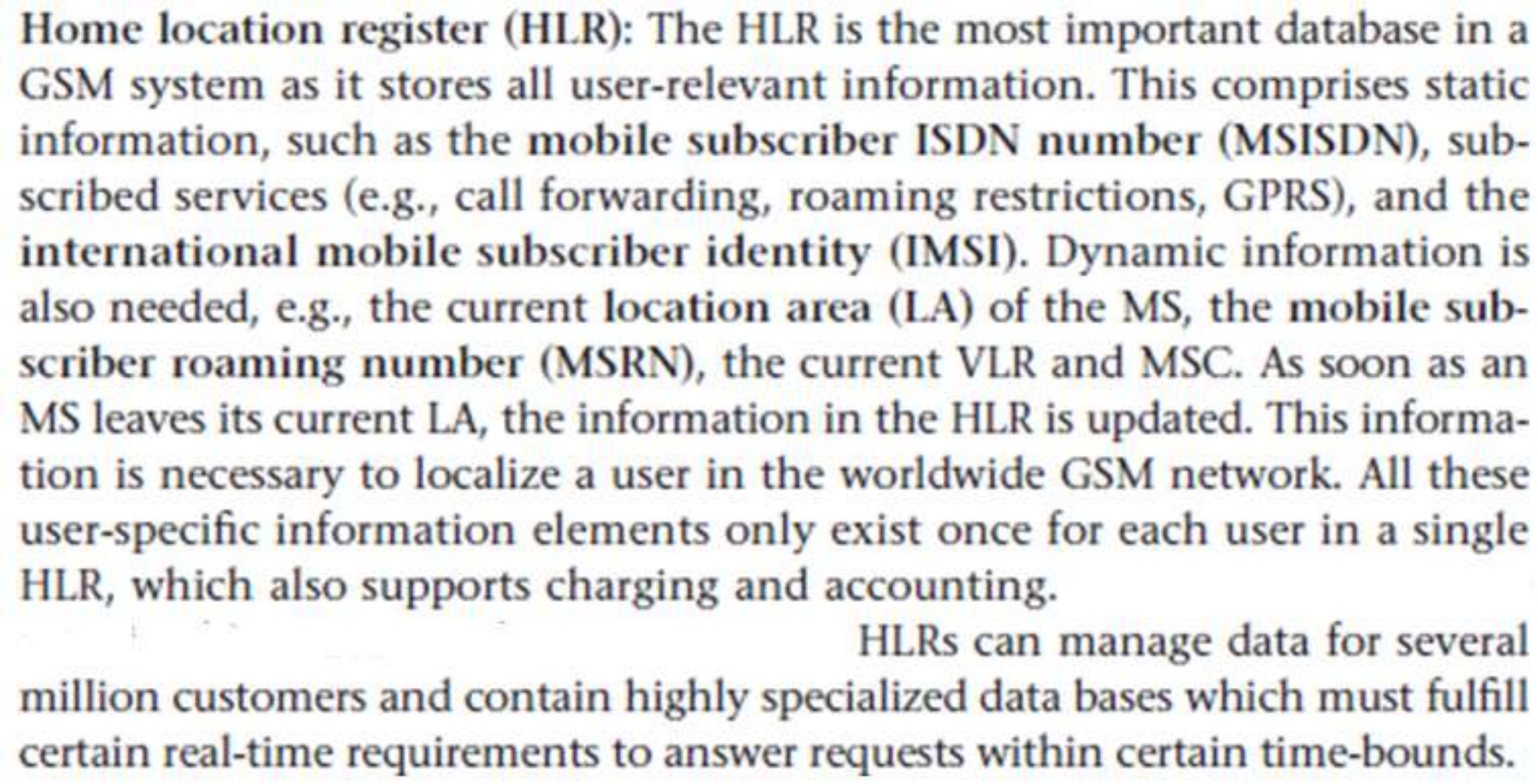
MS
MS
MS
TRX
BTS
BSC
MSC
Visitor Location Register
VLR
Home Location Register
HLR
Operation Maintenance Center
OMC
AuC
Authentication Centre
EIR
Equipment Identity Register
PSTN
Public Switched Telephone Network
NOW

2. NSS



- **Mobile services switching center (MSC):** MSCs are high-performance digital ISDN switches. They set up connections to other MSCs and to the BSCs via the A interface, and form the fixed backbone network of a GSM system. Typically, an MSC manages several BSCs in a geographical region. A **gateway MSC (GMSC)** has additional connections to other fixed networks, such as **PSTN** and **ISDN**. Using additional **interworking functions (IWF)**, an MSC can also connect to **public data networks (PDN)** such as X.25. An MSC handles all signaling needed for connection setup, connection release and handover of connections to other MSCs. The **standard signaling system No. 7 (SS7)** is used for this purpose. SS7 covers all aspects of control signaling for digital networks (reliable routing and delivery of control messages, establishing and monitoring of calls). Features of SS7 are number portability, free phone/toll/collect/credit calls, call forwarding, three-way calling etc. An MSC also performs all functions needed for supplementary services such as call forwarding, multi-party calls, reverse charging etc.

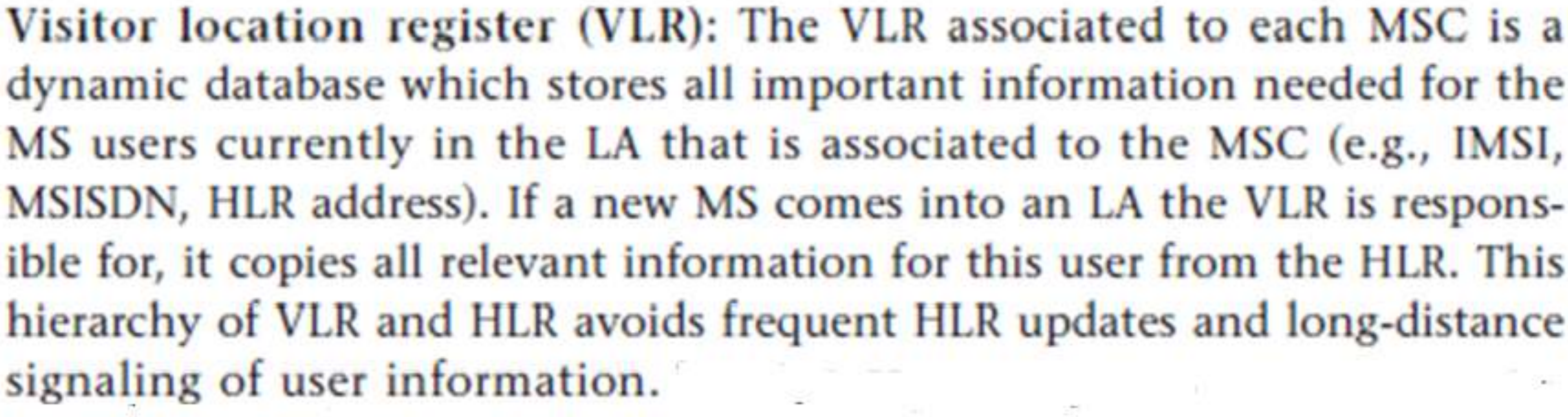
HLR



Home location register (HLR): The HLR is the most important database in a GSM system as it stores all user-relevant information. This comprises static information, such as the **mobile subscriber ISDN number** (MSISDN), subscribed services (e.g., call forwarding, roaming restrictions, GPRS), and the **international mobile subscriber identity** (IMSI). Dynamic information is also needed, e.g., the current **location area** (LA) of the MS, the **mobile subscriber roaming number** (MSRN), the current VLR and MSC. As soon as an MS leaves its current LA, the information in the HLR is updated. This information is necessary to localize a user in the worldwide GSM network. All these user-specific information elements only exist once for each user in a single HLR, which also supports charging and accounting.

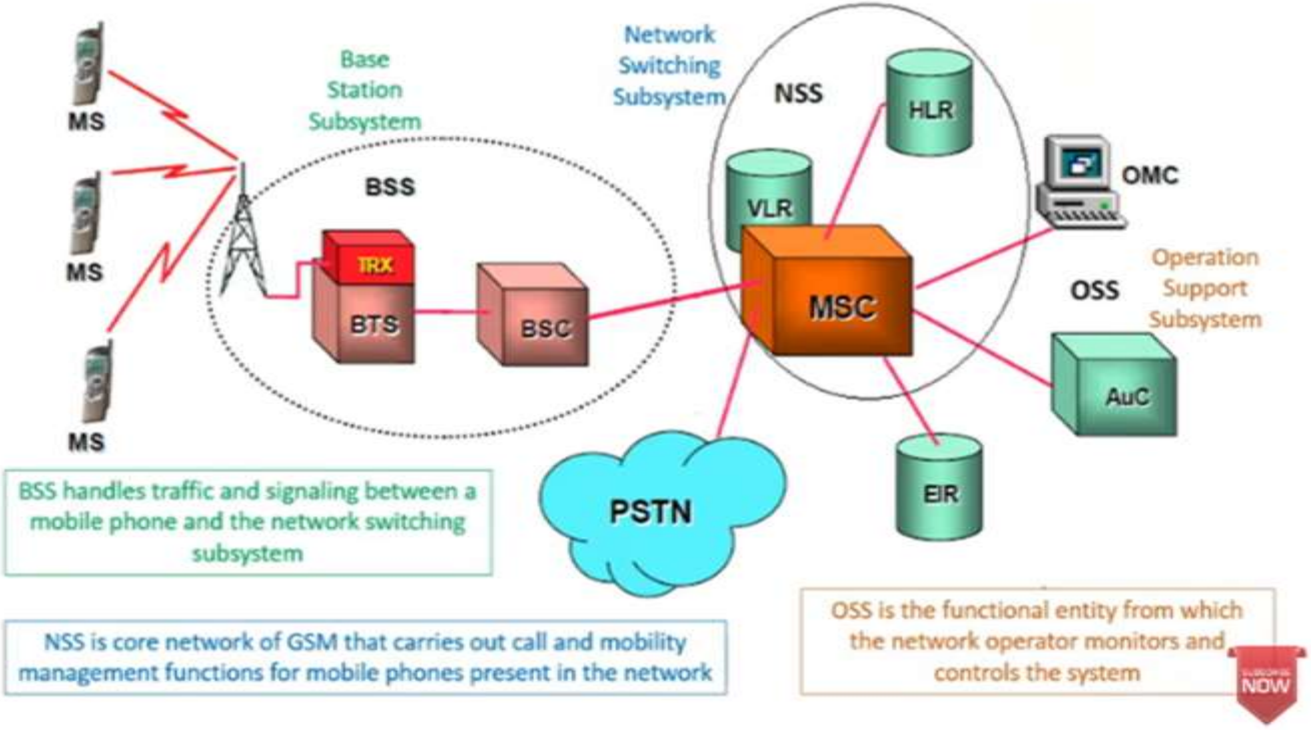
HLRs can manage data for several million customers and contain highly specialized data bases which must fulfill certain real-time requirements to answer requests within certain time-bounds.

VLR



Visitor location register (VLR): The VLR associated to each MSC is a dynamic database which stores all important information needed for the MS users currently in the LA that is associated to the MSC (e.g., IMSI, MSISDN, HLR address). If a new MS comes into an LA the VLR is responsible for, it copies all relevant information for this user from the HLR. This hierarchy of VLR and HLR avoids frequent HLR updates and long-distance signaling of user information.

OSS

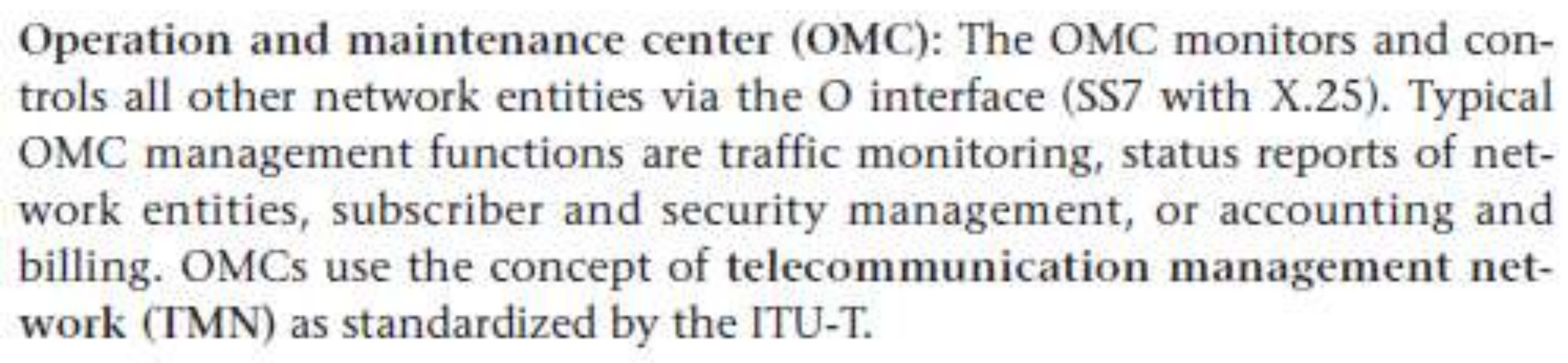


BSS handles traffic and signaling between a mobile phone and the network switching subsystem

NSS is core network of GSM that carries out call and mobility management functions for mobile phones present in the network

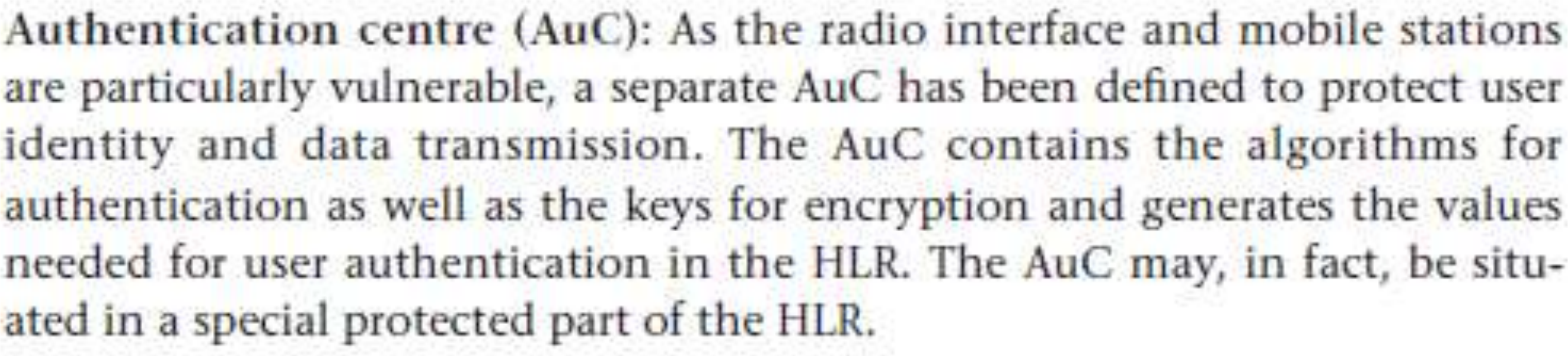
OSS is the functional entity from which the network operator monitors and controls the system

3. OSS



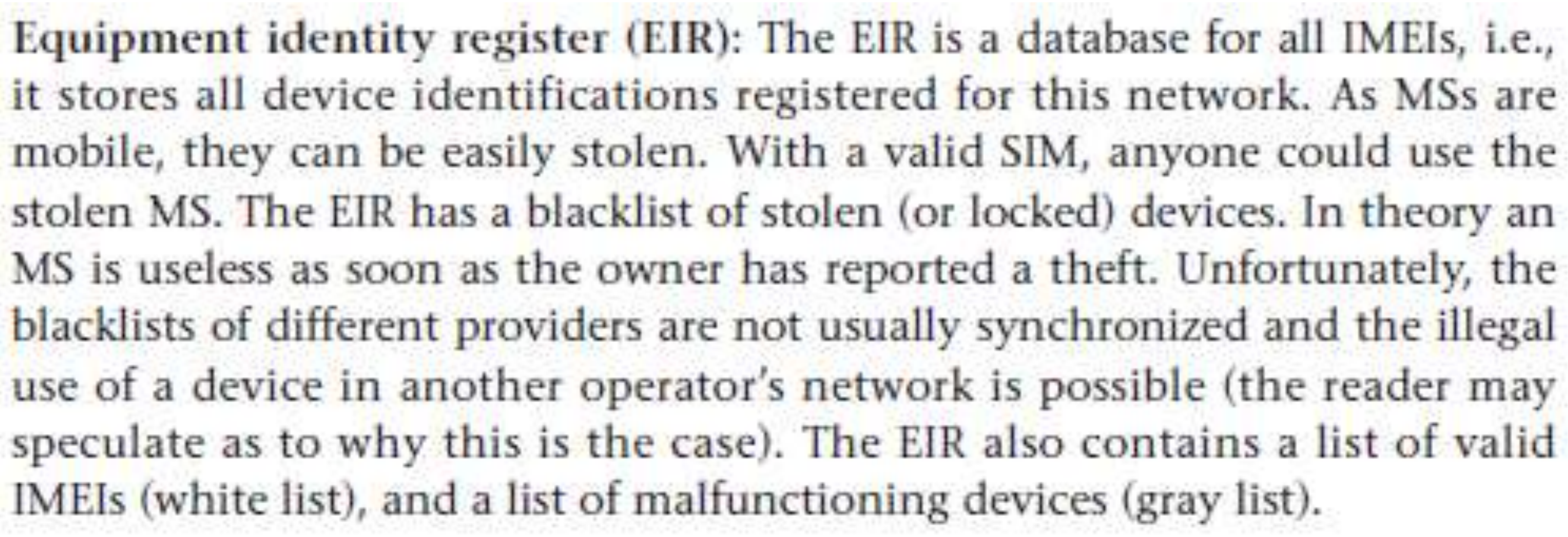
Operation and maintenance center (OMC): The OMC monitors and controls all other network entities via the O interface (SS7 with X.25). Typical OMC management functions are traffic monitoring, status reports of network entities, subscriber and security management, or accounting and billing. OMCs use the concept of **telecommunication management network (TMN)** as standardized by the ITU-T.

AuC



Authentication centre (AuC): As the radio interface and mobile stations are particularly vulnerable, a separate AuC has been defined to protect user identity and data transmission. The AuC contains the algorithms for authentication as well as the keys for encryption and generates the values needed for user authentication in the HLR. The AuC may, in fact, be situated in a special protected part of the HLR.

EIR



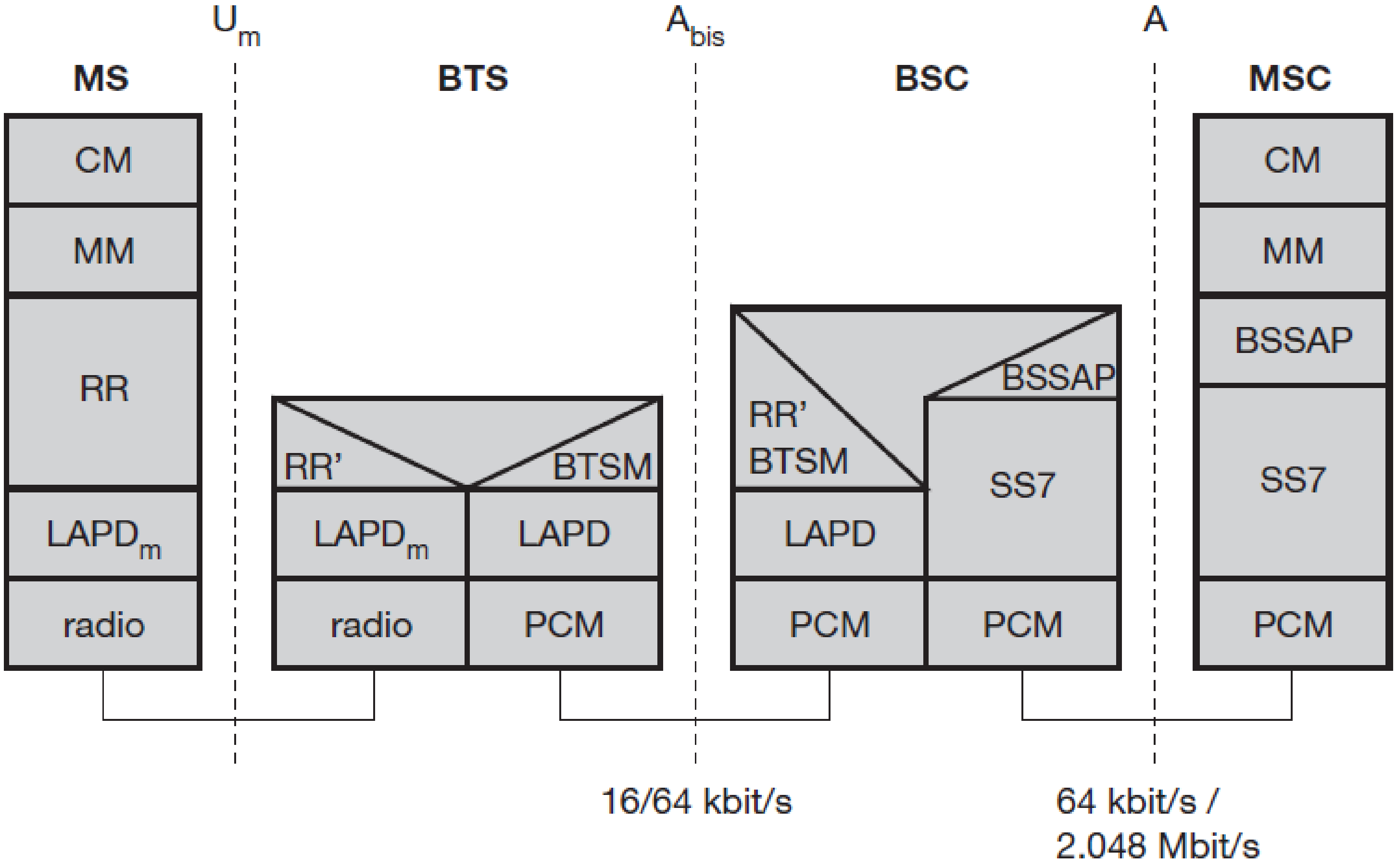
Equipment identity register (EIR): The EIR is a database for all IMEIs, i.e., it stores all device identifications registered for this network. As MSs are mobile, they can be easily stolen. With a valid SIM, anyone could use the stolen MS. The EIR has a blacklist of stolen (or locked) devices. In theory an MS is useless as soon as the owner has reported a theft. Unfortunately, the blacklists of different providers are not usually synchronized and the illegal use of a device in another operator's network is possible (the reader may speculate as to why this is the case). The EIR also contains a list of valid IMEIs (white list), and a list of malfunctioning devices (gray list).

How Does GSM Work?

- GSM is a globally recognized digital cellular communication protocol. The GSM standard was developed by the European Telecommunications Standards Institute to describe the procedures for second-generation digital mobile networks, such as those used by mobile phones. It is a broad-area communications technology programme that uses digital radio channeling to provide audio, information, and multimedia communication systems. Every GSM radio channel is 200 kHz broad and is further divided into frames of eight time slots. The GSM system consists of mobile stations, base stations, and interweaving switching systems.
- The GSM programme allows 8 to 16 audio users to share a single radio channel, and each radio transmission station can have numerous radio channels. Because of its simplicity, cost, and accessibility, GSM is now the most often utilized network technology in the Internet of Things (IoT).

GSM Part II

Protocol Architecture



- GSM architecture is a layered model that is designed to allow communications between two different systems. The lower layers assure the services of the upper-layer protocols. Each layer passes suitable notifications to ensure the transmitted data has been formatted, transmitted, and received accurately.

CM is further subdivided into three protocol entities: Call Control (CC), Supplementary Services (SS) and SMS. Additional multiplexing functions within Layer 3 are required between these sublayers.

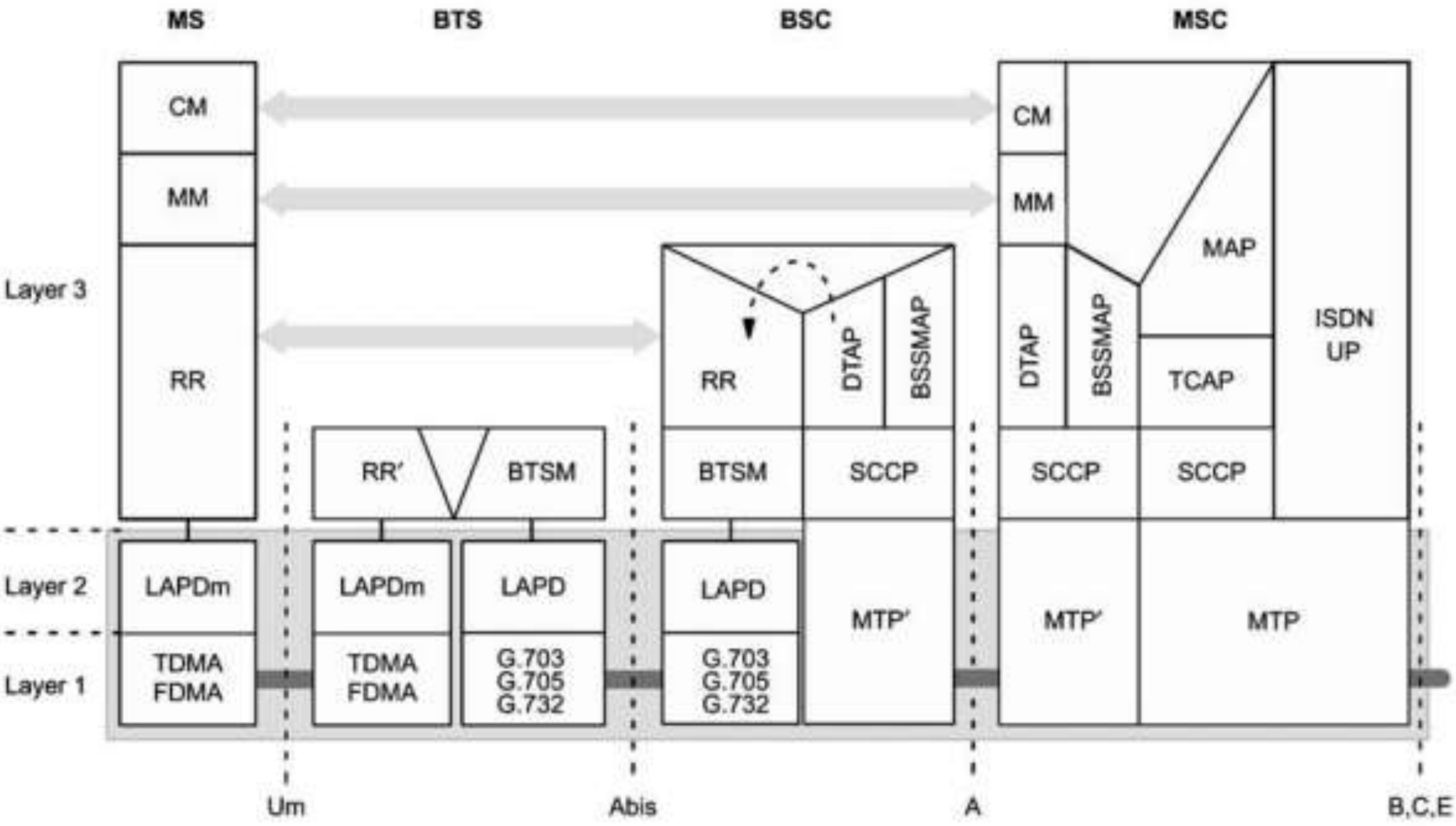
The call-independent SS and the SMS are offered to higher layers at two Service Access Points (SAPs), MNSS and MNSMS.

- Based on the interface, the GSM signaling protocol is assembled into three general layers −
- **Layer 1** − The physical layer. It uses the channel structures over the air interface.
- **Layer 2** − The data-link layer. Across the Um interface, the data-link layer is a modified version of the Link access protocol for the D channel (LAP-D) protocol used in ISDN, called Link access protocol on the D channel (LAP-D). Across the A interface, the Message Transfer Part (MTP), Layer 2 of SS7 is used.
- **Layer 3** − GSM signaling protocol’s third layer is divided into three sub layers −
  - Radio Resource Management (RR),
  - Mobility Management (MM), and
  - Connection Management (CM).

- The RR layer is the lower layer that manages a link, both radio and fixed, between the MS and the MSC. For this formation, the main components involved are the MS, BSS, and MSC. The responsibility of the RR layer is to manage the RR-session, the time when a mobile is in a dedicated mode, and the radio channels including the allocation of dedicated channels.

- The MM layer is stacked above the RR layer. It handles the functions that arise from the mobility of the subscriber, as well as the authentication and security aspects. Location management is concerned with the procedures that enable the system to know the current location of a powered-on MS so that incoming call routing can be completed.

- The CM layer is the topmost layer of the GSM protocol stack. This layer is responsible for Call Control, Supplementary Service Management, and Short Message Service Management. Each of these services are treated as individual layer within the CM layer. Other functions of the CC sub layer include call establishment, selection of the type of service (including alternating between services during a call), and call release.



CM	Connection Management
MM	Mobility Management
RR	Radio Resource Management
LAPDm	Link Protocol (specially adaptet for air interface Um)
BTSM	Base Transceiving Station Management

BSSMAP	Base Station System Management Application Part
DTAP	Direct Transfer Application Part
SCCP	Signaling Connection Control Part
TCAP	Transaction Capabilities Application Part
MTP	Message Transfer Part
MAP	Mobile Application Part
UP	User Part

- Physical layer
- In the control plane, the lowest layer of the protocol model at the air interface, the physical
- layer, implements the logical signaling channels (TDMA/FDMA, multiframes, channel
- coding, etc

- Layer 2: LAPDm
- On Layer 2 of the logical signaling channels across the air interface a data link protocol
- entity is implemented, the Link Access Procedure on Dm Channels (LAPDm). LAPDm is
- a derivative of Link Access Protocol Channel D (LAPD) which is specifically adapted to
- the air interface. This data link protocol is responsible for the protected transfer of signaling
- messages between MS and BTS over the air interface, i.e. LAPDm is terminated in MS and
- base station.

The main task of LAPDm is the transparent transport of messages between protocol entities of Layer 3 with special support for:

- multiple entities in Layer 3 and Layer 2;
- signaling for broadcasting (BCCH);
- signaling for paging (PCH);
- signaling for channel assignment (AGCH);
- signaling on dedicated channels (SDCCH).

Layer 3

- In the MS, the LAPDm services are used at Layer 3 of the signaling protocol architecture.
- Layer 3 is divided into three sublayers: Radio Resource Management (RR), Mobility
- Management (MM) and Connection Management (CM)

Radio resource management

The general objective of the RR is to set up, maintain and take down RR connections which enable point-to-point communication between MS and network. This also includes cell selection in idle mode and handover procedures. Furthermore, the RR is responsible for monitoring BCCH and CCCH on the downlink when no RR connections are active.

The following functions are realized in the RR module:

- monitoring of BCCH and PCH (readout of system information and paging messages)
- RACH administration: MSs send their requests for connections and replies to paging announcements to the BSS
- requests for and assignments of data and signaling channels
- periodic measurement of channel quality (quality monitoring)
- transmitter power control and synchronization of the MS

Mobility management

MM encompasses all of the tasks resulting from mobility. The MM activities are exclusively

performed in cooperation between MS and MSC, and they include:

- TMSI assignment;
- localization of the MS;
- location updating of the MS, sometimes known as roaming functions
- identification of the MS (IMSI, IMEI);
- authentication of the MS;
- IMSI attach and detach procedures (at insertion or removal of SIM);
- ensuring confidentiality of subscriber identity

Connection management

CM consists of three entities: CC, SS and SMS. CC handles all tasks related to setting up,

maintaining and taking down calls. The services of CC are provided at the MNCC-SAP, and

they encompass:

• establishment of normal calls (MS-originating and MS-terminating);

• establishment of emergency calls (only MS-originating);

136 GSM – ARCHITECTURE, PROTOCOLS AND SERVICES

• termination of calls;

• Dual-Tone Multifrequency (DTMF) signaling;

• call-related supplementary services;

• in-call modification: the service may be changed during a connection (e.g. speech and

transparent/nontransparent data are alternating; or speech and fax alternate)

BSC Protocols

- The BSC uses a different set of protocols after receiving the data from the BTS. The Abis interface is used between the BTS and BSC. At this level, the radio resources at the lower portion of Layer 3 are changed from the RR to the Base Transceiver Station Management (BTSM). The BTS management layer is a relay function at the BTS to the BSC.
- The RR protocols are responsible for the allocation and reallocation of traffic channels between the MS and the BTS. These services include controlling the initial access to the system, paging for MT calls, the handover of calls between cell sites, power control, and call termination. The BSC still has some radio resource management in place for the frequency cordination, frequency allocation, and the management of the overall network layer for the Layer 2 interfaces.
- To transit from the BSC to the MSC, the BSS mobile application part or the direct application part is used.

MSC Protocols

- At the MSC, starting from the BSC, the information is mapped across the A interface .Here, Base Station System Management Application Part (BSS MAP) is said to be the equivalent set of radio resources. The relay process is finished by the layers that are stacked on top of Layer 3 protocols, they are BSS MAP/DTAP, MM, and CM. This completes the relay process. To find and connect to the users across the network, MSCs interact using the control signaling network. Location registers are included in the MSC databases to assist in the role of determining how and whether connections are to be made to roaming users.
- Each GSM MS user is given a HLR that in turn comprises of the user's location and subscribed services. VLR is a separate register that is used to track the location of a user. When the users move out of the HLR covered area, the VLR is notified by the MS to find the location of the user. The VLR in turn, with the help of the control network, signals the HLR of the MS's new location. With the help of location information contained in the user's HLR, the MT calls can be routed to the user.

LOCALIZATION

Localization is the process by which MS is identified, authenticated and provide service by MSC.

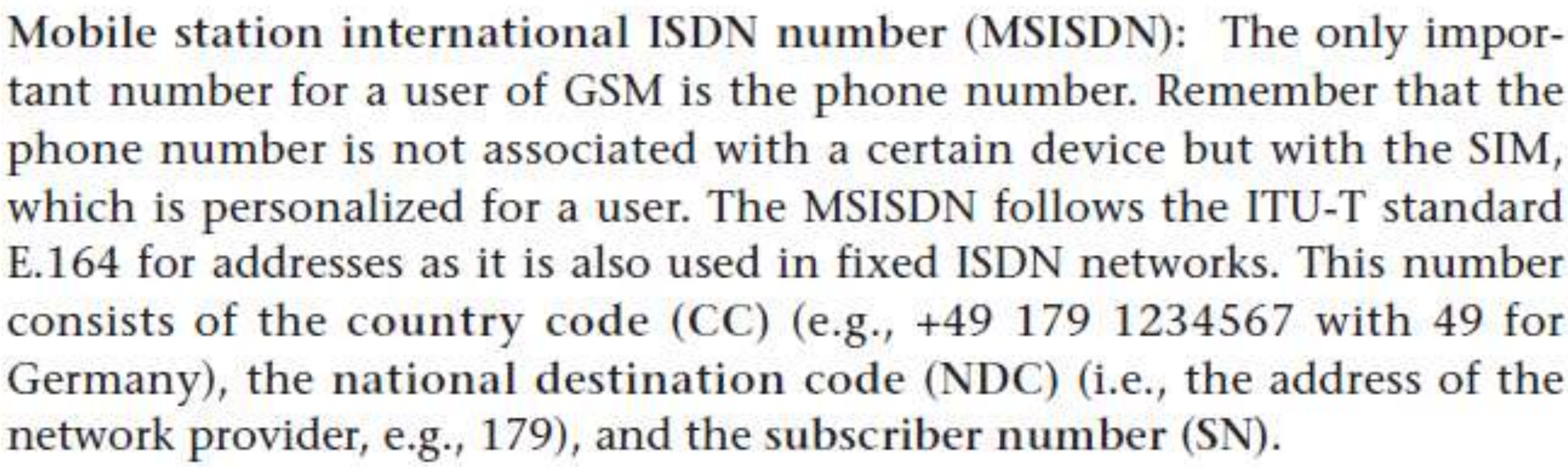
When users are moving service providers provide service only after

- Identifying MS
- Verifying Services subscribed by user

Following Numbers are needed to Locate and address MS

1. MSISDN
2. IMSI
3. TMSI
4. MSRN

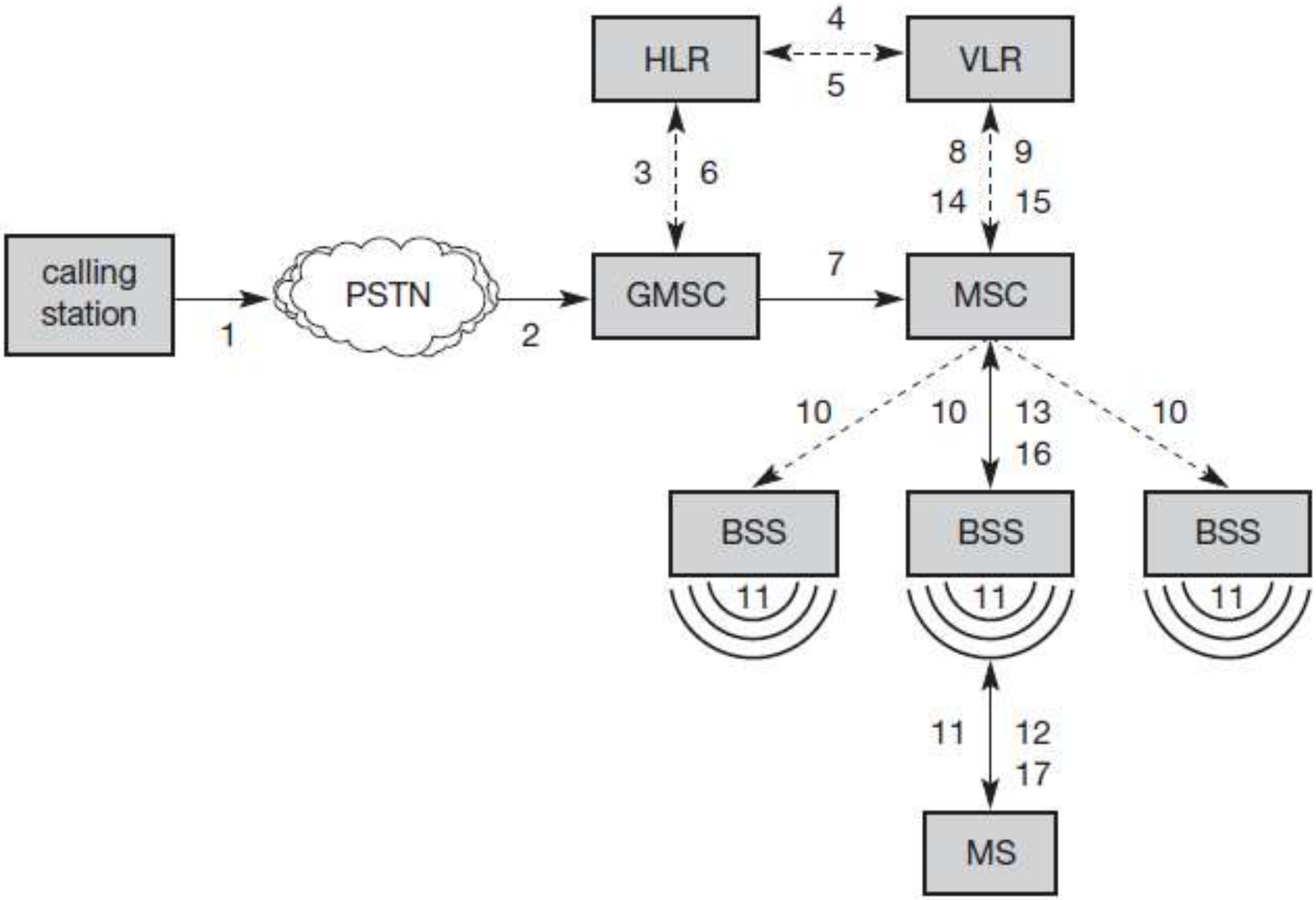
Localisation and Calling



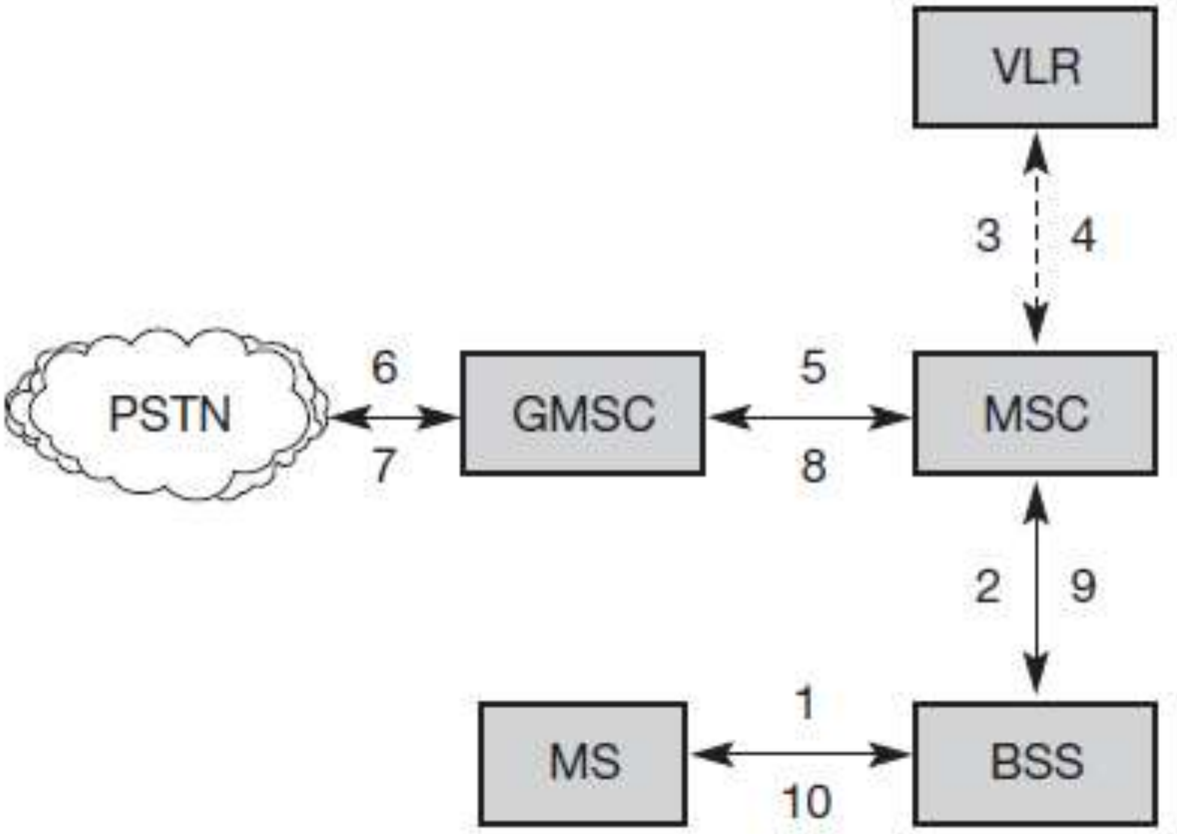
Mobile station international ISDN number (MSISDN): The only important number for a user of GSM is the phone number. Remember that the phone number is not associated with a certain device but with the SIM, which is personalized for a user. The MSISDN follows the ITU-T standard E.164 for addresses as it is also used in fixed ISDN networks. This number consists of the country code (CC) (e.g., +49 179 1234567 with 49 for Germany), the national destination code (NDC) (i.e., the address of the network provider, e.g., 179), 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) (e.g., 240 for Sweden, 208 for France), the mobile network code (MNC) (i.e., the code of the network provider), 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 signaling over the air interface, GSM uses the 4 byte TMSI for local subscriber identification. TMSI is selected by the current VLR and is only valid temporarily and within the location area of the VLR (for an ongoing communication TMSI and LAI are sufficient to identify a user; the IMSI is not needed). Additionally, a VLR may change the TMSI periodically.
- Mobile station[7] 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.

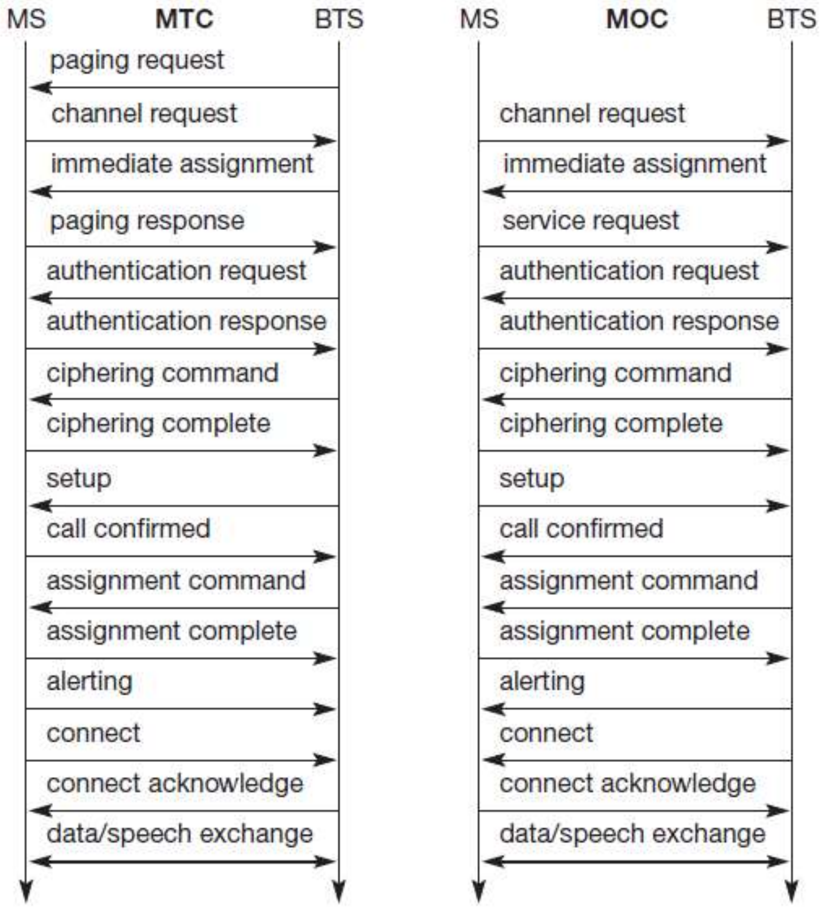
Mobile Terminated Call



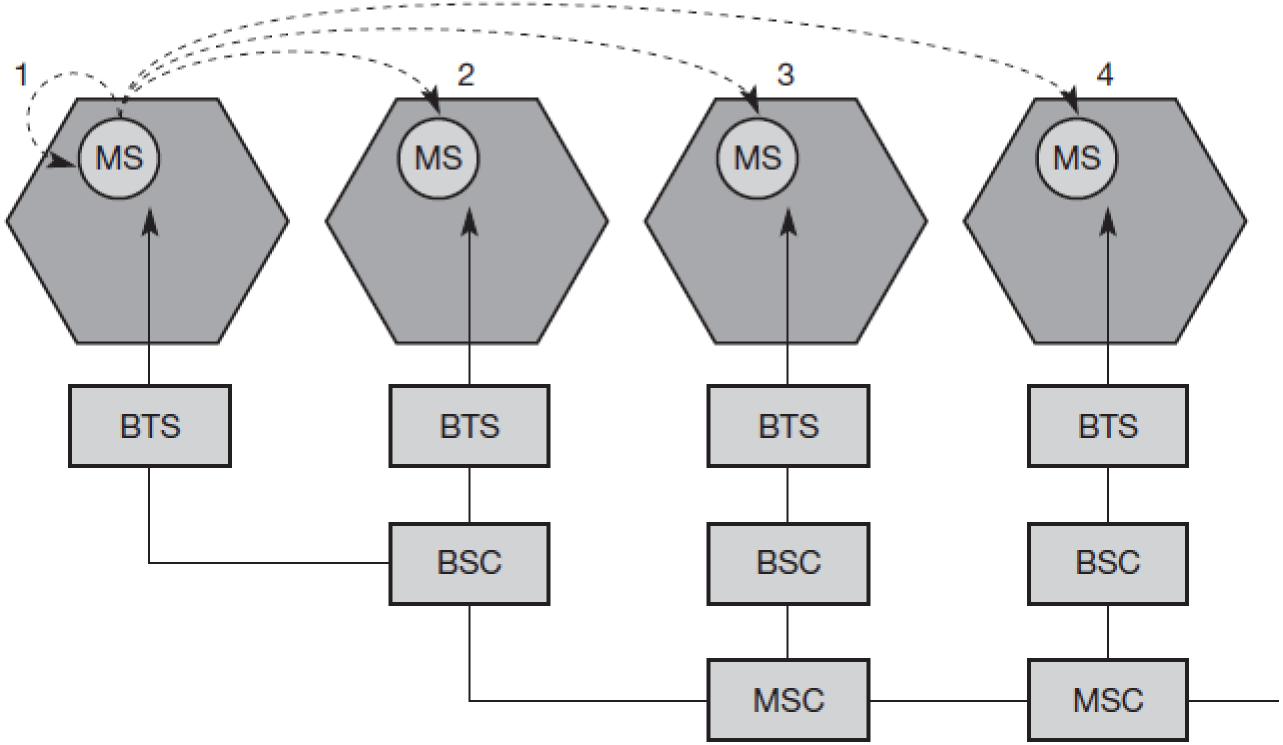
Mobile Originated call



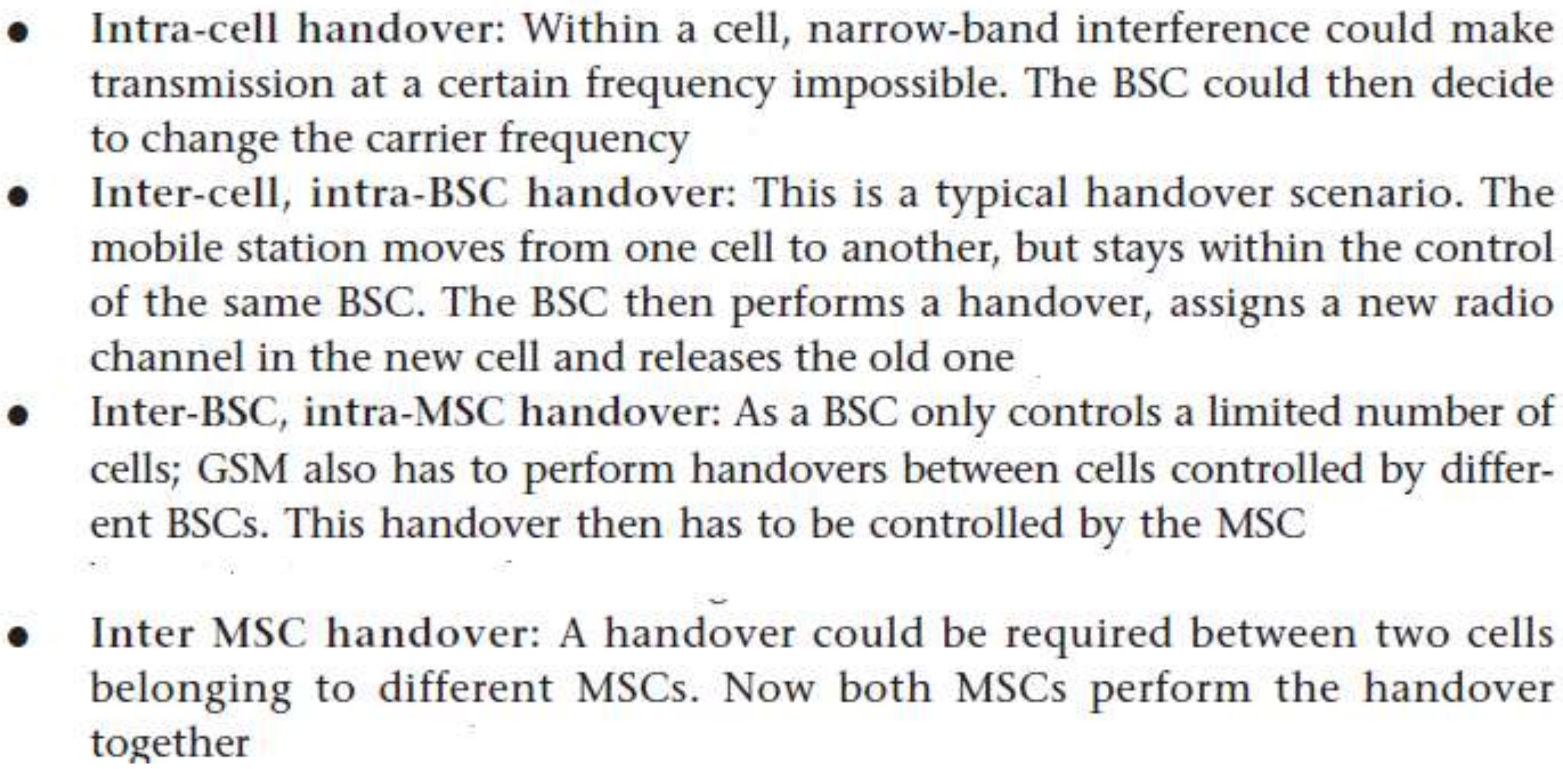
Message Flow



Handover



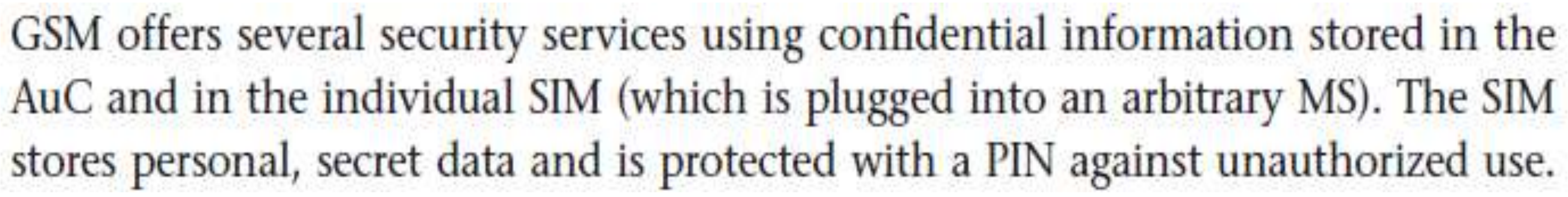
Types of Handover



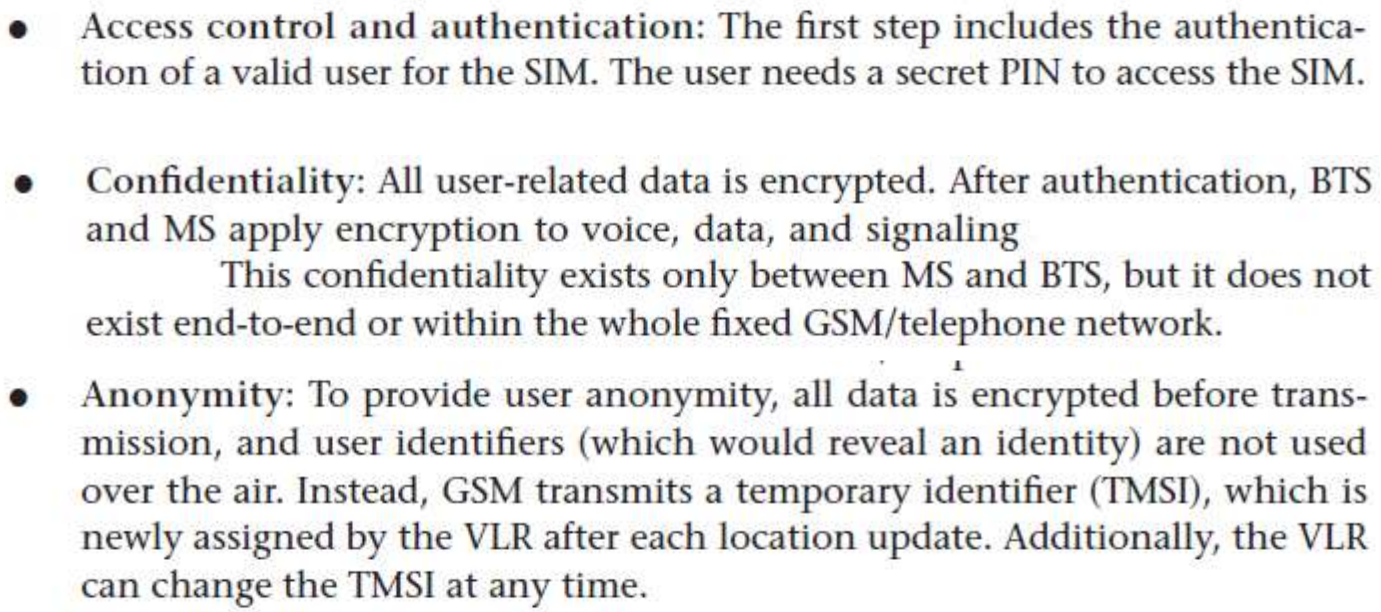
- **Intra-cell handover:** Within a cell, narrow-band interference could make transmission at a certain frequency impossible. The BSC could then decide to change the carrier frequency
- **Inter-cell, intra-BSC handover:** This is a typical handover scenario. The mobile station moves from one cell to 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
- **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
- **Inter MSC handover:** A handover could be required between two cells belonging to different MSCs. Now both MSCs perform the handover together

http://vlabs.iitkgp.ac.in/fcmc/exp8/index.html

Security



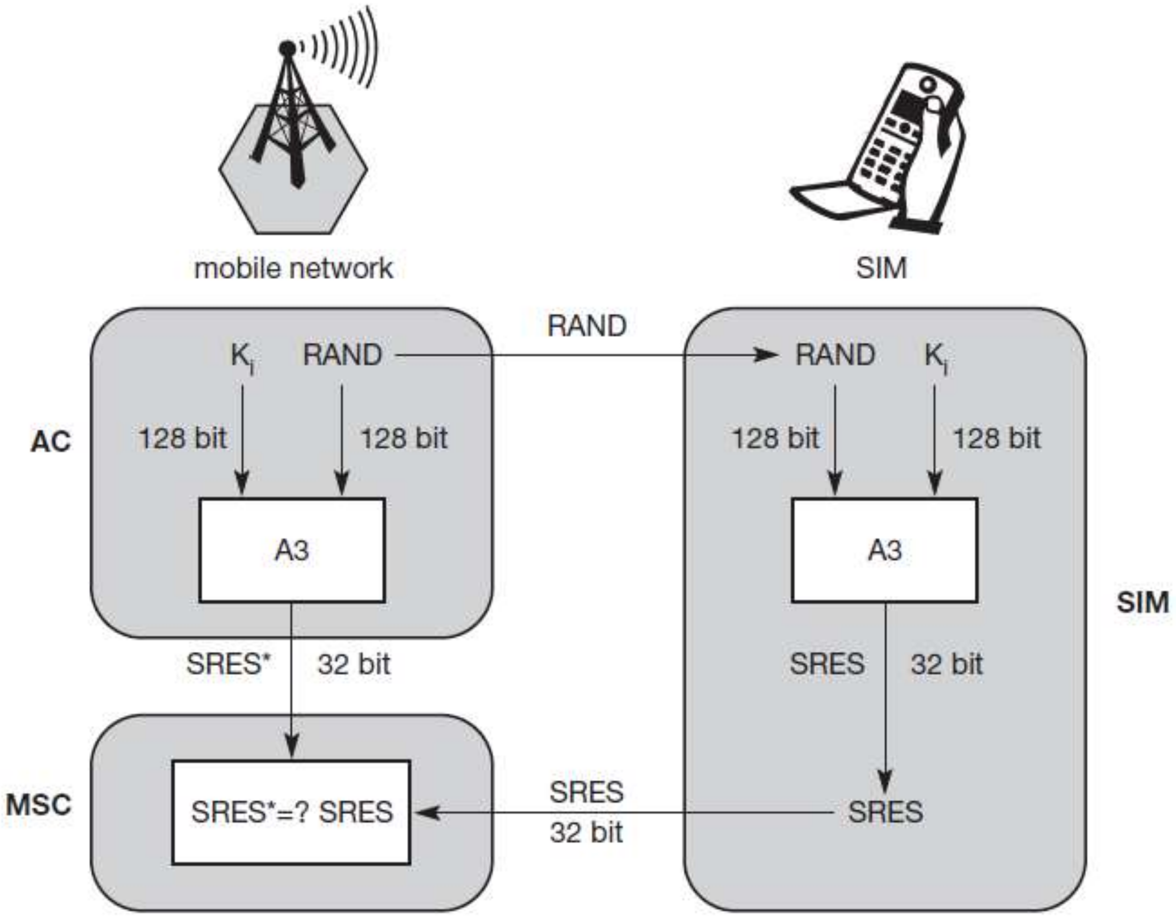
GSM offers several security services using confidential information stored in the AuC and in the individual SIM (which is plugged into an arbitrary MS). The SIM stores personal, secret data and is protected with a PIN against unauthorized use.



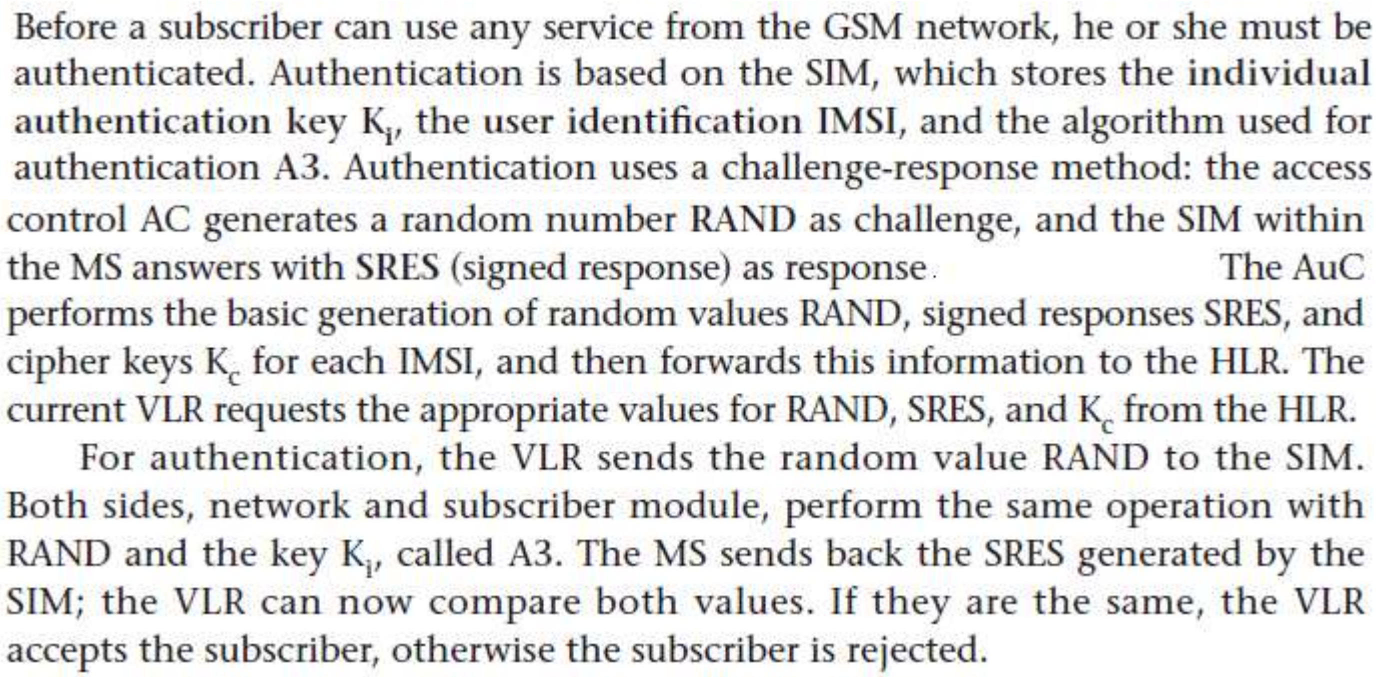
- Access control and authentication: The first step includes the authentication of a valid user for the SIM. The user needs a secret PIN to access the SIM.
- Confidentiality: All user-related data is encrypted. After authentication, BTS and MS apply encryption to voice, data, and signaling
  This confidentiality exists only between MS and BTS, but it does not exist end-to-end or within the whole fixed GSM/telephone network.
- Anonymity: To provide user anonymity, all data is encrypted before transmission, and user identifiers (which would reveal an identity) 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.

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. In the GSM standard only algorithm A5 was publicly available, whereas A3 and A8 were secret, but standardized with open interfaces. Both A3 and A8 are no longer secret, but were published on the internet in 1998. This demonstrates that security by obscurity does not really work. As it turned out, the algorithms are not very strong. However, network providers can use stronger algorithms for authentication – or users can apply stronger end-to-end encryption. Algorithms A3 and A8 (or their replacements) are located on the SIM and in the AuC and can be proprietary. Only A5 which is implemented in the devices has to be identical for all providers.

Subscriber Authentication



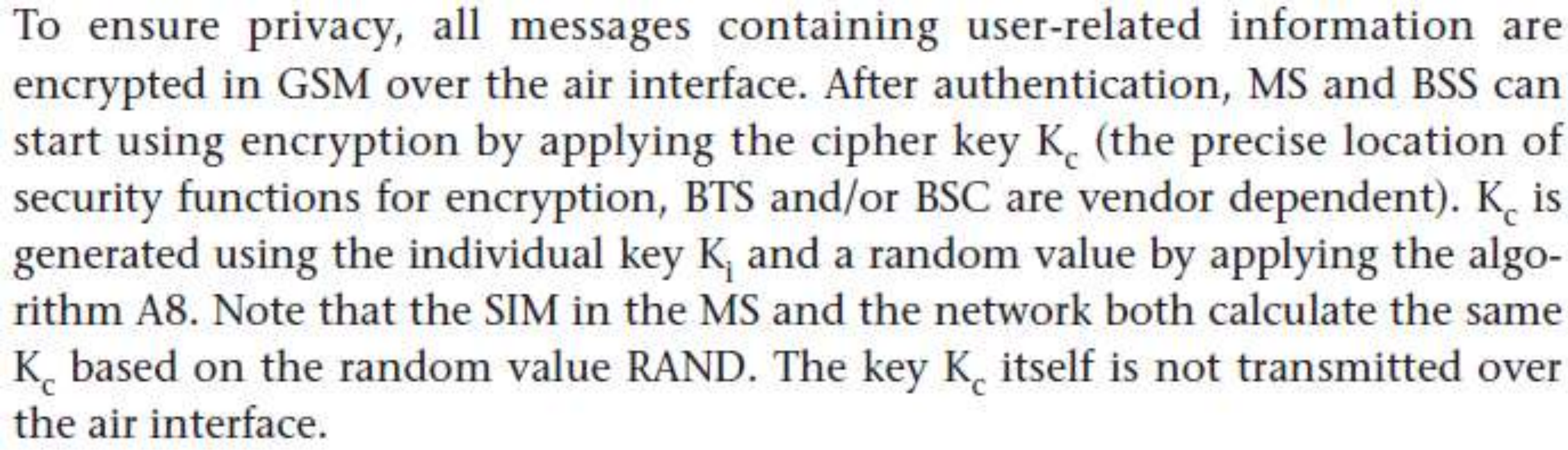
Authentication



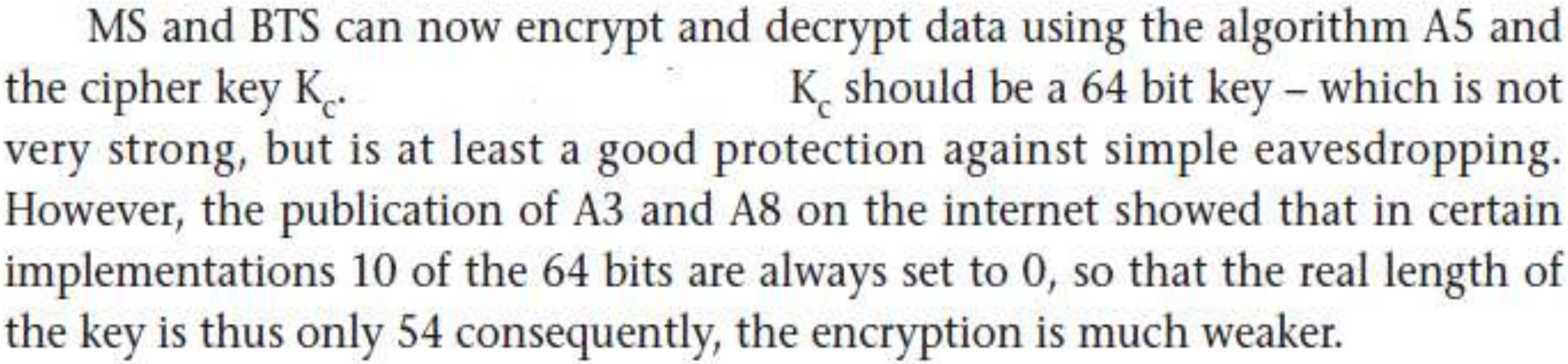
Before a subscriber can use any service from the GSM network, he or she must be authenticated. Authentication is based on the SIM, which stores the individual authentication key K_i, the user identification IMSI, and the algorithm used for authentication A3. Authentication uses a challenge-response method: the access control AC generates a random number RAND as challenge, and the SIM within the MS answers with SRES (signed response) as response. The AuC performs the basic generation of random values RAND, signed responses SRES, and cipher keys K_c for each IMSI, and then forwards this information to the HLR. The current VLR requests the appropriate values for RAND, SRES, and K_c 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 K_i, 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.

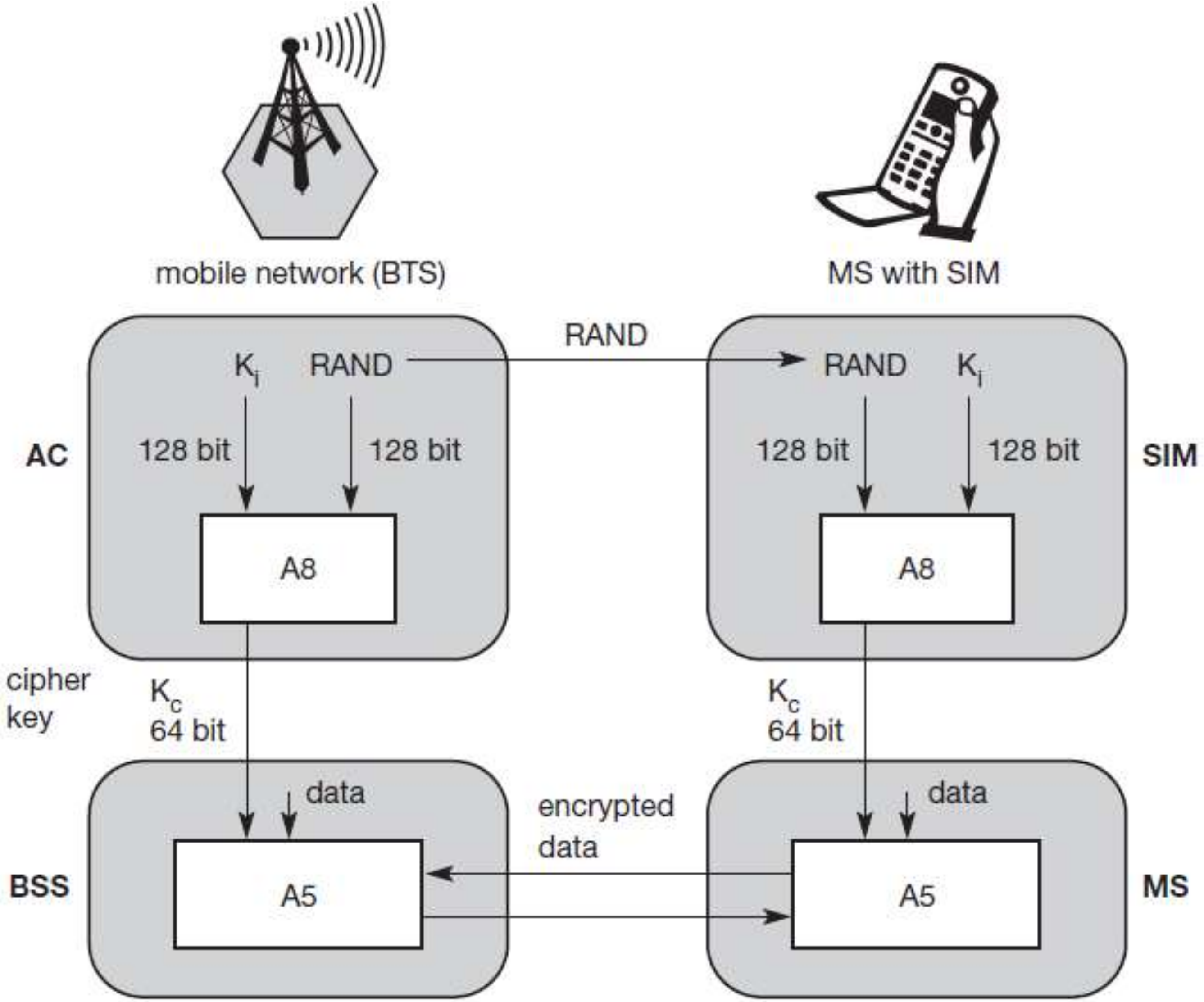
Encryption



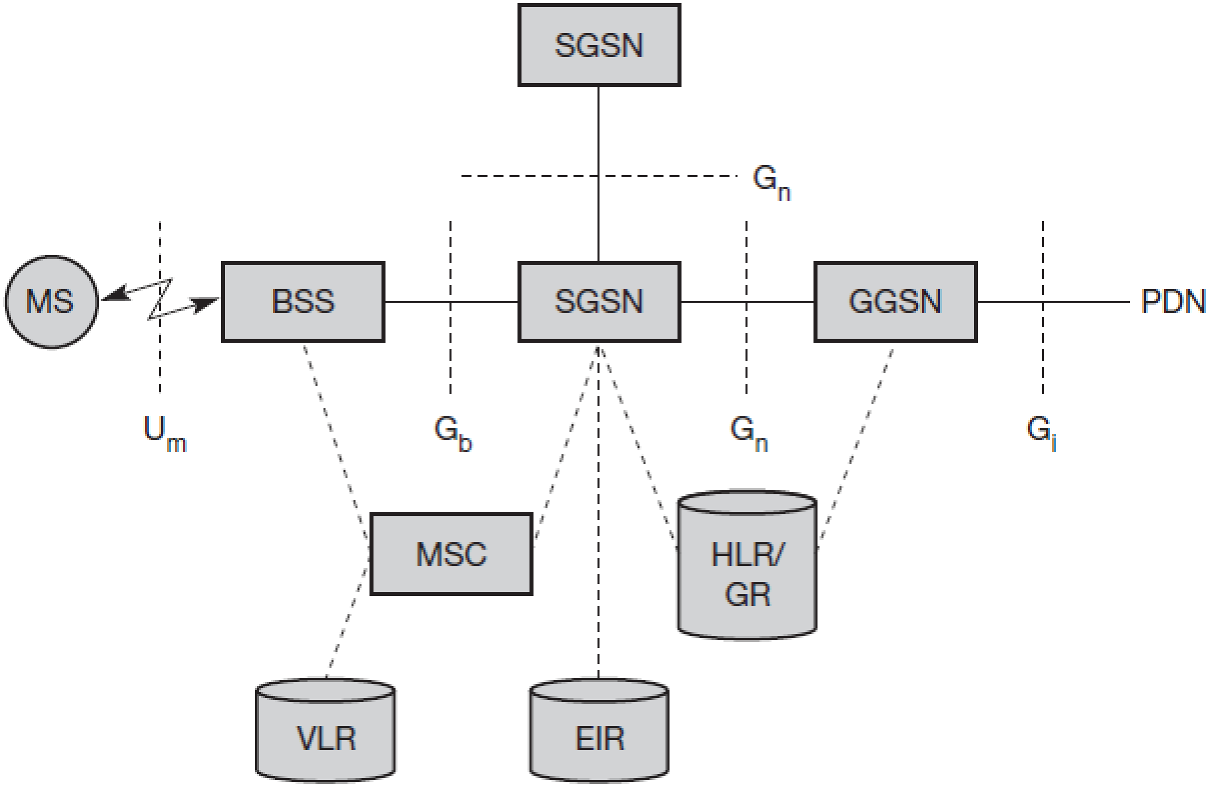
To ensure privacy, all messages containing user-related information are encrypted in GSM over the air interface. After authentication, MS and BSS can start using encryption by applying the cipher key K_c (the precise location of security functions for encryption, BTS and/or BSC are vendor dependent). K_c is generated using the individual key K_i and a random value by applying the algorithm A8. Note that the SIM in the MS and the network both calculate the same K_c based on the random value RAND. The key K_c 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 K_c. K_c should be a 64 bit key – which is not very strong, but is at least a good protection against simple eavesdropping. However, the publication of A3 and A8 on the internet showed that in certain implementations 10 of the 64 bits are always set to 0, so that the real length of the key is thus only 54 consequently, the encryption is much weaker.


mobile network (BTS)
MS with SIM
AC
K_i
RAND
128 bit
128 bit
A8
RAND
RAND
K_i
128 bit
128 bit
A8
SIM
cipher key
K_c 64 bit
K_c 64 bit
data
encrypted data
data
BSS
A5
A5
MS

GPRS



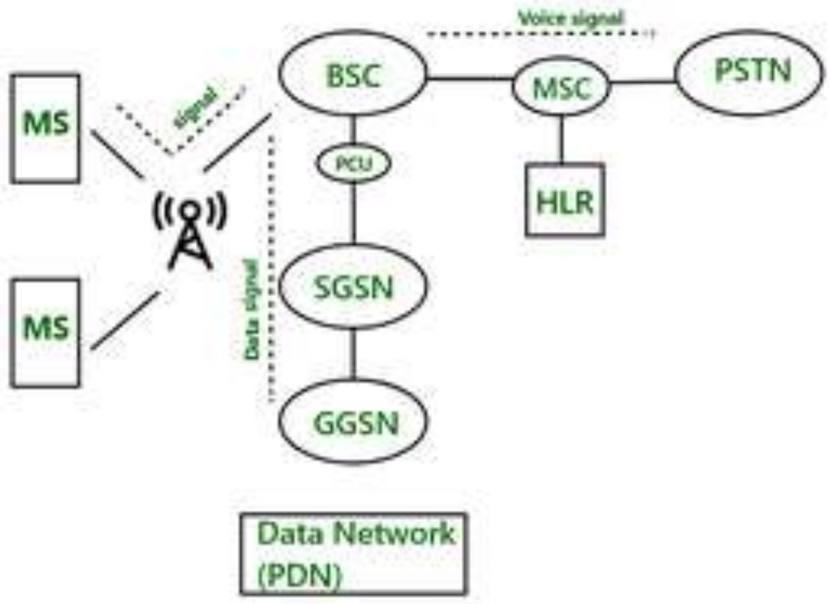
The GPRS architecture introduces two new network elements, which are called GPRS support nodes (GSN) and are in fact routers. All GSNs are integrated into the standard GSM architecture, and many new interfaces have been defined The gateway GPRS support node (GGSN) is the interworking unit between the 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 G_i interface and transfers packets to the SGSN via an IP-based GPRS backbone network (G_n interface).

The other new element is the serving GPRS support node (SGSN) which supports the MS via the G_b 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. GGSNs and SGSNs can be compared with home and foreign agents, respectively, in a mobile IP network

As shown in Figure , 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. This context comprises the status of the MS (which can be ready, idle, or standby; ETSI, 1998b), the CKSN, a flag indicating if compression is used, and routing data (TLLI, the routing area RA, a cell identifier, and a packet data channel, PDCH, identifier). Besides attaching and detaching, mobility management also comprises functions for authentication, location management, and ciphering (here, the scope of ciphering lies between MS and SGSN, which is more than in standard GSM). In idle mode an MS is not reachable and all context is deleted. In the standby state only movement across routing areas is updated to the SGSN but not changes of the cell. Permanent updating would waste battery power, no updating would require system-wide paging. The update procedure in standby mode is a compromise. Only in the ready state every movement of the MS is indicated to the SGSN.

In GSM architecture there is one component called BSC. But in GPRS there is one component is added to BSC called PCU. PCU stands for Packet Control Unit. If signal comes to BSC and that signal contains data, then PCU routes to the SGSN. Interface is used between BSC and PCU is FRI interface. After signal comes to SGSN, it delivers the data packet to the GGSN. GGSN routes the data packet to the data network (PDN-Predefined Data Network).



GPRS Architecture

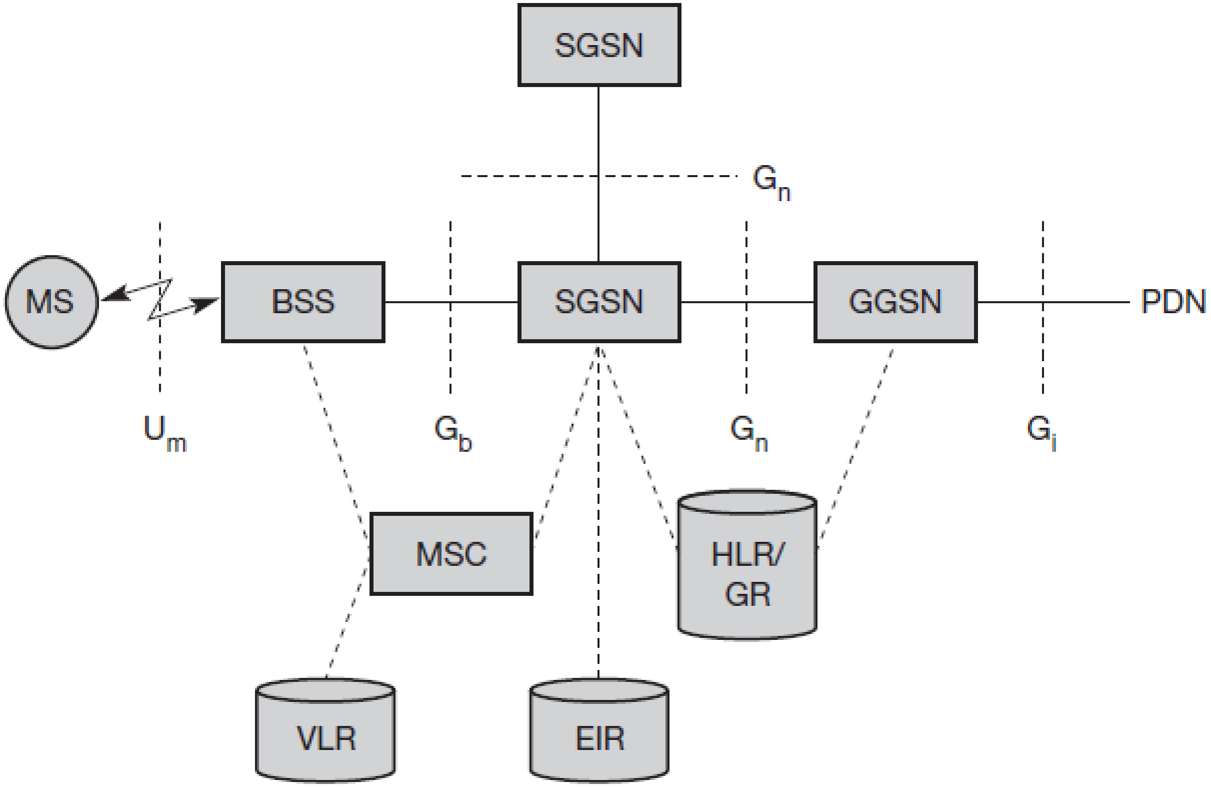
Task of SGSN :

- Packet Delivery
- Mobility management
  - apply/ sign off of terminals
  - localization
- LLC (Logical Link Control) management
- Authentication
- billing

Task of GGSN :

- Mediator between GPRS between backbone and external data networks.
- Saves current data for the SGSN address of the participant as well as their profile and data for authentication and invoice.

GPRS



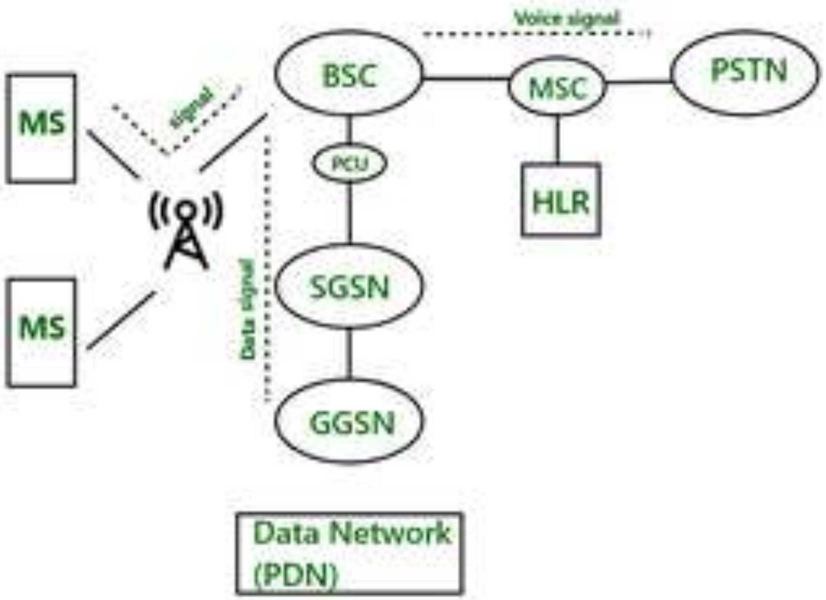
The GPRS architecture introduces two new network elements, which are called GPRS support nodes (GSN) and are in fact routers. All GSNs are integrated into the standard GSM architecture, and many new interfaces have been defined The gateway GPRS support node (GGSN) is the interworking unit between the 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 G_i interface and transfers packets to the SGSN via an IP-based GPRS backbone network (G_n interface).

The other new element is the serving GPRS support node (SGSN) which supports the MS via the G_b 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. GGSNs and SGSNs can be compared with home and foreign agents, respectively, in a mobile IP network

As shown in Figure , 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. This context comprises the status of the MS (which can be ready, idle, or standby; ETSI, 1998b), the CKSN, a flag indicating if compression is used, and routing data (TLLI, the routing area RA, a cell identifier, and a packet data channel, PDCH, identifier). Besides attaching and detaching, mobility management also comprises functions for authentication, location management, and ciphering (here, the scope of ciphering lies between MS and SGSN, which is more than in standard GSM). In idle mode an MS is not reachable and all context is deleted. In the standby state only movement across routing areas is updated to the SGSN but not changes of the cell. Permanent updating would waste battery power, no updating would require system-wide paging. The update procedure in standby mode is a compromise. Only in the ready state every movement of the MS is indicated to the SGSN.

In GSM architecture there is one component called BSC. But in GPRS there is one component is added to BSC called PCU. PCU stands for Packet Control Unit. If signal comes to BSC and that signal contains data, then PCU routes to the SGSN. Interface is used between BSC and PCU is FRI interface. After signal comes to SGSN, it delivers the data packet to the GGSN. GGSN routes the data packet to the data network (PDN-Predefined Data Network).



GPRS Architecture

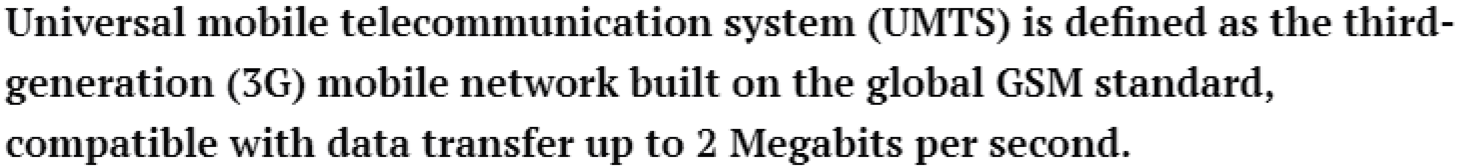
Task of SGSN :

- Packet Delivery
- Mobility management
  - apply/ sign off of terminals
  - localization
- LLC (Logical Link Control) management
- Authentication
- billing

Task of GGSN :

- Mediator between GPRS between backbone and external data networks.
- Saves current data for the SGSN address of the participant as well as their profile and data for authentication and invoice.

UMTS

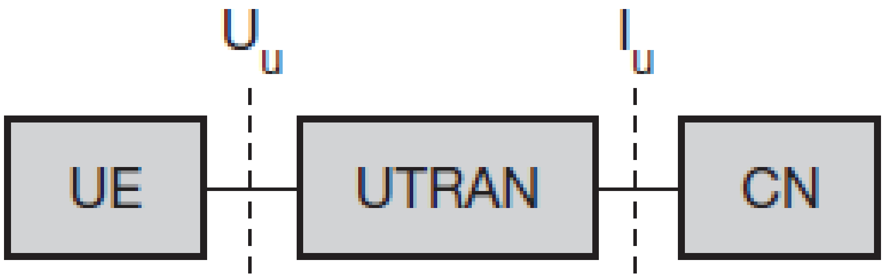


Universal mobile telecommunication system (UMTS) is defined as the third-generation (3G) mobile network built on the global GSM standard, compatible with data transfer up to 2 Megabits per second.

The **UTRA network (UTRAN)** handles cell level mobility and comprises several **radio network subsystems (RNS)**.

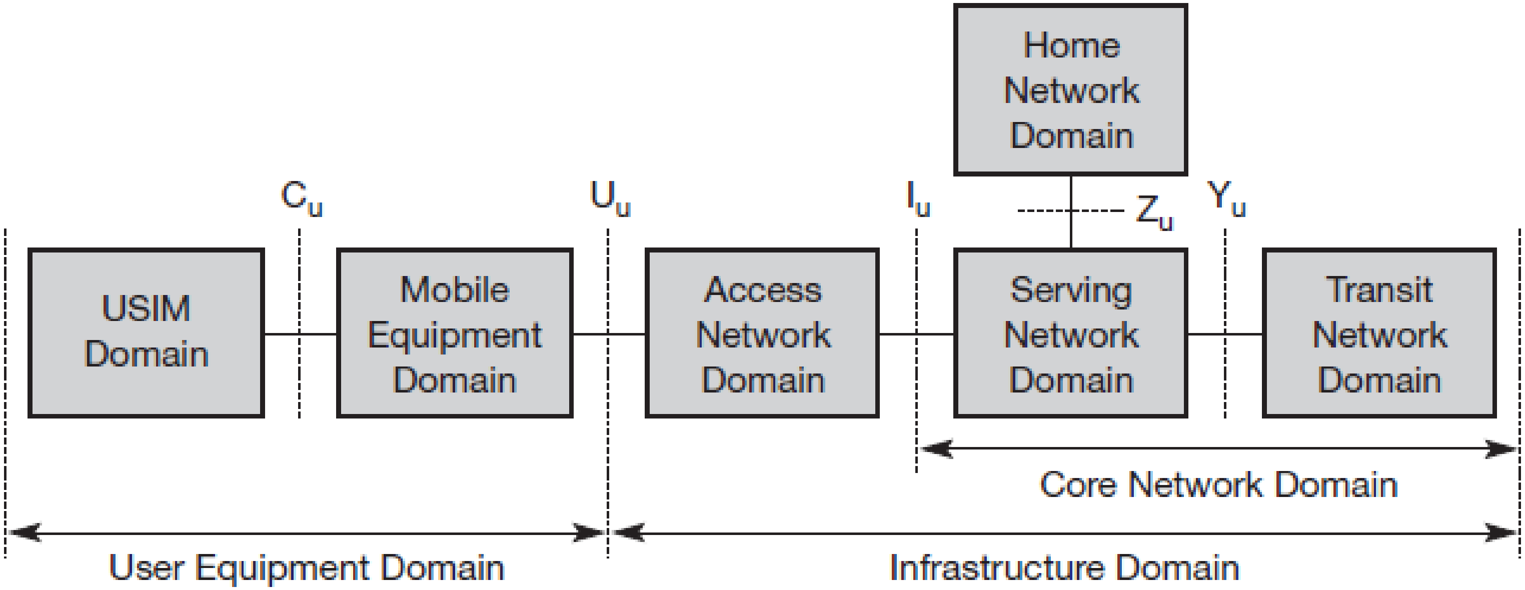
The functions of the RNS include radio channel ciphering and deciphering, handover control, radio resource management

UMTS system architecture



- ***User Equipment (UE):*** The User Equipment or UE is the name given to what was previous termed the mobile, or cellphone. The new name was chosen because the considerably greater functionality that the UE could have. It could also be anything between a mobile phone used for talking to a data terminal attached to a computer with no voice capability.
- ***Radio Network Subsystem (RNS):*** The RNS also known as the UMTS Radio Access Network, UTRAN, was the equivalent of the previous Base Station Subsystem or BSS in GSM. It provided and manages the air interface for the overall network.
- ***Core Network:*** The core network provided all the central processing and management for the system. It was the equivalent of the GSM Network Switching Subsystem or NSS.

UMTS DOMAINS



- The **user equipment** domain is assigned to a single user and comprises all the functions that are needed to access UMTS services. Within this domain are the USIM domain and the mobile equipment domain. The **USIM** domain contains the SIM for UMTS which performs functions for encryption and authentication of users, and stores all the necessary user-related data for UMTS. Typically, this USIM belongs to a service provider and contains a micro processor for an enhanced program execution environment (USAT, UMTS SIM application toolkit). All functions for radio transmission as well as user interfaces are located here.

- The **infrastructure** domain is shared among all users and offers UMTS services to all accepted users. This domain consists of the **access network** domain, which contains the radio access networks (RAN), and the core network domain, which contains access network independent functions.

The **core network** domain can be separated into three domains with specific tasks.

- The **serving network** domain comprises all functions currently used by a user for accessingUMTS services.
- All functions related to the home network of a user, e.g., user data look-up, fall into the **home network** domain.
- The **transit network** domain may be necessary if the serving network cannot directly contact the home network.

All three domains within the core network may be in fact the same physical network. These domains only describe functionalities

User Equipment(UE)

- The USER Equipment or UE was a major element of the overall 3G UMTS network architecture. It formed the final interface with the user.
- In view of the far greater number of applications and facilities that it could perform, the decision was made to call it a user equipment rather than a mobile.

it consists of a variety of different elements including RF circuitry, processing, antenna, battery, etc.

There were a number of elements within the UE that can be described separately:

1.UE RF circuitry: The RF areas handled all elements of the signal, both for the receiver and for the transmitter. One of the major challenges for the RF power amplifier was to reduce the power consumption.

The form of modulation used for W-CDMA required the use of an RF linear amplifier. These inherently take more current than non linear amplifiers which could be used for the form of modulation used on GSM.

Accordingly to maintain battery life, measures were introduced into many of the designs to ensure the optimum efficiency.

2. Baseband processing: The base-band signal processing consisted mainly of digital circuitry. This was considerably more complicated than that used in phones for previous generations.

Again this had been optimized to reduce the current consumption as far as possible.

3. Battery: While current consumption has been minimized as far as possible within the circuitry of the phone, there had been an increase in current drain on the battery.

With users expecting the same lifetime between charging batteries as experienced on the previous generation phones, this had necessitated the use of new and improved battery technology. Lithium Ion (Li-ion) batteries started to be more widely used to address this issue.

These phones needed to remain small and relatively light while still retaining or even improving the overall life between charges.

Universal Subscriber Identity Module, USIM

4. Universal Subscriber Identity Module, USIM: The UE also contained a SIM card, although in the case of UMTS it was termed a USIM (Universal Subscriber Identity Module).

This was a more advanced version of the SIM card used in GSM and other systems, but embodied the same types of information. It contained the International Mobile Subscriber Identity number (IMSI) as well as the Mobile Station International ISDN Number (MSISDN).

Other information that the USIM held included the preferred language to enable the correct language information to be displayed, especially when roaming, and a list of preferred and prohibited Public Land Mobile Networks (PLMN).

The USIM also contained a short message storage area that allowed messages to stay with the user even when the phone was changed. Similarly "phone book" numbers and call information of the numbers of incoming and outgoing calls were stored.

Radio Network Subsystem

- This was the section of the 3G UMTS / WCDMA network that interfaced to both the UE and the core network - it handled the wireless communications elements of the network.
- The overall radio access network, i.e. collectively all the Radio Network Subsystem was known as the UTRAN or UMTS Radio Access Network.

An **RNC** in UMTS has a broad spectrum of tasks as listed in the following:

- **Call admission control:** It is very important for CDMA systems to keep the interference below a certain level. The RNC calculates the traffic within each cell and decides, if additional transmissions are acceptable or not.
- **Congestion control:** During packet-oriented data transmission, several sta- tions share the available radio resources. The RNC allocates bandwidth to each station in a cyclic fashion and must consider the QoS requirements.
- **Encryption/decryption:** The RNC encrypts all data arriving from the fixed network before transmission over the wireless link (and vice versa).

- **ATM switching and multiplexing, protocol conversion:** Typically, the connections between RNCs, node Bs, and the CN are based on ATM. An RNC has to switch the connections to multiplex different data streams. Several protocols have to be converted – this is explained later.
- **Radio resource control:** The RNC controls all radio resources of the cells connected to it via a node B. This task includes interference and load mea- surements. The priorities of different connections have to be obeyed.
- **Radio bearer setup and release:** An RNC has to set-up, maintain, and release a logical data connection to a UE (the so-called UMTS radio bearer).
- **Code allocation:** The CDMA codes used by a UE are selected by the RNC. These codes may vary during a transmission.

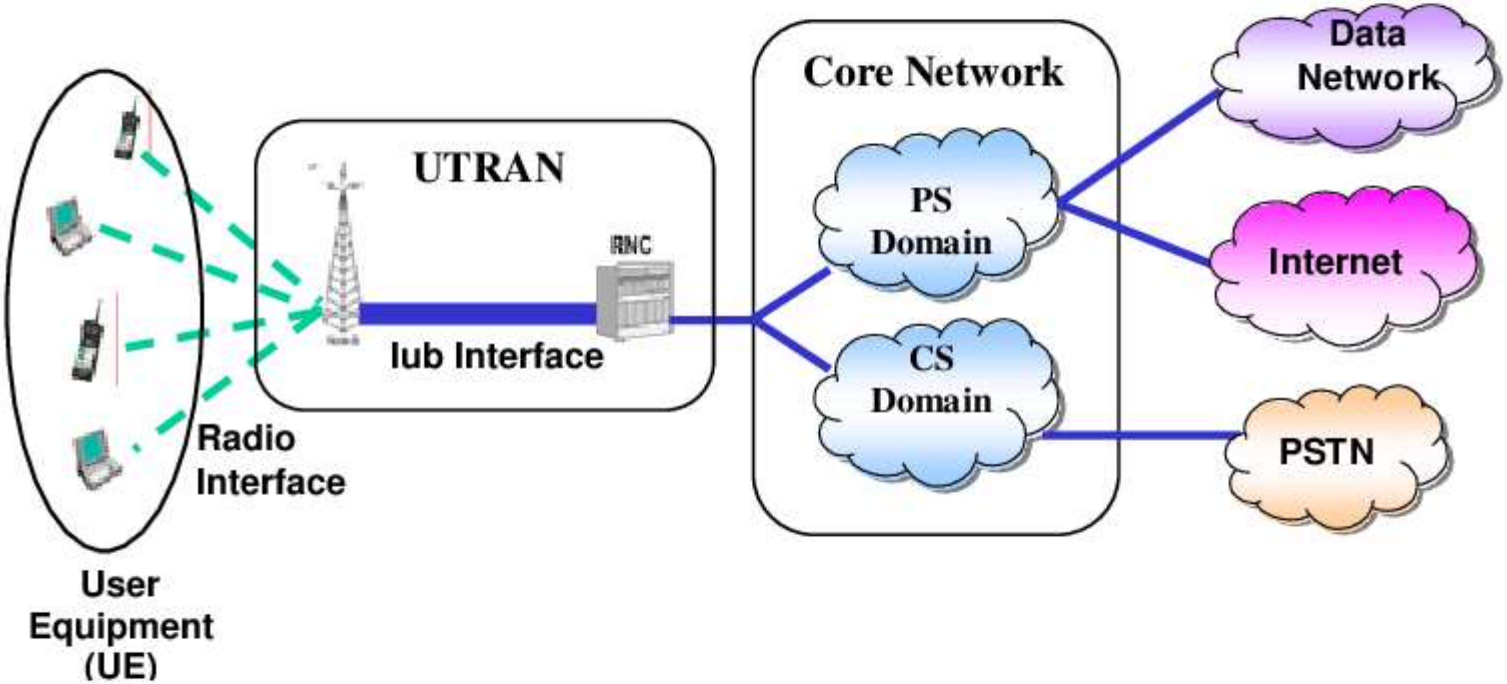
- **Power control:** The RNC only performs a relatively loose power control (the outer loop). This means that the RNC influences transmission power based on interference values from other cells or even other RNCs. But this is not the tight and fast power control performed 1,500 times per second. This is carried out by a node B. This outer loop of power control helps to mini- mize interference between neighbouring cells or controls the size of a cell.
- **Handover control and RNS relocation:** Depending on the signal strengths received by UEs and node Bs, an RNC can decide if another cell would be better suited for a certain connection. If the RNC decides for handover it informs the new cell and the UE as explained in subsection 4.4.6. If a UE moves further out of the range of one RNC, a new RNC responsible for the UE has to be chosen. This is called RNS relocation.
- **Management:** Finally, the network operator needs a lot of information regarding the current load, current traffic, error states etc. to manage its net- work. The RNC provides interfaces for this task, too.

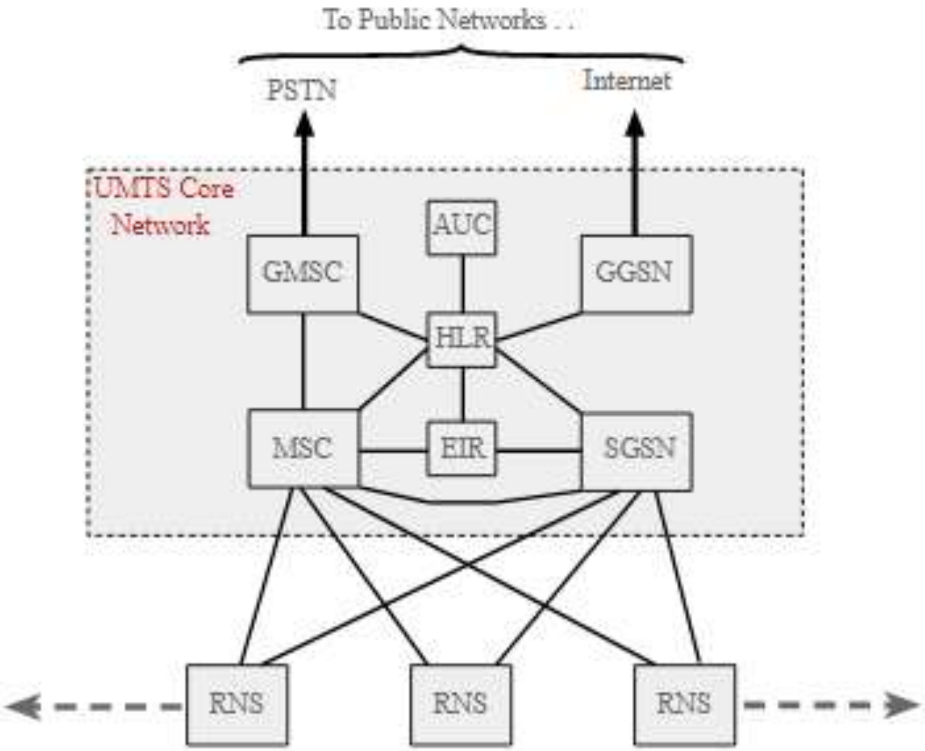
Core Network

The 3G UMTS core network architecture was a migration of that used for GSM with further elements overlaid to enable the additional functionality demanded by UMTS.

In view of the different ways in which data could be carried, the UMTS core network was split into two different areas:

- Circuit switched elements: These elements were primarily based on the GSM network entities and carry data in a circuit switched manner, i.e. a permanent channel for the duration of the call.
- Packet switched elements: These network entities were designed to carry packet data. This enabled much higher network usage as the capacity could be shared and data was carried as packets which were routed according to their destination


Core Network
Data Network
UTRAN
PS Domain
Internet
RNC
Iub Interface
CS Domain
PSTN
Radio Interface
User Equipment (UE)



UMTS Network Architecture Overview

Circuit switched elements

The circuit switched elements of the UMTS core network architecture included the following network entities:

- ***Mobile switching centre (MSC):*** This was essentially the same as that within GSM, and it managed the circuit switched calls under way.
- ***Gateway MSC (GMSC):*** This was effectively the interface to the external networks.

Packet switched elements

1. ***Serving GPRS Support Node (SGSN):*** As the name implies, this entity was first developed when GPRS was introduced, and its use has been carried over into the UMTS network architecture. The SGSN provided a number of functions within the UMTS network architecture.

   I. Mobility management When a UE attached to the Packet Switched domain of the UMTS Core Network, the SGSN generates MM information based on the mobile's current location.

   II. Session management: The SGSN managed the data sessions providing the required quality of service and it also managed what were termed the PDP (Packet data Protocol) contexts, i.e. the pipes over which the data was sent.

   III. Interaction with other areas of the network: The SGSN was able to manage its elements within the network only by communicating with other areas of the network, e.g. MSC and other circuit switched areas.

   IV. Billing: The SGSN was also responsible billing. It achieved this by monitoring the flow of user data across the GPRS network. CDRs (Call Detail Records) were generated by the SGSN before being transferred to the charging entities (Charging Gateway Function, CGF).

2. Gateway GPRS Support Node (GGSN): Like the SGSN, this entity was also first introduced into the GPRS network. The Gateway GPRS Support Node (GGSN) was the central element within the UMTS packet switched network. It handled inter-working between the UMTS packet switched network and external packet switched networks, and could be considered as a very sophisticated router. In operation, when the GGSN received data addressed to a specific user, it checked if the user was active and then forwarded the data to the SGSN serving the particular UE.

Shared elements

The shared elements of the 3G UMTS core network architecture included the following network entities:

i. ***Home location register (HLR):*** This database contained all the administrative information about each subscriber along with their last known location. In this way, the UMTS network was able to route calls to the relevant RNC / Node B. When a user switched on their UE, it registered with the network and from this it was possible to determine which Node B it communicated with so that incoming calls could be routed appropriately. Even when the UE was not active (but switched on) it re-registered periodically to ensure that the network (HLR) was aware of its latest position with their current or last known location on the network.

ii. ***Equipment identity register (EIR):*** The EIR was the entity that decided whether a given UE equipment could be allowed onto the network. Each UE equipment had a number known as the International Mobile Equipment Identity. This number, as mentioned above, was installed in the equipment and was checked by the network during registration.

iii. ***Authentication centre (AuC) :*** The AuC was a protected database that contained the secret key also contained in the user's USIM card

Handover

UMTS knows two basic classes of handovers:

1. Hard Hand over
2. Soft Hand over

Hard handover:

This handover type is already known from GSM and other TDMA/FDMA systems. Switching between different antennas or different systems is performed at a certain point in time. **UTRA TDD** can only use this type. Switching between TDD cells is done between the slots of different frames. **Inter frequency handover**, i.e., changing the carrier frequency, is a hard handover.

Receiving data at different frequencies at the same time requires a more complex receiver compared to receiving data from different sources at the same carrier frequency.

Typically, all **inter system handovers** are hard handovers in UMTS. This includes handovers to and from GSM or other IMT-2000 systems. A special type of handover is the handover to a satellite system (inter-segment handover), which is also a hard handover, as different frequencies are used. However, it is unclear what technology will be used for satellite links if it will ever come.

To enable a UE to listen into GSM or other frequency bands, UMTS specifies a **compressed mode** transmission for UTRA FDD. During this mode a UE stops all transmission. To avoid data loss, either the spreading factor can be lowered before and after the break in transmission (i.e., more data can be sent in shorter time) or less data is sent using different coding schemes.

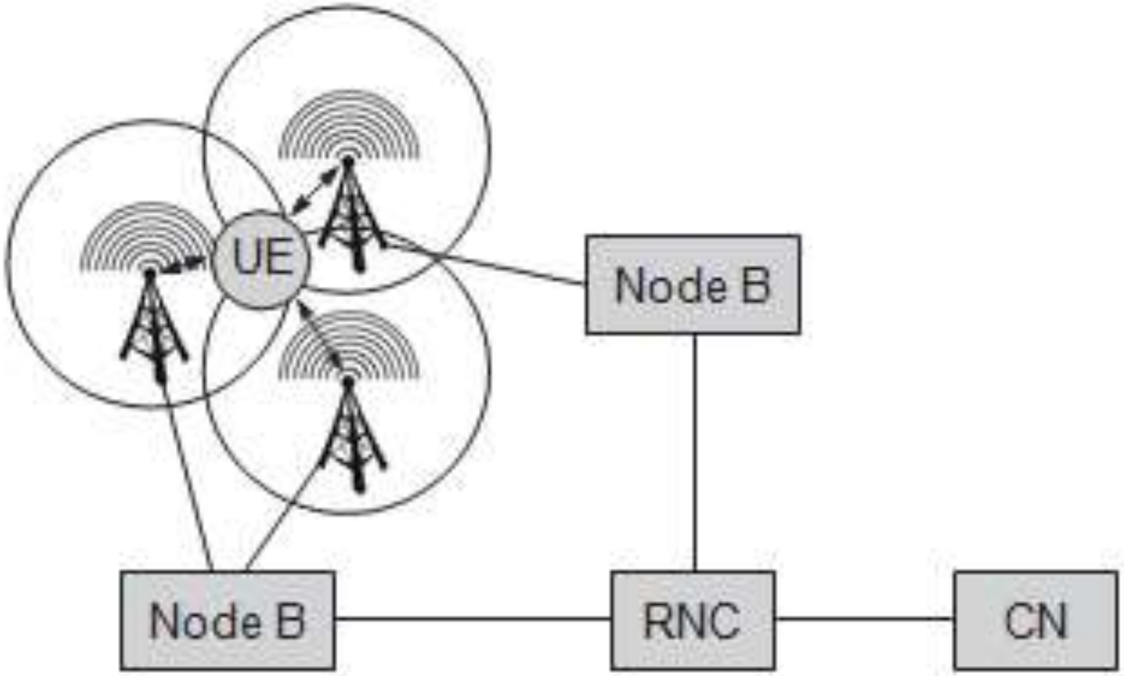
Soft handover:

This is the real new mechanism in UMTS compared to GSM and is only available in the FDD mode. Soft handovers are well known from traditional CDMA networks as they use **macro diversity**, a basic property of CDMA.

UE can receive signals from up to three different antennas, which may belong to different node Bs. Towards the UE the RNC splits the data stream and forwards it to the node Bs. The UE combines the received data again.

In the other direction, the UE simply sends its data which is then received by all node Bs involved. The RNC combines the data streams received from the node Bs.

The fact that a UE receives data from different antennas at the same time makes a handover soft. Moving from one cell to another is a smooth, not an abrupt process.


UE
Node B
Node B
RNC
CN

Mobile IP

- Need of mobile IP, IP packet delivery, Agent Discovery, Registration, Tunnelling and encapsulation, Route optimization, IP Handoff

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 the user’s sessions or connections being dropped.

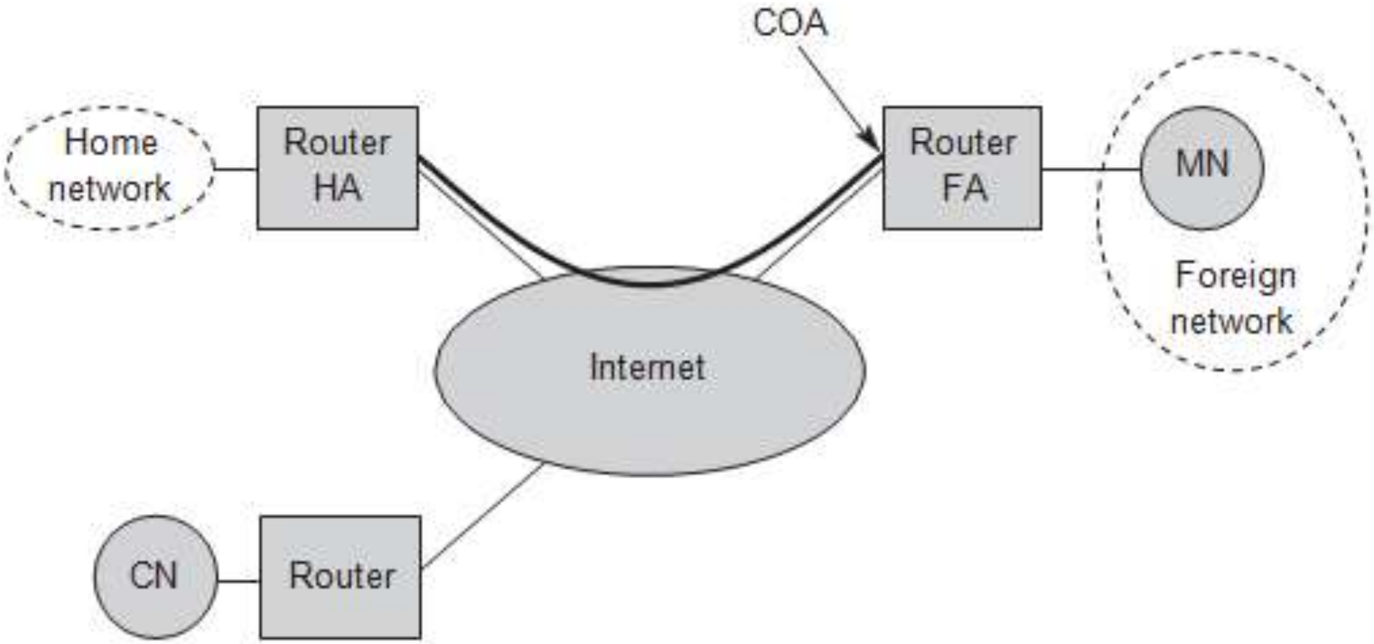
Entities and terminology

1. Mobile Node
2. Correspondent Node
3. Home network
4. Foreign Network
5. Foreign agent
6. Care of address
   1. Foreign Agent COA
   2. Co located COA
7. Home Agent address

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

Mobile Node

- ✓ 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. Mobile nodes are not necessarily small devices such as laptops with antennas or mobile phones; a router onboard an aircraft can be a powerful mobile node.


COA
Home network
Router HA
Router FA
MN
Foreign network
Internet
CN
Router

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 this example.

- ✓ **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 (defined below), 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. For mobile IP functioning, FAs are not necessarily needed. Typically, an 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, as explained later. To be more precise, the COA marks the tunnel endpoint, i.e., the address where packets exit the tunnel.

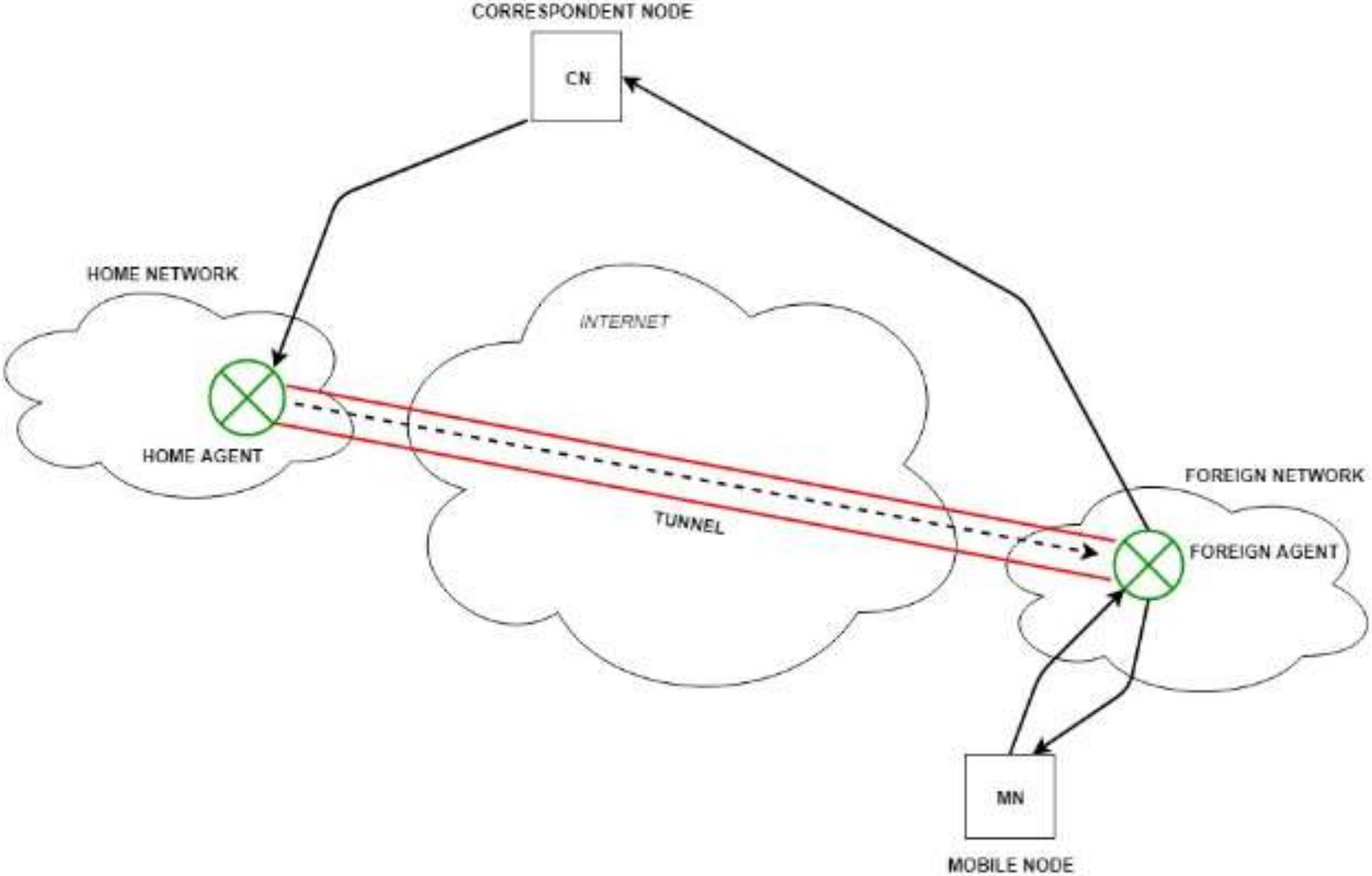
There are two different possibilities for the location of the COA:

- **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. One problem associated with this approach is the need for additional addresses if MNs request a COA. This is not always a good idea considering the scarcity of IPv4 addresses.

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.

  - 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.
  - 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.


CORRESPONDENT NODE
CN
HOME NETWORK
INTERNET
HOME AGENT
TUNNEL
FOREIGN NETWORK
FOREIGN AGENT
MN
MOBILE NODE

- The correspondent node sends the data to the mobile node. Data packets contain the correspondent node’s address (Source) and home address (Destination). Packets reach the home agent. But now mobile node is not in the home network, it has moved into the foreign network. The 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, the 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 another side of the tunnel, receives the data packets, decapsulates them, and sends them to the mobile node. The mobile node in response to the data packets received sends a reply in response to the foreign agent. The foreign agent directly sends the reply to the correspondent node.

Key Mechanisms in Mobile IP:

- **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.
- **Agent Registration:** Mobile node after discovering the foreign agent sends a registration request (RREQ) to the foreign agent. The foreign agent, in turn, sends the registration request to the home agent with the care-of-address. The home agent sends a registration reply (RREP) to the foreign agent. Then it forwards the registration reply to the mobile node and completes the process of registration.
- **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 the home agent receives a packet from the correspondent node, it encapsulates the packet with source address as home address and destination as care-of-address.

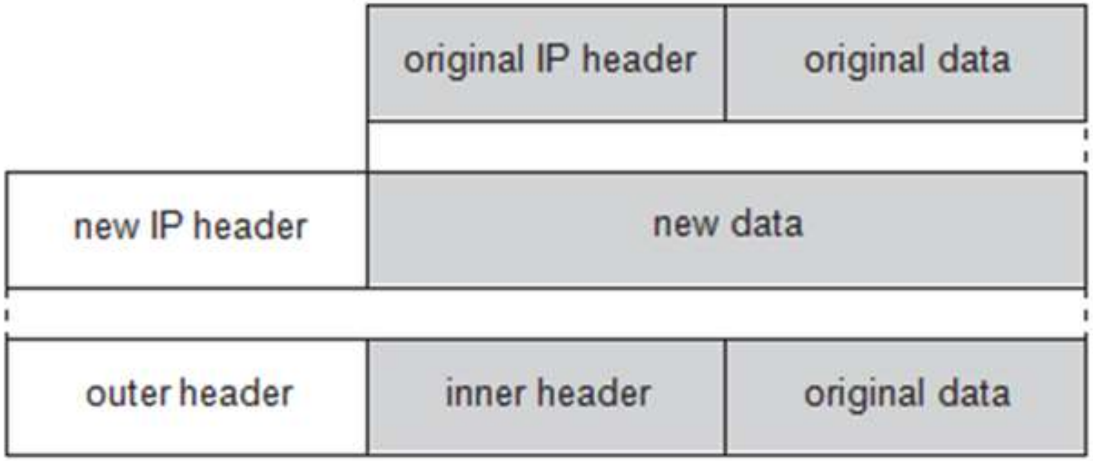
Encapsulation

Encapsulation is the mechanism of taking a packet consisting of packet header and data and putting it into the data part of a new packet.

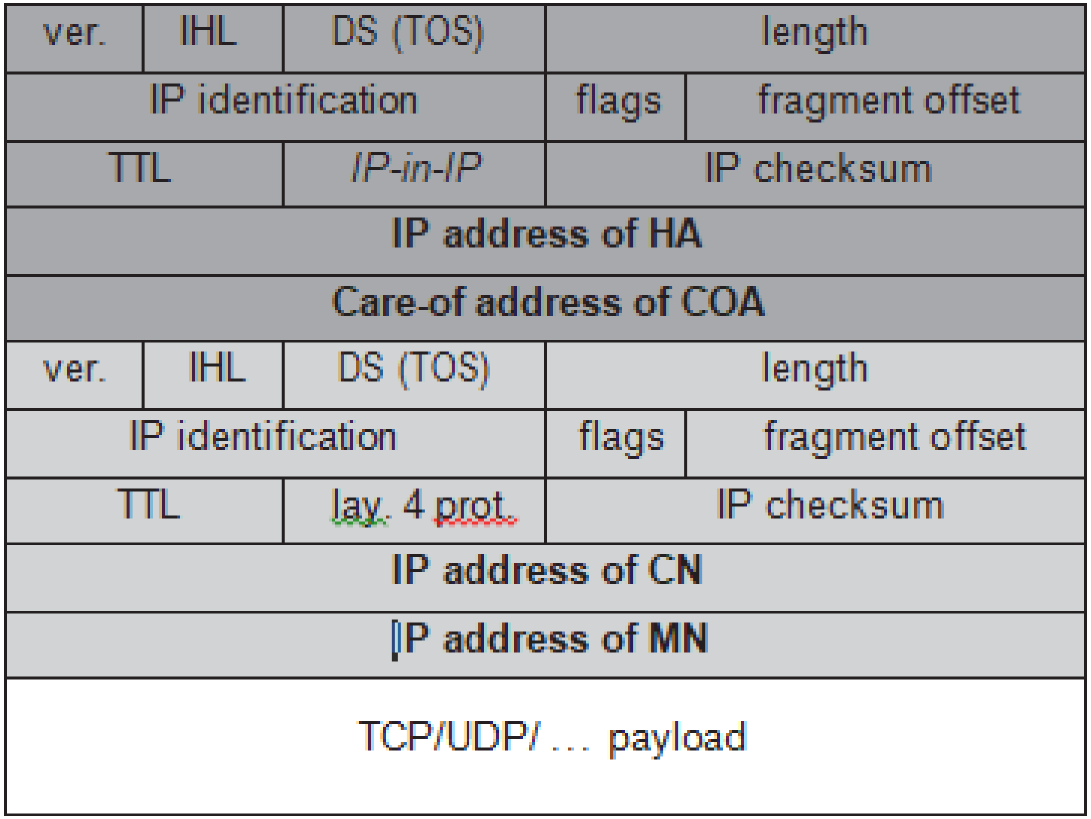
The reverse operation, taking a packet out of the data part of another packet, is called **decapsulation**.

A **tunnel** establishes a virtual pipe for data packets between a tunnel entry and a tunnel endpoint. Packets entering a tunnel are forwarded inside the tunnel and leave the tunnel unchanged. Tunneling (sending a packet through a tunnel) is achieved by using encapsulation.

The HA takes the original packet with the MN as destination, puts it into the data part of a new packet and sets the new IP header in such a way that the packet is routed to the COA. The new header is also called the **outer header.** There are different ways of performing the encapsulation needed for the tunnel between HA and COA.



1. IP-in-IP encapsulation



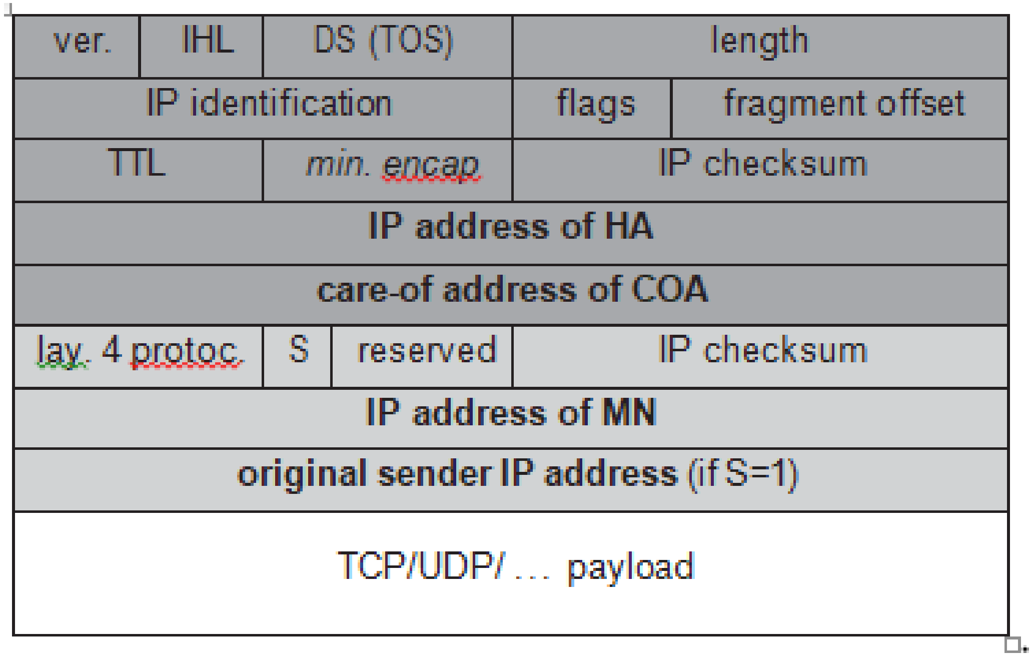
<table>
<tr><td>ver.</td><td>IHL</td><td>DS (TOS)</td><td colspan="2">length</td></tr>
<tr><td colspan="3">IP identification</td><td>flags</td><td>fragment offset</td></tr>
<tr><td colspan="2">TTL</td><td><i>IP-in-IP</i></td><td colspan="2">IP checksum</td></tr>
<tr><td colspan="5"><b>IP address of HA</b></td></tr>
<tr><td colspan="5"><b>Care-of address of COA</b></td></tr>
<tr><td>ver.</td><td>IHL</td><td>DS (TOS)</td><td colspan="2">length</td></tr>
<tr><td colspan="3">IP identification</td><td>flags</td><td>fragment offset</td></tr>
<tr><td colspan="2">TTL</td><td>lay. 4 prot.</td><td colspan="2">IP checksum</td></tr>
<tr><td colspan="5"><b>IP address of CN</b></td></tr>
<tr><td colspan="5"><b>IP address of MN</b></td></tr>
<tr><td colspan="5">TCP/UDP/ … payload</td></tr>
</table>

- The version field **ver** is 4 for IP version 4.
- The internet header length (**IHL**) denotes the length of the outer header in 32 bit words.
- **DS(TOS)** is just copied from the inner header, the **length** field covers the complete encapsulated packet.
- **TTL** must be high enough so the packet can reach the tunnel endpoint.
- **IP-in-IP**, is the type of the protocol used in the IP payload. This field is set to 4, the protocol type for IPv4 because again an IPv4 packet follows after this outer header.
- **IP address of the HA** is the tunnel entry as source address and
- **COA** is the tunnel exit point as destination address
- **IP checksum** is calculated **by the sender using a specific algorithm**. It is then stored in the header and sent as part of the data stream. The receiving side calculates the checksum on the data that is received using the same algorithm as the sender and compares its value to the checksum passed in the header

If no options follow the outer header, the inner header starts with the same fields as just explained. This header remains almost unchanged during encapsulation, thus showing the original sender CN and the receiver MN of the packet.

The only change is TTL which is decremented by 1. This means that the whole tunnel is considered a single hop from the original packet's point of view. This is a very important feature of tunneling as it allows the MN to behave as if it were attached to the home network. No matter how many real hops the packet has to take in the tunnel, it is just one (logical) hop away for the MN. Finally, the payload follows the two headers.

2. Minimal encapsulation

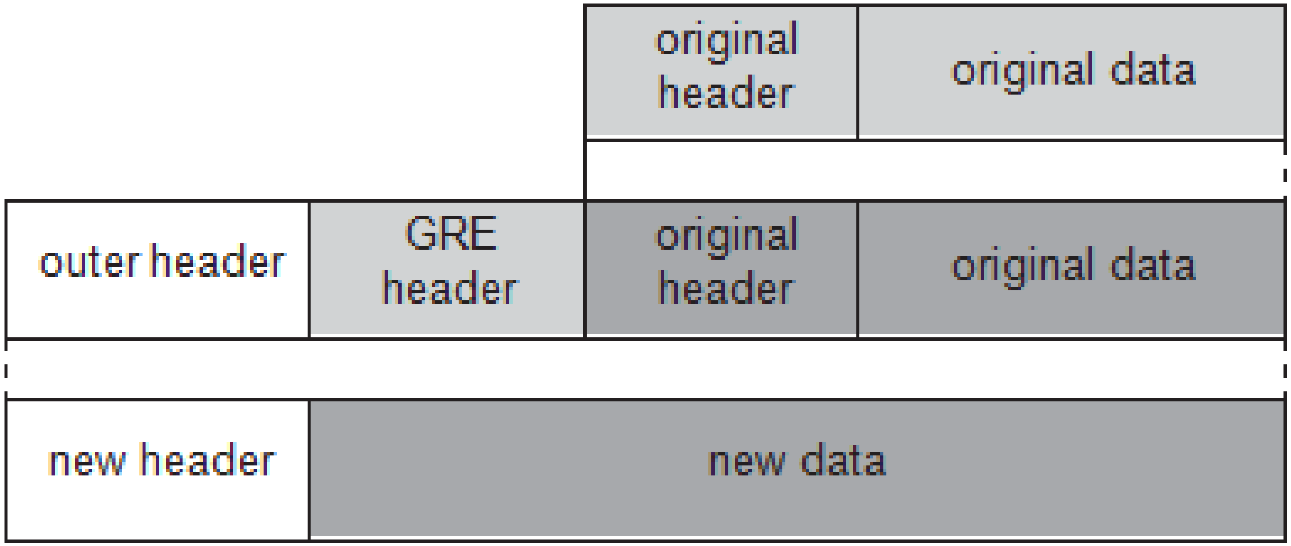


<table>
<tr><td>ver.</td><td>IHL</td><td colspan="2">DS (TOS)</td><td colspan="2">length</td></tr>
<tr><td colspan="4">IP identification</td><td>flags</td><td>fragment offset</td></tr>
<tr><td colspan="2">TTL</td><td colspan="2"><i>min. encap</i></td><td colspan="2">IP checksum</td></tr>
<tr><td colspan="6">IP address of HA</td></tr>
<tr><td colspan="6">care-of address of COA</td></tr>
<tr><td colspan="2">lay. 4 protoc.</td><td>S</td><td>reserved</td><td colspan="2">IP checksum</td></tr>
<tr><td colspan="6">IP address of MN</td></tr>
<tr><td colspan="6">original sender IP address (if S=1)</td></tr>
<tr><td colspan="6">TCP/UDP/ … payload</td></tr>
</table>

In IP-in-IP encapsulation, several fields are redundant. Therefore, minimal encapsulation can be used . The tunnel entry point and endpoint are specified.

The inner header is different for minimal encapsulation. The type of the following protocol and the address of the MN are needed. If the **S** bit is set, the original sender address of the CN is included as omitting the source is quite often not an option. No field for frag- mentation offset is left in the inner header and minimal encapsulation does not work with already fragmented packets.

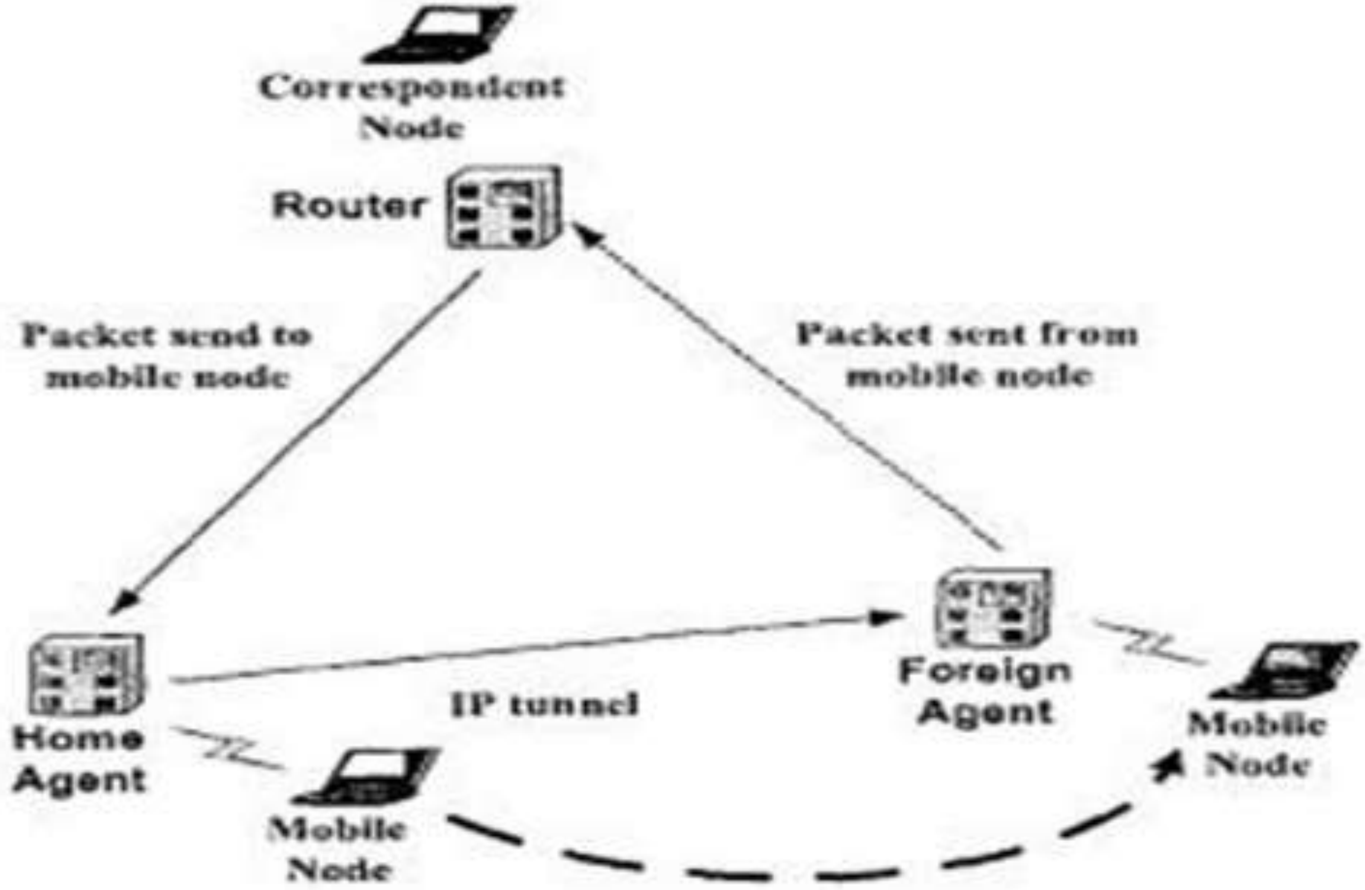
3. **Generic routing encapsulation**



Generic routing encapsulation (GRE) allows the encapsulation of packets of one protocol suite into the payload portion of a packet of another protocol suite. IP-in-IP encapsulation and minimal encapsulation work only for IP.

The packet of one protocol suite with the original packet header and data is taken and a new GRE header is prepended. Together this forms the new data part of the new packet. Finally, the header of the second protocol suite is put in front.

Triangular routing



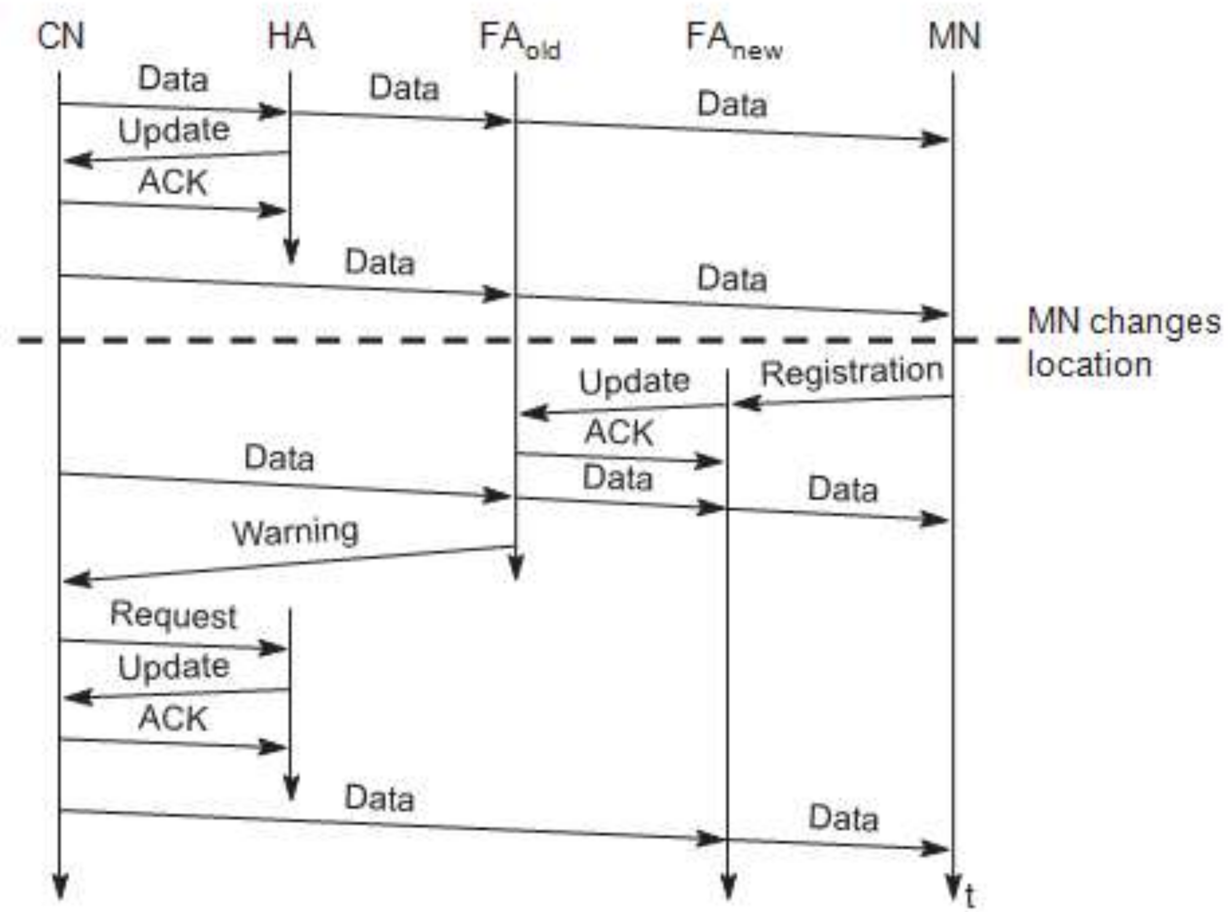
Optimization

- An inefficient behavior of a non- optimized mobile IP is called **triangular routing**. The triangle is made of the three segments, CN to HA, HA to COA/MN, and MN back to CN.
- With the basic mobile IP protocol all packets to the MN have to go through the HA. This can cause unnecessary overheads for the network between CN and HA, but also between HA and COA, depending on the current location of the MN. The latency can increase dramatically.
- One way to optimize the route is to inform the CN of the current location of the MN. The CN can learn the location by caching it in a **binding cache** which is a part of the local routing table for the CN. The appropriate entity to inform the CN of the location is the HA.

The optimized mobile IP protocol needs four additional messages.

1. **Binding request**: Any node that wants to know the current location of an MN can send a binding request to the HA. The HA can check if the MN has allowed dissemination of its current location. If the HA is allowed to reveal the location it sends back a binding update.
2. **Binding update:** This message sent by the HA to CNs reveals the current location of an MN. The message contains the fixed IP address of the MN and the COA. The binding update can request an acknowledgement.
3. **Binding acknowledgement**: If requested, a node returns this acknowledgement after receiving a binding update message.
4. **Binding warning:** If a node decapsulates a packet for an MN, but it is not the current FA for this MN, this node sends a binding warning. The warning contains MN’s home address and a target node address, i.e., the address of the node that has tried to send the packet to this MN. The recipient of the warning then knows that the target node could benefit from obtaining a fresh binding for the MN. The recipient can be the HA, so the HA should now send a binding update to the node that obviously has a wrong COA for the MN.

Change of the foreign agent with an optimized mobile IP



The route optimization adds a conceptual data structure, the binding cache, to the correspondent node. The binding cache contains bindings for the mobile node’s home address and its current care-of-address. Every time the home agent receives an 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. Mobile IP is provided by the network providers.

Unit IV

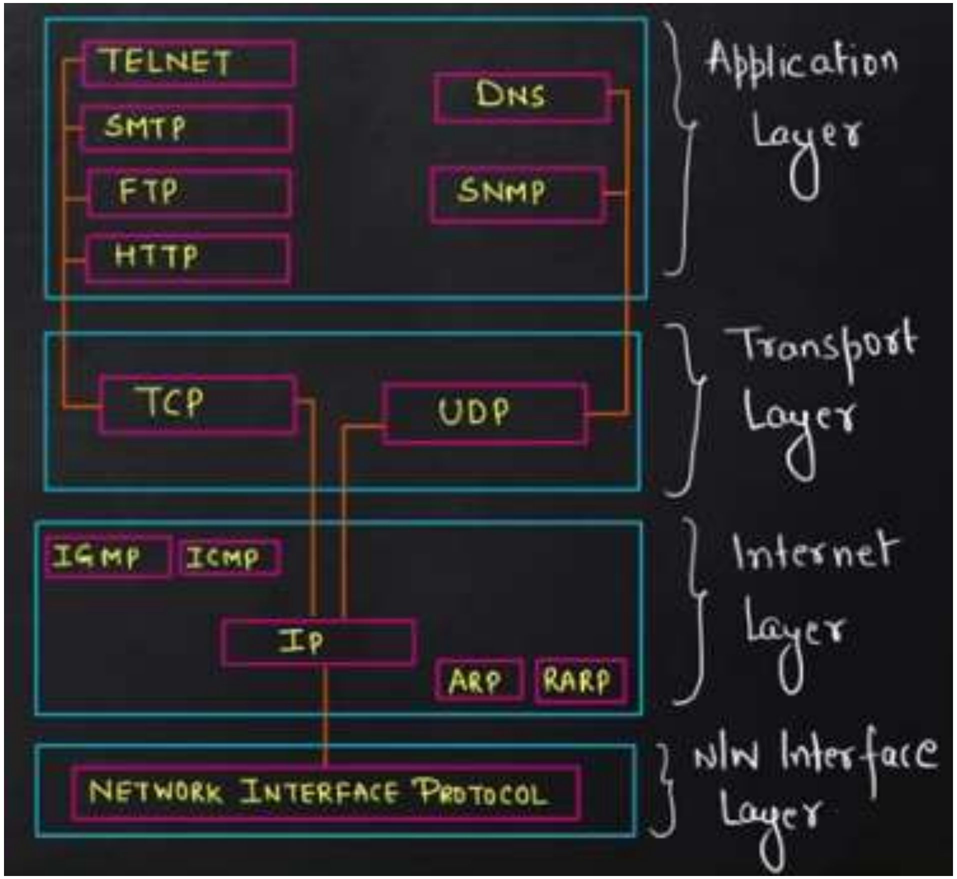
Overview of Traditional TCP and implications of mobility control. Improvement of TCP: Indirect TCP, Snoop TCP, Mobile TCP, Fast Retransmit/fast recovery, Time-out freezing, Selective retransmission, Transaction-oriented TCP.

Overview of Traditional TCP

- Supporting mobility only on lower layers up to the network layer is not enough to provide mobility support for applications. Most applications rely on a transport layer, such as TCP (transmission control protocol) or UDP (user datagram protocol) in the case of the internet. Two functions of the transport layer in the internet are

I. check summing over user data

II. multiplexing/demultiplexing of data from/to applications.


TELNET
SMTP
FTP
HTTP
DNS
SNMP
Application Layer
TCP
UDP
Transport Layer
IGMP
ICMP
IP
ARP
RARP
Internet layer
NETWORK INTERFACE PROTOCOL
N/W Interface Layer

difference between UDP and TCP

- TCP offers connections between two applications. Within a connection TCP can give certain guarantees, such as in-order delivery or reliable data transmission using retransmission techniques. TCP has built-in mechanisms to behave in a ‘network friendly’ manner. If, for example, TCP encounters packet loss, it assumes network internal congestion and slows down the transmission rate.
- UDP requires that applications handle reliability, in-order delivery etc. UDP does not behave in a network friendly manner, i.e., does not pull back in case of congestion and continues to send packets into an already congested network.

Traditional TCP

- Congestion control
- Slow start
- Fast retransmit/fast recovery
- Implications on mobility

Congestion control

- A transport layer protocol such as TCP has been designed for fixed networks with fixed end-systems. Data transmission takes place using network adapters, fiber optics, copper wires, special hardware for routers etc. The probable reason for a packet loss in a fixed network is a temporary overload some point in the transmission path, i.e., a state of congestion at a node.
- 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. Although it is not guaranteed that all packets of the TCP connection take the same way through the network, this assumption holds for most of the packets. To mitigate congestion, TCP slows down the transmission rate dramatically.

Slow start

- 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). After arrival of the two corresponding acknowledgements, the sender again adds 2 to the congestion window, one for each of the acknowledgements. Now the congestion window equals 4. This scheme doubles the congestion window every time the acknowledgements come back, which takes one round trip time. This is called the exponential growth of the congestion window in the slow start mechanism.
- It is too dangerous to double the congestion window each time because the steps might become too large. 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

- In TCP, a receiver sends acknowledgements only if it receives any packets from the sender. Receiving acknowledgements from a receiver also shows that the receiver continuously receives something from the sender. The gap in the packet stream is not due to severe 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.
- 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.

Implications on mobility

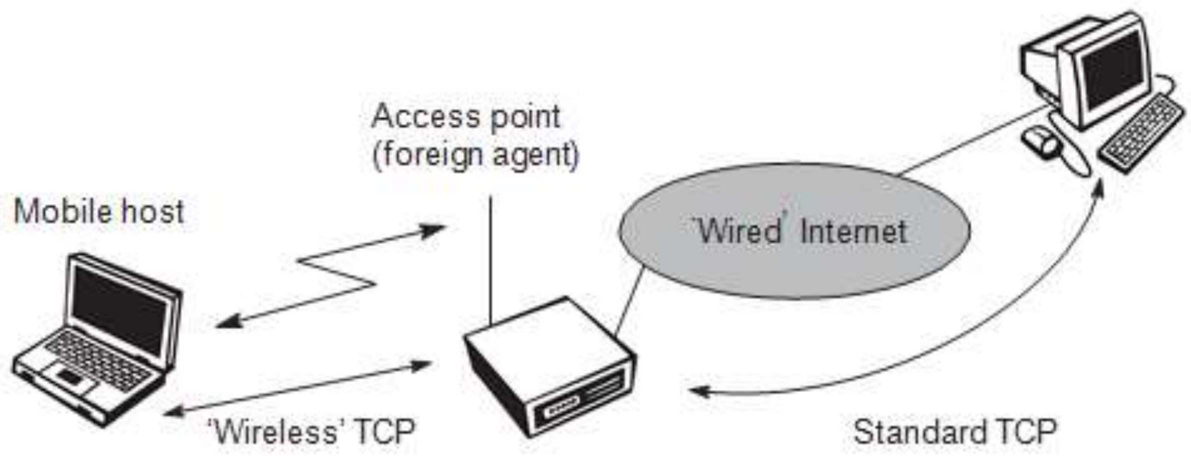
- Mobility itself can cause packet loss. There are many situations where a soft handover from one access point to another is not possible for a mobile end- system. For example, when using mobile IP, there could still be some packets in transit to the old foreign agent while the mobile node moves to the new foreign agent. The old foreign agent may not be able to forward those packets to the new foreign agent or even buffer the packets if disconnection of the mobile node takes too long. This packet loss has nothing to do with wireless access but is caused by the problems of rerouting traffic.
- The TCP mechanism detecting missing acknowledgements via time-outs and concluding packet loss due to congestion cannot distinguish between the different causes. This is a fundamental design problem in TCP: An error control mechanism (missing acknowledgement due to a transmission error) is misused for congestion control (missing acknowledgement due to network overload). In both cases packets are lost (either due to invalid checksums or to dropping in routers). However, the reasons are completely different. TCP cannot distinguish between these two different reasons.

Classical TCP improvements

- Indirect TCP
- Snooping TCP
- Mobile TCP
- Fast retransmit/fast recovery
- Transmission/time-out freezing
- Selective retransmission
- Transaction-oriented TCP

Indirect TCP

Indirect TCP segments a TCP connection into two parts



- 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.

- The correspondent host in the fixed network does not notice the wireless link or the segmentation of the connection. The foreign agent acts as a proxy and relays all data in both directions. If the correspondent host sends a packet, the foreign agent acknowledges this packet and tries to forward the packet to the mobile host. If the mobile host receives the packet, it acknowledges the packet. However, this acknowledgement is only used by the foreign agent. If a packet is lost on the wireless link due to a transmission error, the correspondent host would not notice this. In this case, the foreign agent tries to retransmit this packet locally to maintain reliable data transport.
- Similarly, if the mobile host sends a packet, the foreign agent acknowledges this packet and tries to forward it to the correspondent host. 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.

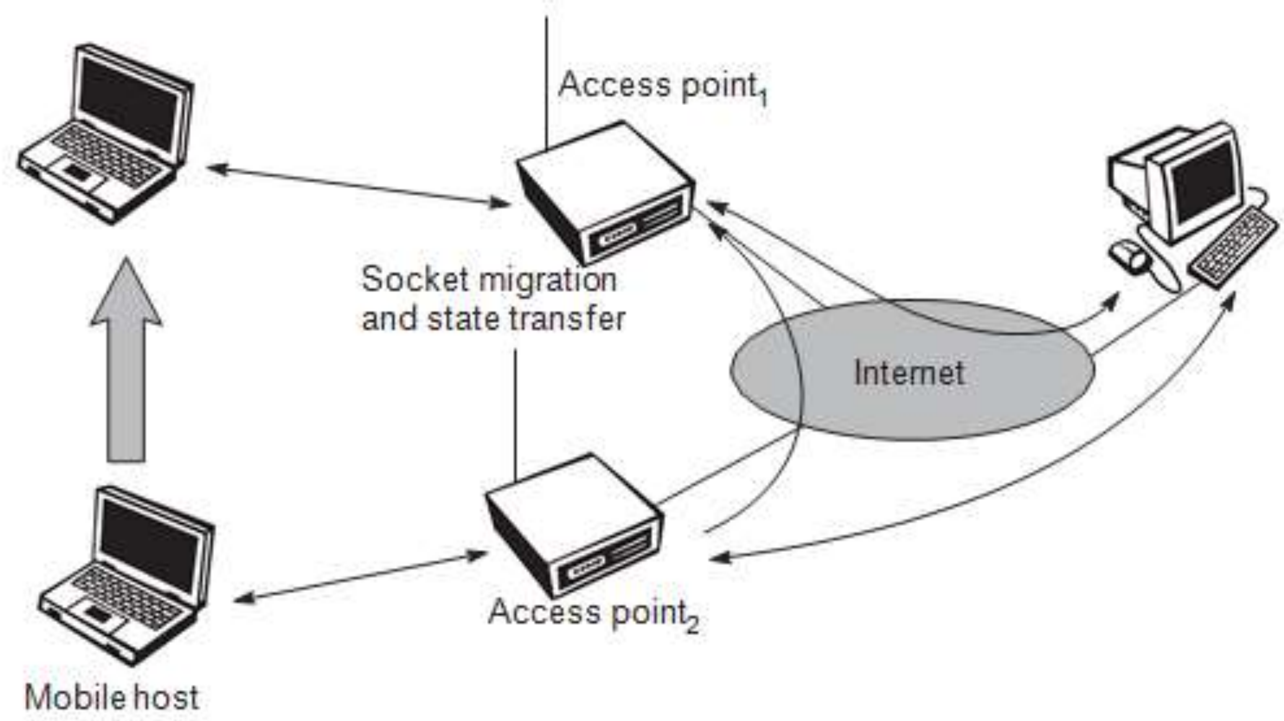
Advantages of I-TCP

- ●I-TCP does not require any changes in the TCP protocol as used by the hosts in the fixed network or other hosts in a wireless network that do not use this optimization.
- ●Due to the strict partitioning into two connections, transmission errors on the wireless link, i.e., lost packets, cannot propagate into the fixed network.
- ● new mechanisms are needed to improve TCP only between the mobile host and the foreign agent.
- ●An optimized TCP could use precise time-outs to guarantee
- retransmission as fast as possible. Even standard TCP could benefit from the short round trip time, so recovering faster from packet loss.
- ●Partitioning into two connections also allows the use of a different transport layer protocol between the foreign agent and the mobile host or the use of compressed headers etc. The foreign agent can now act as a gateway to translate between the different protocols.

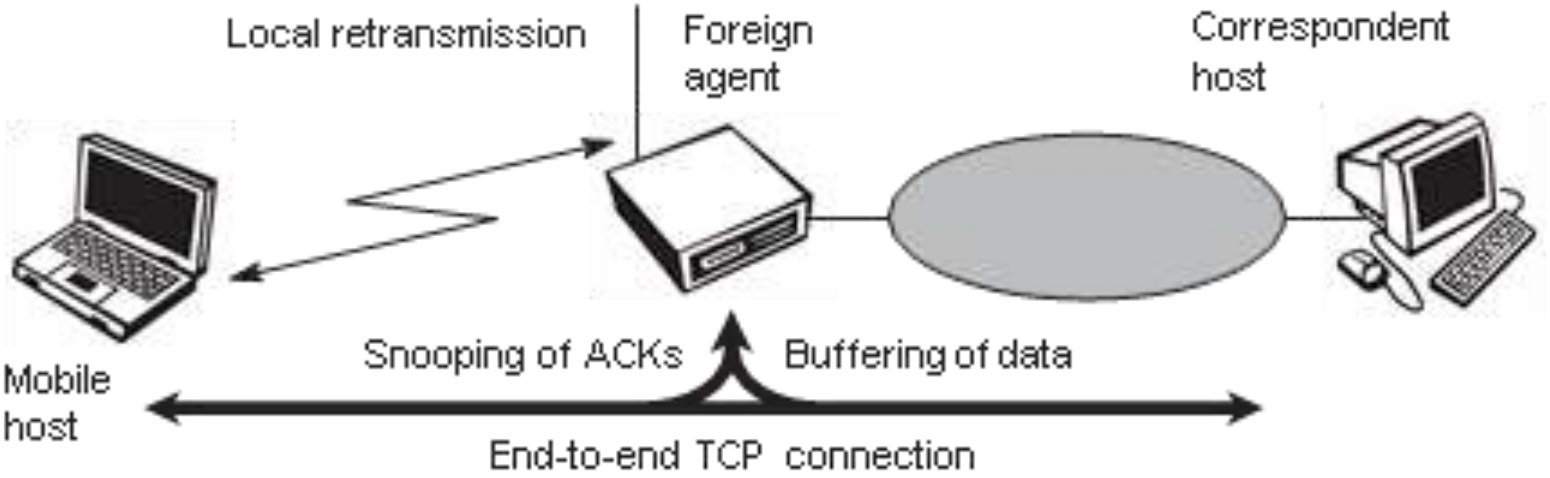
Disadvantages

- The correspondent node does not know anything about the partitioning, so a **crashing** access node may also crash applications running on the correspondent node assuming reliable end-to-end delivery.
- Increased handover **latency**: All packets sent by the correspondent host are buffered by the foreign agent besides forwarding them to the mobile host. The foreign agent removes a packet from the buffer as soon as the appropriate acknowledgement arrives. If the mobile host now performs a handover to another foreign agent, it takes a while before the old foreign agent can forward the buffered data to the new foreign agent.
- The foreign agent must be a trusted entity because the TCP connections end at this point. If users apply end-to-end encryption, the foreign agent has to be integrated into all **security** mechanisms.

Socket and state migration after handover of a mobile host



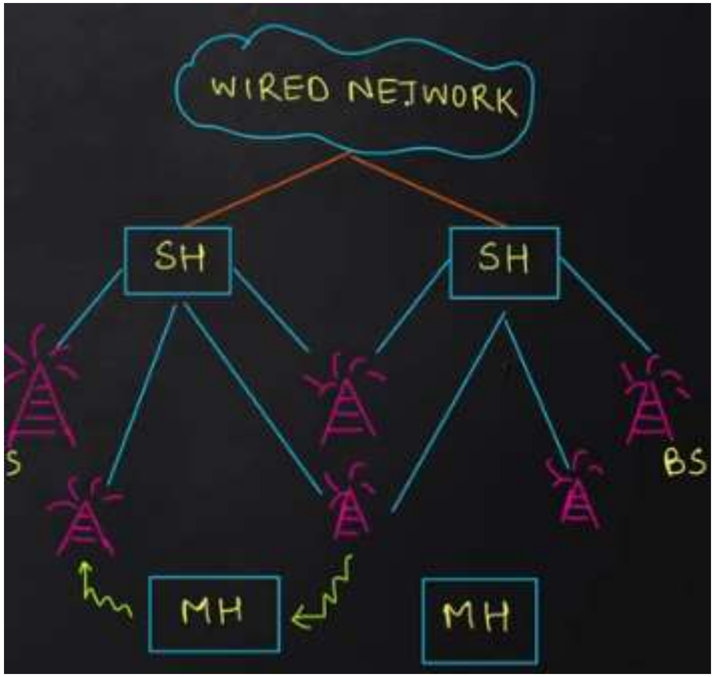
Snooping TCP



- In this approach, the foreign agent buffers all packets with destination mobile host and additionally ‘snoops’ the packet flow in both directions to recognize acknowledgements. The reason for buffering packets toward the mobile node is to enable the foreign agent to per- form a local retransmission in case of packet loss on the wireless link. The foreign agent buffers every packet until it receives an acknowledgement from the mobile host. If the foreign agent 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 foreign agent retransmits the packet directly from the buffer, performing a much faster retransmission compared to the correspondent host. The time out for acknowledgements can be much shorter, because it reflects only the delay of one hop plus processing time.

Mobile TCP

- The M-TCP (mobile TCP)1 approach has the same goals as I-TCP and snooping TCP: to prevent the sender window from shrinking if bit errors or dis- connection 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 supervisory host is responsible for exchanging data between both parts similar to the proxy in I- TCP. The M-TCP approach assumes a relatively low bit error rate on the wireless link. Therefore, it does not perform caching/retransmission of data via the SH. If a packet is lost on the wireless link, it has to be retransmitted by the original sender. This maintains the TCP end-to-end semantics.


WIRED NETWORK
SH
SH
S
BS
MH
MH

- 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.

Fast retransmit/fast recovery

- As soon as the mobile host registers at a new foreign agent using mobile IP, it starts sending duplicated acknowledgements to correspondent hosts. The proposal is to send three dupli- cates. This forces the corresponding host to go into fast retransmit mode and not to start slow start, i.e., the correspondent host continues to send with the same rate it did before the mobile host moved to another foreign agent.
- As the mobile host may also go into slow start after moving to a new for- eign agent, this approach additionally puts the mobile host into fast retransmit. The mobile host retransmits all unacknowledged packets using the current con- gestion window size without going into slow start.

- The advantage of this approach is its simplicity. Only minor changes in the mobile host’s software already result in a performance increase. No foreign agent or correspondent host has to be changed.
- The main disadvantage of this scheme is the insufficient isolation of packet losses. Forcing fast retransmission increases the efficiency, but retransmitted packets still have to cross the whole network between correspondent host and mobile host. If the handover from one foreign agent to another takes a longer time, the correspondent host will have already started retransmission. The approach focuses on loss due to handover: packet loss due to problems on the wireless link is not considered. This approach requires more cooperation between the mobile IP and TCP layer making it harder to change one without influencing the other.

Transmission/time-out freezing

Quite often, the MAC layer has already noticed connection problems, before the connection is actually interrupted from a TCP point of view. Additionally, the MAC layer knows the real reason for the interruption and does not assume congestion, as TCP would. 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.

Selective retransmission

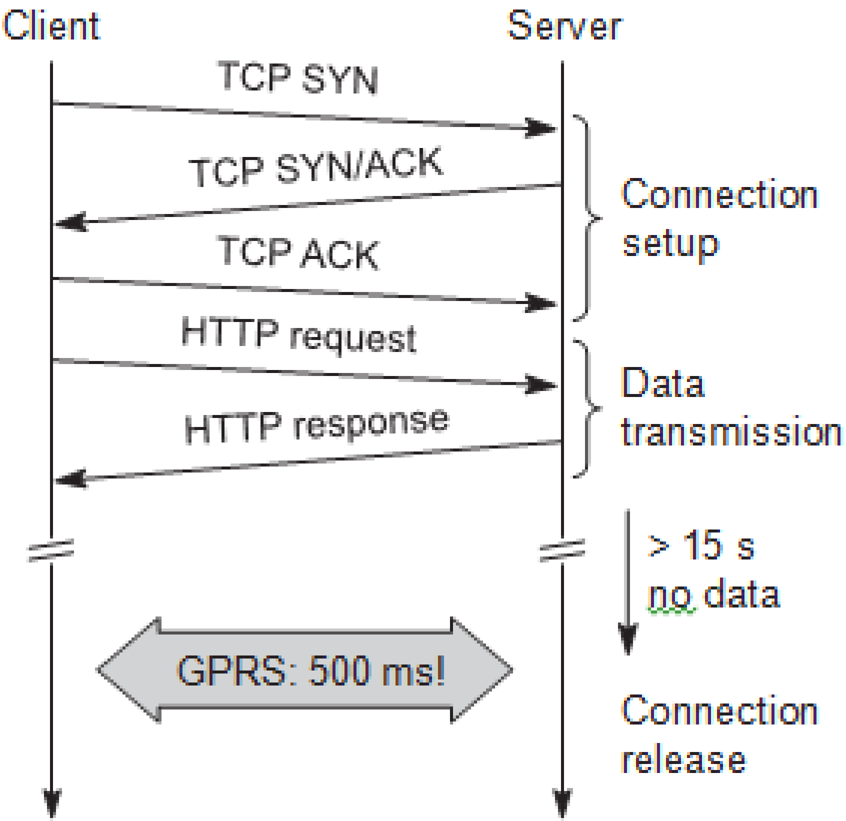
A very useful extension of TCP is the use of selective retransmission. TCP acknowledgements are cumulative, i.e., they acknowledge in-order receipt of packets up to a certain packet. If a single packet is lost, the sender has to retrans- mit everything starting from the lost packet (go-back-n retransmission). This obviously wastes bandwidth, not just in the case of a mobile network, but for any network

- 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.
- Advantage : a sender retransmits only the lost packets. This lowers bandwidth requirements and is extremely helpful in slow wireless links. The gain in efficiency is not restricted to wireless links and mobile environments. Using selective retransmission is also beneficial in all other networks.
- Disadvantage :- more complex software on the receiver side, because now more buffer is necessary to re sequence data and to wait for gaps to be filled. But while memory sizes and CPU performance permanently increase, the bandwidth of the air interface remains almost the same.

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. If the application requires reliable transport of the packets, it may use TCP (many applications of this kind use UDP and solve reliability on a higher, application-oriented layer).
- Using TCP now requires several packets over the wireless link. 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.

TCP connection setup overhead



- 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.

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, security problems
Snooping TCP	Snoops data and acknowledgements, local retransmission	Transparent for end-to-end connection, MAC integration possible	Insufficient isolation of wireless link, security problems
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, security problems
Fast retransmit/ fast recovery	Avoids slow-start after roaming	Simple and efficient	Mixed layers, not transparent
Transmission/ time-out freezing	Freezes TCP state at disconnection, resumes after reconnection	Independent of content, works for longer interruptions	Changes in TCP required, MAC dependent
Selective retransmission	Retransmits only lost data	Very efficient	Slightly more complex receiver software, more buffer space needed
Transaction-oriented TCP	Combines connection setup/release and data transmission	Efficient for certain applications	Changes in TCP required, not transparent, security problems