WIRELESS SENSOR NETWORKS

UNIT 1

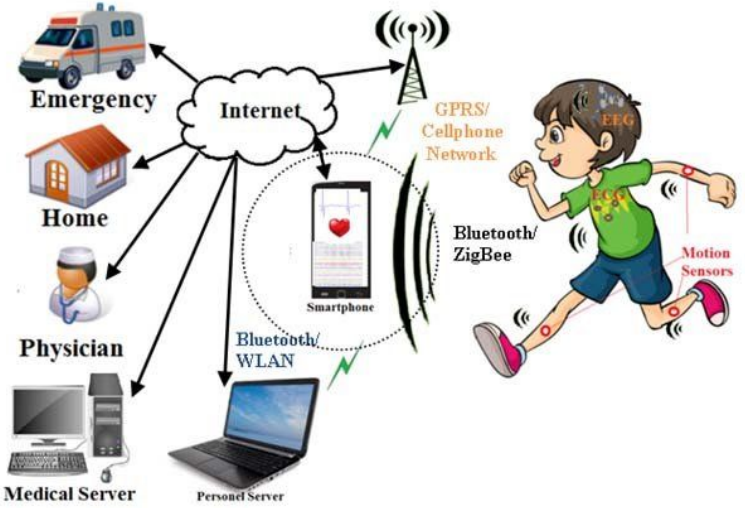
Asst. Prof. Jesica D’cruz

Introduction to Sensor Networks

- A wireless sensor network (WSN) can be defined as a network of small embedded devices, called sensors, which communicate wirelessly following an ad hoc configuration.
- These networks consist of individual nodes that are able to interact with their environment by sensing or controlling physical parameters; these nodes have to collaborate to fulfill their tasks as, usually, a single node is incapable of doing so; and they use wireless communication to enable this collaboration.
- Despite the fact that these networks also often include actuators, the term wireless sensor network has become the commonly accepted name.

Constraints & Challenges

- There is no single set of requirements that clearly classifies all WSNs, and there is also not a single technical solution that encompasses the entire design space
- For example, in many WSN applications, individual nodes in the network cannot easily be connected to a wired power supply but rather have to rely on onboard batteries.
- In such an application, the energy efficiency of any proposed solution is hence a very important figure of merit as a long operation time is usually desirable.



Constraints & Challenges

- In other applications, power supply might not be an issue and hence other metrics, for example, the accuracy of the delivered results, can become more important.
- Also, the acceptable size and costs of an individual node can be relevant in many applications.

Advantages of Wireless Sensor Networks

1. It is scalable and hence can accommodate any new nodes or devices at any time.
2. It is flexible and hence open to physical partitions.
3. All the WSN nodes can be accessed through centralized monitoring system.
4. As it is wireless in nature, it does not require wires or cables.
5. Wireless can be applied on large scale and in various domains such as mines, healthcare, surveillance, agriculture etc.
6. It uses different security algorithms as per underlying wireless technologies and hence provide reliable network for consumers or users.

Disadvantages of Wireless Sensor Networks

1. As it is wireless in nature, it is prone to hacking by hackers.
2. It cannot be used for high-speed communication as it is designed for low-speed applications.
3. It is expensive to build such network and hence cannot be affordable by all.
4. There are various challenges to be considered in WSN such as energy efficiency, limited bandwidth, node costs, deployment model, Software/hardware design constraints and so on.
5. In star topology based WSN, failure of central node leads to whole network shutdown.

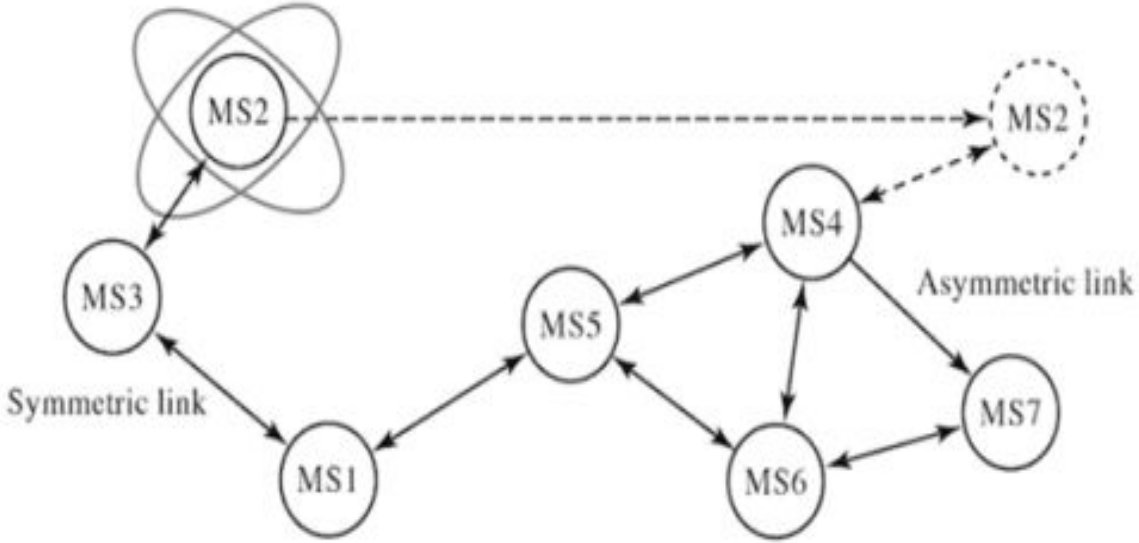
Applications of Sensor Networks

- Forest Fire Detection
- Industrial Control and Monitoring
- Home Applications
- Military and Security Applications
- Asset Tracking
- Heath Monitoring

Mobile Adhoc NETworks (MANETs) & WSN

- A MANET can be defined as an autonomous system of nodes or MSs(also serving as routers) that come together to form a network as needed connected by wireless links
- These networks are without any support from any existing internet infrastructure or any other kind of fixed stations.
- MANETs are basically peer-to-peer, multi-hop wireless networks in which information packets are transmitted in a store and forward manner from a source to an arbitrary destination, via intermediate nodes.

- Mobile nodes are free to move randomly. Network topology changes frequently.
- The resulting change in the network topology known at the local level must be passed on to other nodes so that old topology information can be updated.



Characteristics of MANET

1. **Dynamic topologies**

➔ Nodes are free to move arbitrarily; thus the network topology may be changed randomly and unpredictably and primarily consists of bidirectional links.

➔ In some cases where the transmission power of two nodes is different, a unidirectional link may exist.

2. **Bandwidth-constrained and variable capacity links**

➔ Wireless links continue to have significantly lower capacity than infrastructure networks.

Characteristics of MANET

3. Energy-constrained operation:

- ➔ Some or all of the MSs in a MANET may rely on batteries or other exhaustible means for their energy.
- ➔ For these nodes or devices, the most important system design optimization criteria may be energy conservation.

4. Limited physical security:

- ➔ MANETs are generally more prone to physical security threats than wire line networks.
- ➔ The increased possibility of eavesdropping, spoofing, and denial of services (DoS) attacks should be considered carefully.
- ➔ To reduce security threats, many existing link security techniques are often applied within wireless networks.

Let's look at the similarities between MANET and WSN

1. Both are infrastructure-less, distributed wireless networks
2. Routing Techniques are more or less the same
3. Both are Ad-hoc networks
4. Topology can change over a period
5. Nodes can be operated on a battery

What makes them different?

1. The data rate of MANETs is more than WSN
2. The number of nodes in the WSN is more than MANETs
3. Mobility is very high in MANETs(since nodes are less) than WSN
4. Sensor nodes of WSN are generally static and cooperate together to transfer the sensed data
5. Sensor nodes usually consume less energy than MANET's nodes
6. MANETs are usually close to civilization
7. Public-key cryptography is used in MANETs whereas symmetric key cryptography used in WSNs for security purposes

Enabling technologies for WSN

Building such wireless sensor networks has only become possible with some fundamental advances in enabling technologies.

1) **Miniaturization of hardware:** Reduced chip size and improved energy efficiency is accompanied by reduced cost.

2) **Network:** processing and communication

3) **Sensing:** actual sensing equipment

Sensor Node Hardware and Network Architecture

Key Terms

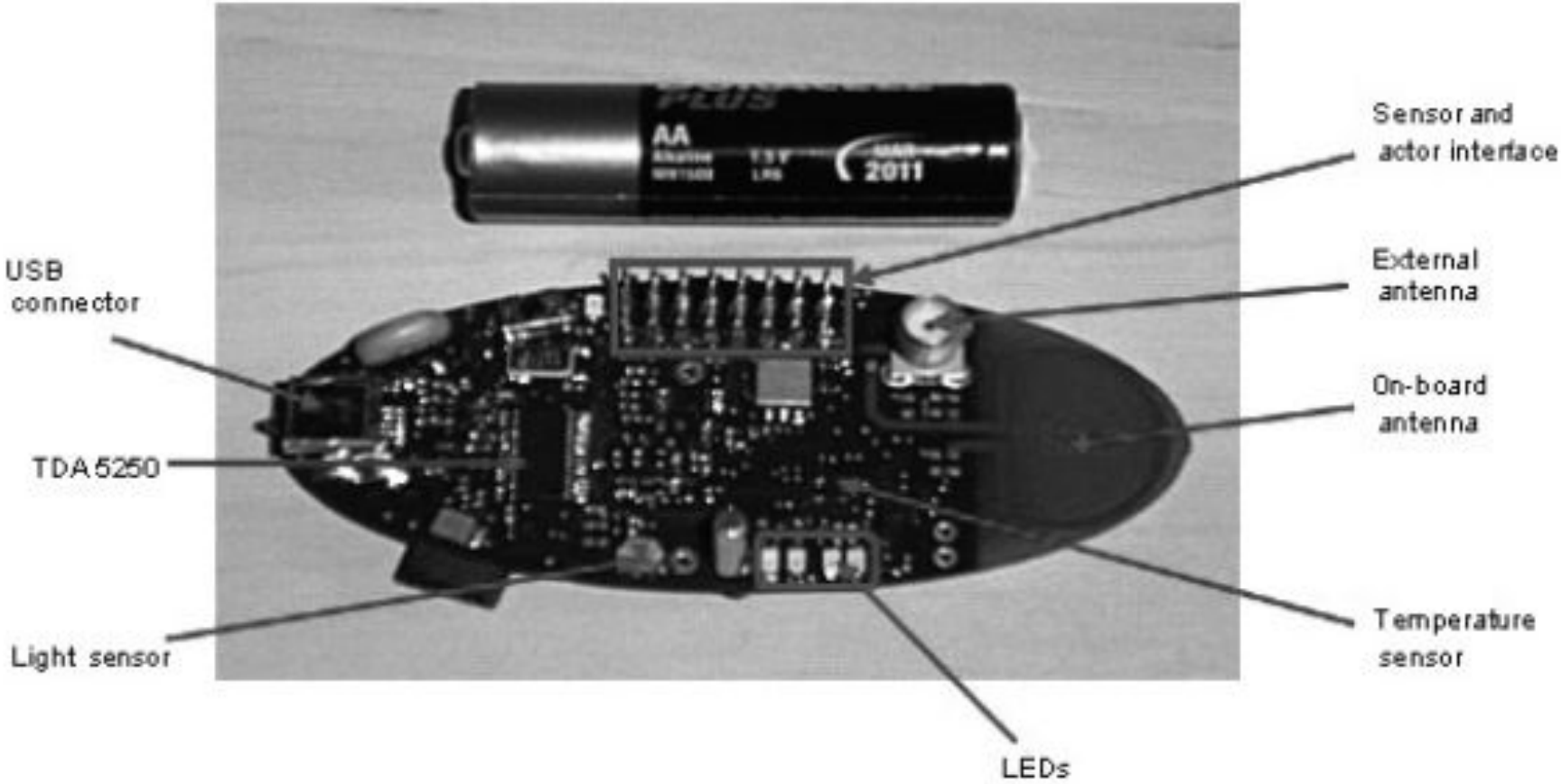
Sensor

A transducer that converts a physical phenomenon such as heat, light, sound, or motion into electrical or other signals that may be further operated by other apparatus.

Sensor node

A basic unit in a sensor network, with on-board sensors, processor, memory, wireless modem, and power supply. It is often abbreviated as node. When a node has only a single sensor on board, the node is sometimes referred as a sensor.

Sensor node


Sensor and actor interface
External antenna
On-board antenna
Temperature sensor
LEDs
Light sensor
TDA5250
USB connector

Hardware components & design constraints

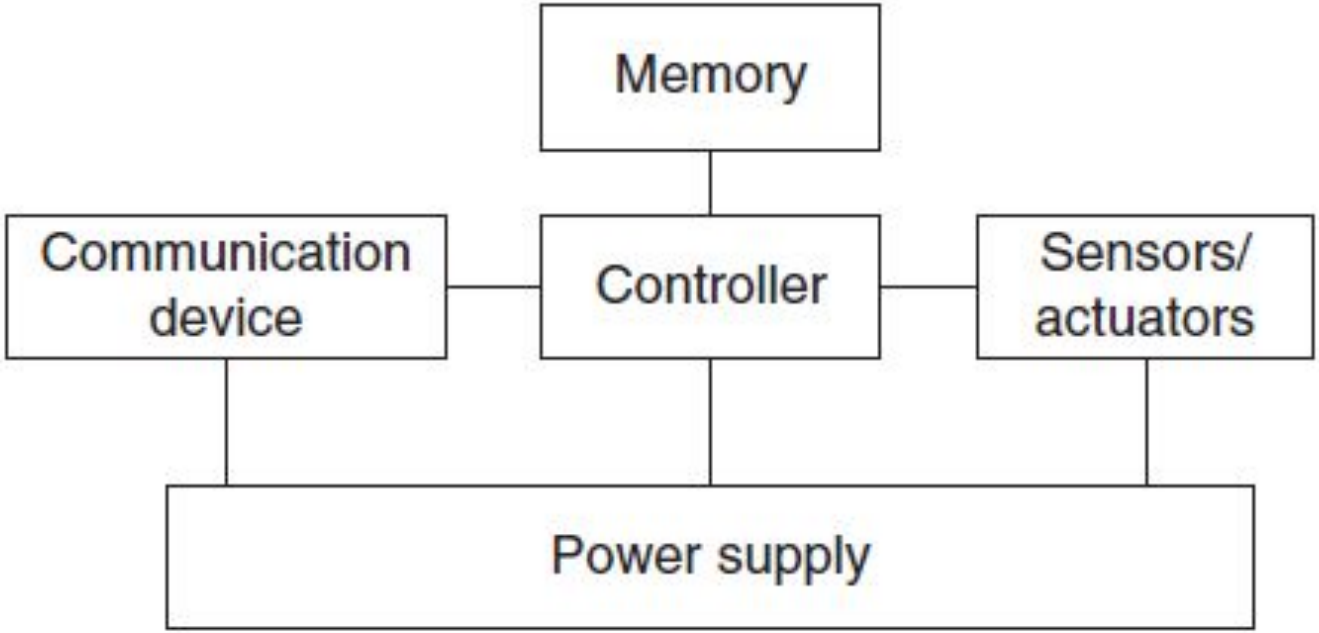


Figure 2.1 Overview of main sensor node hardware components

Hardware components & design constraints

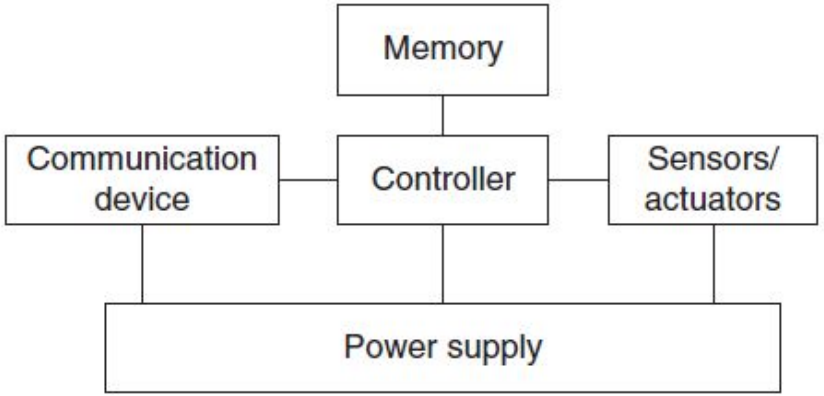


Figure 2.1 Overview of main sensor node hardware components

A basic sensor node comprises five main components:

Controller A controller to process all the relevant data, capable of executing arbitrary code.

Memory Some memory to store programs and intermediate data; usually, different types of memory are used for programs and data.

Sensors and actuators The actual interface to the physical world: devices that can observe or control physical parameters of the environment.

Communication Turning nodes into a network requires a device for sending and receiving information over a wireless channel.

Power supply As usually no tethered power supply is available, some form of batteries are necessary to provide energy. Sometimes, some form of recharging by obtaining energy from the environment is available as well (e.g. solar cells).

Controllers

1. The controller is the core of a wireless sensor node.
2. It collects data from the sensors, processes this data, decides when and where to send it, receives data from other sensor nodes, and decides on the actuator's behavior.
3. It has to execute various programs, ranging from time-critical signal processing and communication protocols to application programs; it is the Central Processing Unit (CPU) of the node.
4. For General-purpose processors applications microcontrollers are used.
5. Examples: Intel Strong ARM, Texas Instruments MSP 430, Atmel ATmeg

Memory

1. Evidently, there is a need for Random Access Memory (RAM) to store intermediate sensor readings, packets from other nodes, and so on.
2. While RAM is fast, its main disadvantage is that it loses its content if power supply is interrupted.
3. Program code can be stored in Read-Only Memory (ROM) or, more typically, in Electrically Erasable Programmable Read-Only Memory (EEPROM) or flash memory (the later being similar to EEPROM but allowing data to be erased or written in blocks instead of only a byte at a time).
4. Flash memory can also serve as intermediate storage of data in case RAM is insufficient or when the power supply of RAM should be shut down for some time.

Communication Device

1. The communication device is used to exchange data between individual nodes.
2. In some cases, wired communication can actually be the method of choice and is frequently applied in many sensor networklike settings (using field buses like Profibus, LON, CAN, or others).
3. In case of wireless communication the usual choices include radio frequencies, optical communication, and ultrasound; other media like magnetic inductance are only used in very specific cases.
4. Of these choices, Radio Frequency (RF)-based communication is by far the most relevant one as it best fits the requirements of most WSN applications
5. It provides relatively long range and high data rates, acceptable error rates at reasonable energy expenditure, and does not require line of sight between sender and receiver.

Sensors and actuators

- **Sensors**

Sensors can be roughly categorized into three categories

1. **Passive, omnidirectional sensors :**

These sensors can measure a physical quantity at the point of the sensor node without actually manipulating the environment by active probing – in this sense, they are passive.

Typical examples for such sensors include thermometer, light sensors, vibration, microphones, humidity, mechanical stress or tension in materials, chemical sensors sensitive for given substances, smoke detectors, air pressure, and so on.

2. **Passive, narrow-beam sensors**

These sensors are passive as well, but have a well-defined notion of direction of measurement. A typical example is a camera, which can “take measurements” in a given direction, but has to be rotated if need be.

Power supply

1. As usually no tethered power supply is available, some form of batteries are necessary to provide energy.
2. Sometimes, some form of recharging by obtaining energy from the environment is available as well (e.g. solar cells).
3. There are essentially two aspects: Storing energy and Energy scavenging.
4. Storing energy: Batteries

Table 2.2 Energy densities for various primary and secondary battery types [703]

Primary batteries			
Chemistry	Zinc-air	Lithium	Alkaline
Energy (J/cm^3)	3780	2880	1200
Secondary batteries			
Chemistry	Lithium	NiMHd	NiCd
Energy (J/cm^3)	1080	860	650

Network Architecture

It introduces the basic principles of turning individual sensor nodes into a wireless sensor network. In this optimization goals of how a network should function are discussed as:

1. Sensor network scenarios
2. Optimization goals and figures of merit
3. Gateway concepts

Sensor Network Scenarios

Points to remember- Sensor Node:

- Multifunctional: The number of sensor node used depends on the application type.
- Short transmission range.
- Have OS(e.g. Tiny O.S)
- Battery Powered Have limited life.

Sensor Network Scenarios

- WSN application have a source and a sink.
- In most of them, there is a clear difference between **sources** of data – the actual nodes that sense data – and **sinks** – nodes where the data should be delivered to.
- These sinks sometimes are part of the sensor network itself; sometimes they are clearly systems “outside” the network (e.g. the firefighter’s PDA communicating with a WSN).
- The interaction patterns between sources and sinks

1. **Event detection:** Sensor nodes should report to the sink(s) once they have detected the occurrence of a specified event.
2. **Periodic measurements:** Sensors can be tasked with periodically reporting measured values. Often, these reports can be triggered by a detected event; the reporting period is application dependent.
3. **Tracking:** The source of an event can be mobile (e.g. an intruder in surveillance scenarios). The WSN can be used to report updates on the event source’s position to the sink(s), potentially with estimates about speed and direction as well.

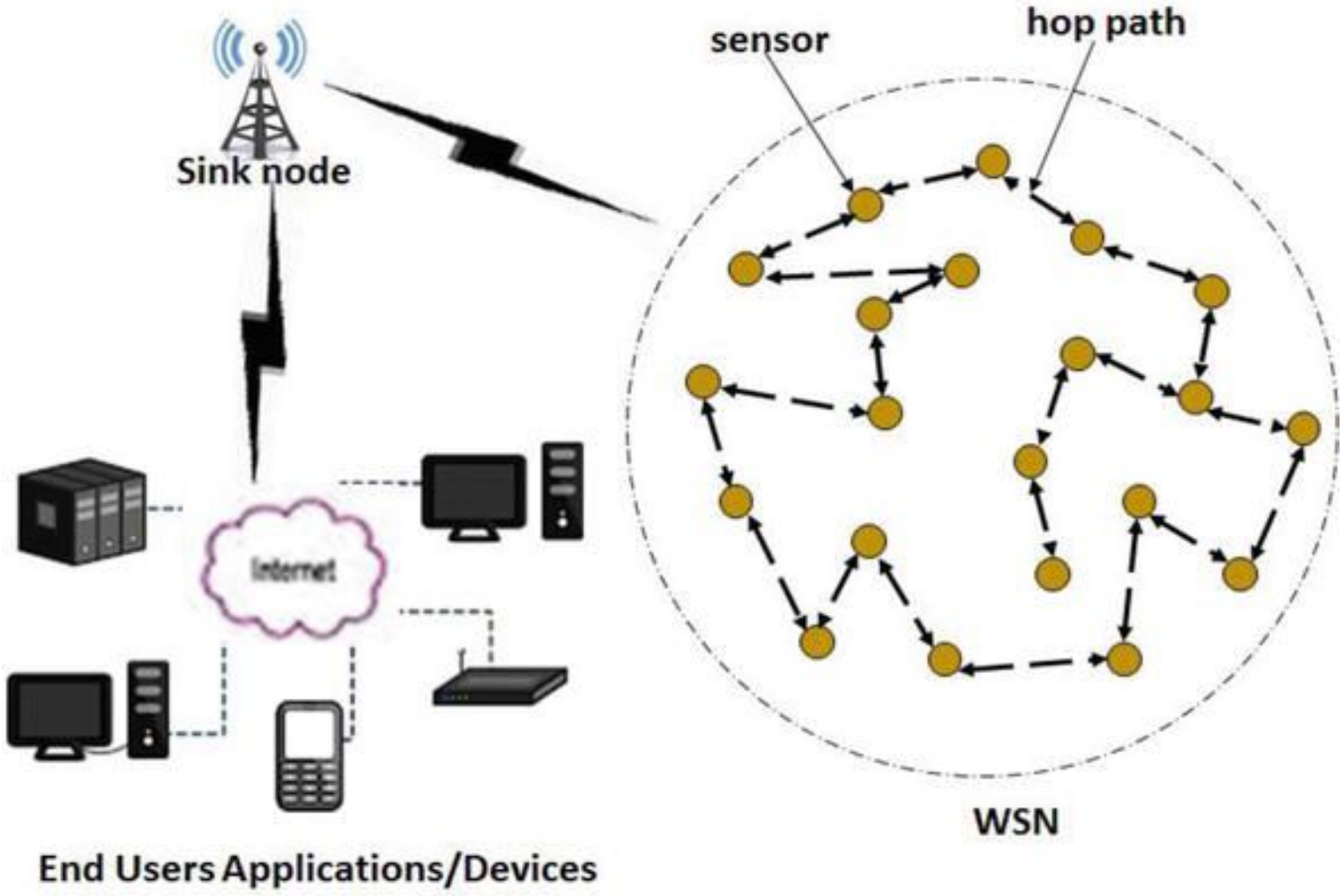
Sensor Network Scenarios

- SOURCE:

  A source is any entity in the network that can provide information, that is, typically a sensor node; it could also be an actuator node that provides feedback about an operation.

- SINK:
  A sink, on the other hand, is the entity where information is required. There are essentially three options for a sink:

1. It could belong to the sensor network as such and be just another sensor/actuator node
2. It could be an entity outside this network.
3. it could also be merely a gateway to another larger network such as the Internet


sensor
hop path
Sink node
Internet
WSN
End Users Applications/Devices

Sensor Network Scenarios

Types of mobility

1. **Node mobility**

- The wireless sensor nodes themselves can be mobile.
- The meaning of such mobility is highly application dependent.
- In examples like environmental control, node mobility should not happen; in livestock surveillance (sensor nodes attached to cattle, for example), it is the common rule.

Sensor Network Scenarios

2. Sink mobility

- The information sinks can be mobile.
- The important aspect is the mobility of an information sink that is not part of the sensor network, for example, a human user requested information via a PDA while walking in an intelligent building.
- The network, possibly with the assistance of the mobile requester, must make provisions that the requested data actually follows and reaches the requester despite its movements

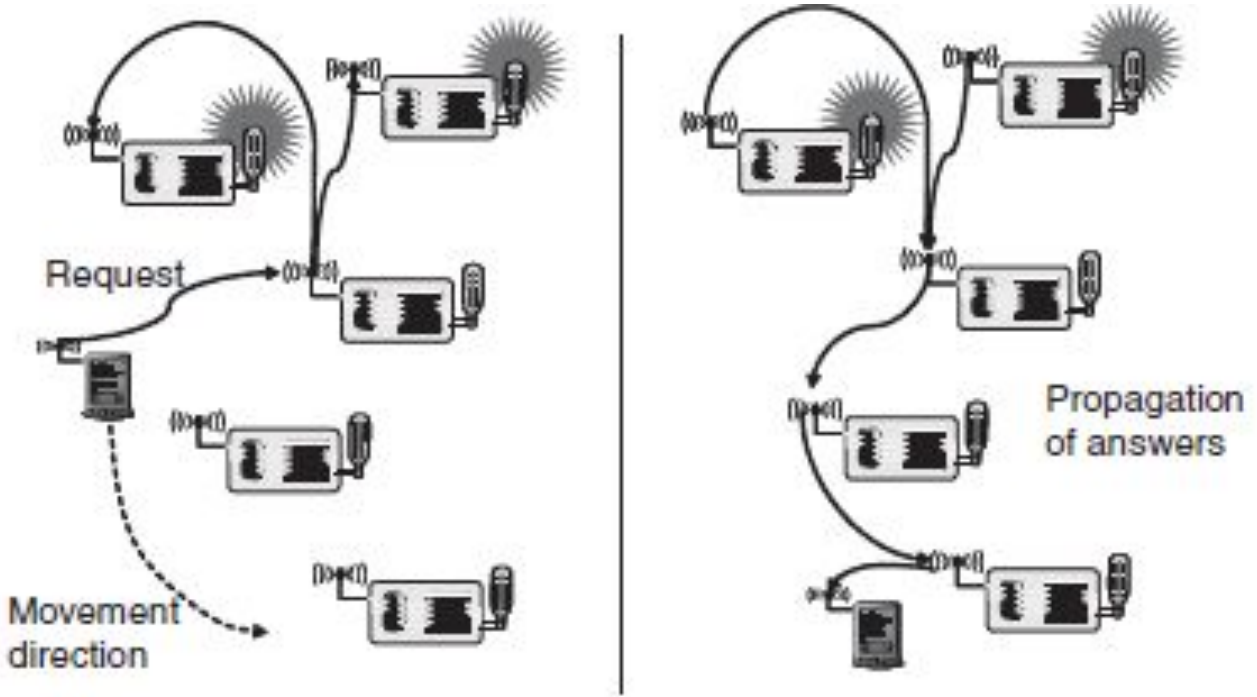
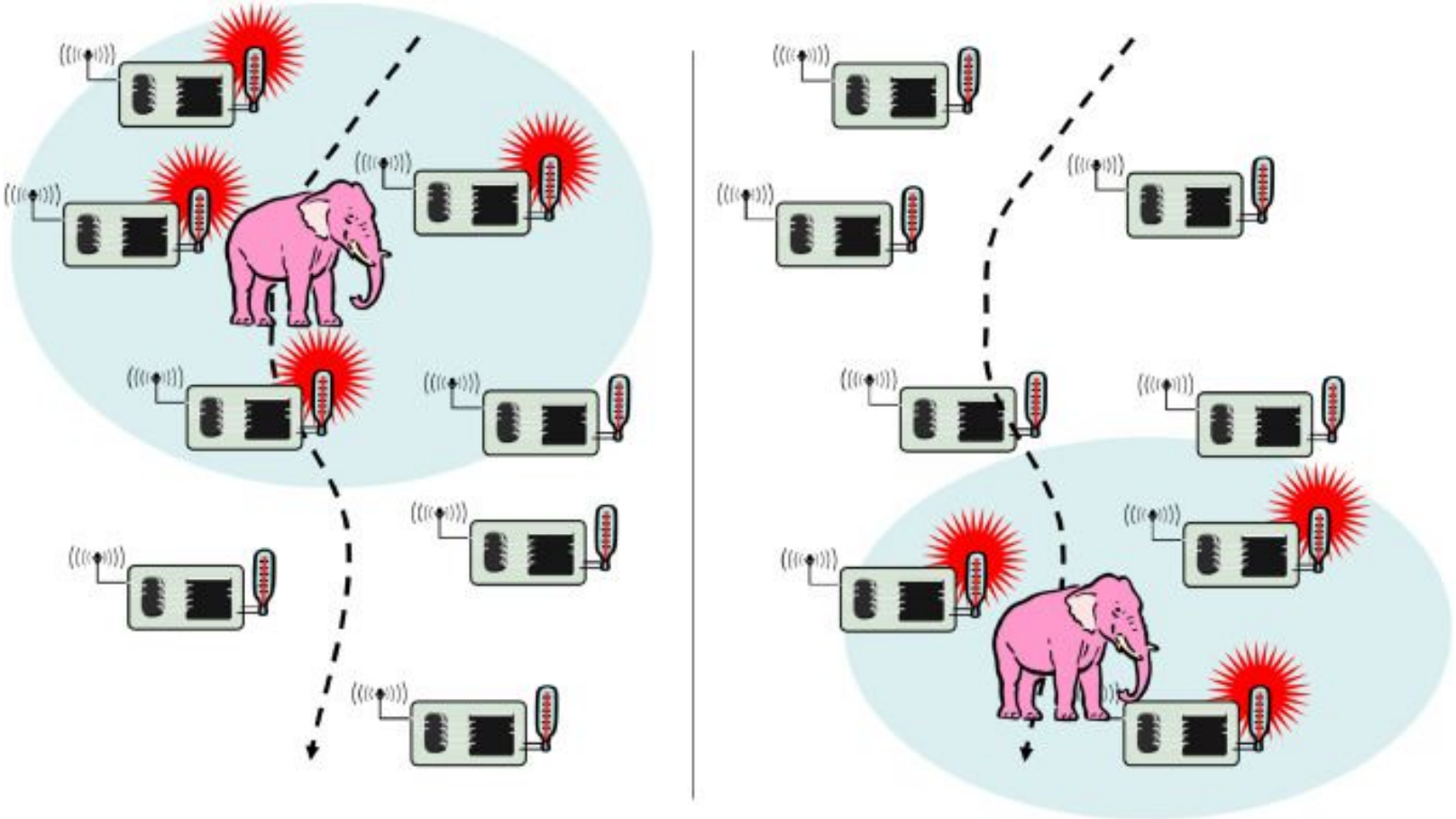


Figure 3.4 A mobile sink moves through a sensor network as information is being retrieved on its behalf

Sensor Network Scenarios

3. Event mobility

- In applications like event detection and in particular in tracking applications, the cause of the events or the objects to be tracked can be mobile.
- As the event source moves through the network, it is accompanied by an area of activity within the network – this has been called the **frisbee model**.
- This notion is described by Figure 3.5, where the task is to detect a moving elephant and to observe it as it moves around.
- Nodes that do not actively detect anything are intended to switch to lower sleep states unless they are required to convey information from the zone of activity to some remote sink



Area of sensor nodes detecting an event – an elephant [378] – that moves through the network along with the event source (dashed line indicate the elephant's trajectory; shaded ellipse the activity area following or even preceding the elephant)

Optimization goals and figures of merit

I. Quality of service

- **Event detection/reporting probability** What is the probability that an event that actually occurred is not detected or, more precisely, not reported to an information sink that is interested in such an event? For example, not reporting a fire alarm to a surveillance station would be a severe shortcoming.
- **Event classification error** If events are not only to be detected but also to be classified, the error in classification must be small.
- **Event detection delay** What is the delay between detecting an event and reporting it to any/all interested sinks?
- **Missing reports** In applications that require periodic reporting, the probability of undelivered reports should be small.
- **Tracking accuracy** Tracking applications must not miss an object to be tracked, the reported position should be as close to the real position as possible, and the error should be small.

Optimization goals and figures of merit

II. Energy efficiency

- The term “energy efficiency” is, in fact, rather an umbrella term for many different aspects of a system, which should be carefully distinguished to form actual, measurable figures of merit.
- The most commonly considered aspects are:

1. **Energy per correctly received bit** How much energy, counting all sources of energy consumption at all possible intermediate hops, is spent on average to transport one bit of information (payload) from the source to the destination? This is often a useful metric for periodic monitoring applications.
2. **Energy per reported (unique) event**

   Similarly, what is the average energy spent to report one event?

3. **Delay/energy trade-offs**

Some applications have a notion of “urgent” events, which can justify an increased energy investment for a speedy reporting of such events.

4. **Network lifetime** The time for which the network is operational

- ★ **Time to first node death-**When does the first node in the network run out of energy or fail and stopoperating?
- ★ **Network half-life-**When have 50 % of the nodes run out of energy and stopped operating
- ★ **Time to partition**-When does the first partition of the network in two (or more) disconnected partsoccur?
- ★ **Time to loss of coverage** the time when for the first time any spot in the deployment region is nolonger covered by any node‘s observations.
- ★ **Time to failure of first event notification** A network partition can be seen as irrelevant if theunreachable part of the network does not want to report any events in the first place.

Optimization goals and figures of merit

III. Scalability

- The ability to maintain performance characteristics irrespective of the size of the network is referred to as scalability.
- With WSN potentially consisting of thousands of nodes, scalability is an evidently indispensable requirement.
- Scalability is ill served by any construct that requires globally consistent state, such as addresses or routing table entries that have to be maintained.
- The need to restrict such information is enforced by and goes hand in hand with the resource limitations of sensor nodes, especially with respect to memory.
- Architectures and protocols should implement appropriate scalability support rather than trying to be as scalable as possible.

Optimization goals and figures of merit

IV. Robustness

- Related to QoS and somewhat also to scalability requirements, wireless sensor networks should also exhibit an appropriate robustness.
- They should not fail just because a limited number of nodes run out of energy, or because their environment changes and severs existing radio links between two nodes – if possible, these failures have to be compensated for, for example, by finding other routes.
- A precise evaluation of robustness is difficult in practice and depends mostly on failure models for both nodes and communication links

Design principles for WSNs

1. Distributed organization
2. In-network processing
3. Data centricity
4. Exploit location information
5. Exploit activity patterns
6. Exploit heterogeneity

Distributed organization

- Both the scalability and the robustness optimization goal are required to organize the network in a distributed fashion.
- When organizing a network in a distributed fashion, it is necessary to know potential shortcomings of this approach
- In many cases, a centralized approach can produce solutions that perform better or require fewer resources.
- One possibility is to use centralized principles in a localized fashion by electing, out of set of equal nodes.
- Such elections result in a dynamic hierarchy.
- The election process should be repeated continuously until the elected node runs out of energy

In Network Processing Techniques

- When organizing a network in a distributed fashion, the nodes in the network are not only passing on packets or executing application programs, they are also actively involved in taking decisions about how to operate the network.
- This is a specific form of information processing that happens in the network, but is limited to information about the network itself.
- It is possible to extend this concept by also taking the concrete data that is to be transported by the network into account in this information processing, making **in-network processing** a first-rank design principle.
- A few example techniques are:

1. Aggregation
2. Distributed source coding and distributed compression
3. Mobile code/Agent-based networking

In Network Processing Techniques

1. Aggregation:

- Suppose a sink is interested in obtaining periodic measurements from all sensors, but it is only relevant to check whether the average value has changed, or whether the difference between minimum and maximum value is too big.
- In such a case, it is evidently not necessary to transport are readings from all sensors to the sink, but suffices to send the average or the minimum and maximum value.
- The name aggregation stems from the fact that in nodes intermediate between sources and sinks, information is aggregated into a condensed form out of information provided by nodes further away from the sink (and potentially, the aggregator’s own readings)

In Network Processing Techniques

1. Aggregation:

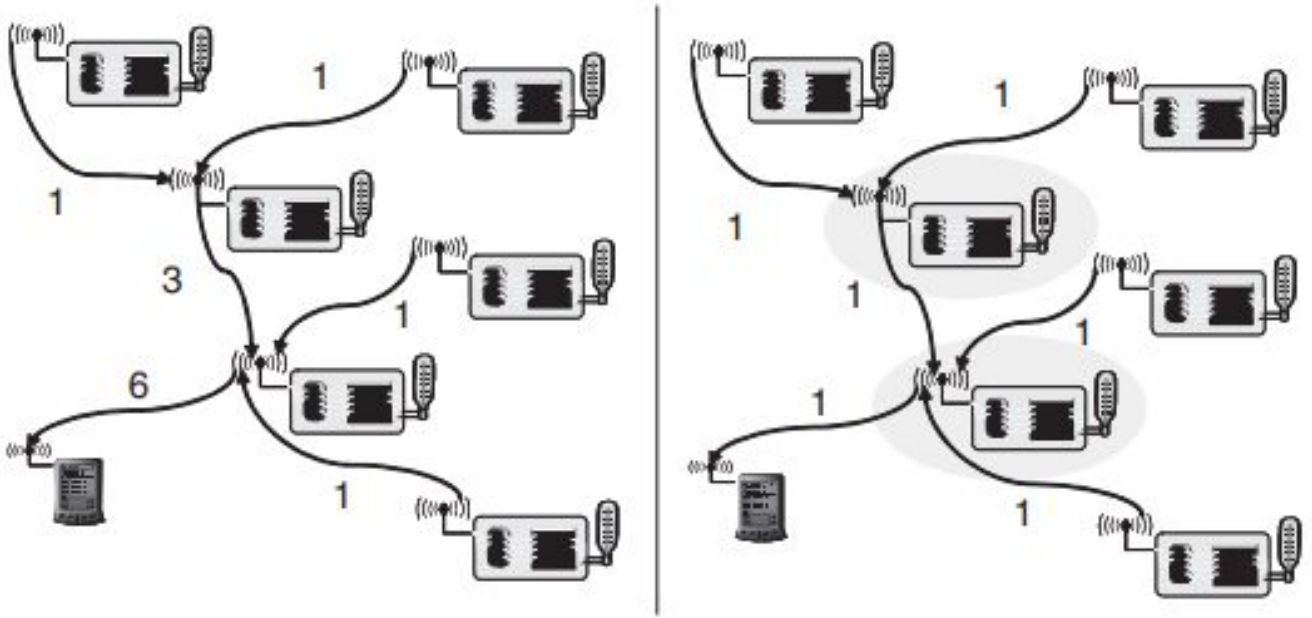


Figure 3.7 Aggregation example

Challenges in this context include how to determine where to aggregate results from which nodes, how long to wait for such results, and determining the impact of lost packets.

Data Centricity

- Address data, not nodes
- In traditional communication networks, the focus of a communication relationship is usually the pair of communicating peers – the sender and the receiver of data.
- In a wireless sensor network, on the other hand, the interest of an application is not so much in the identity of a particular sensor node, it is much rather in the actual information reported about the physical environment.
- This is especially the case when a WSN is redundantly deployed such that any given event could be reported by multiple nodes – it is of no concern to the application precisely which of these nodes is providing data.
- This fact that not the identity of nodes but the data are at the center of attention is called data-centric networking.

Data Centricity

- As an example, consider the elephant-tracking example from Figure 3.5. In a data-centric application, all the application would have to do is state its desire to be informed about events of a certain type – “presence of elephant” – and the nodes in the network that possess “elephant detectors” are implicitly informed about this request.
- In an identity-centric network, the requesting node would have to find out somehow all nodes that provide this capability and address them explicitly.

Exploit location information

- Another useful technique is to exploit location information in the communication protocols whenever such information is present.
- Since the location of an event is a crucial information for many applications, there have to be mechanisms that determine the location of sensor nodes

Exploit activity patterns

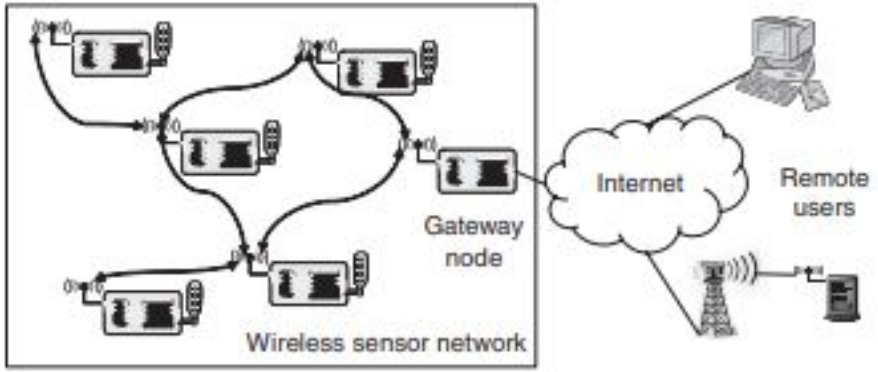
- Once an event has happened, it can be detected by a larger number of sensors, breaking into a frenzy of activity, causing a well-known event shower effect.
- Hence, the protocol design should be able to handle such bursts of traffic by being able to switch between modes of quiescence and of high activity.

Exploit heterogeneity

- Related to the exploitation of activity patterns is the exploitation of heterogeneity in the network.
- Sensor nodes can be heterogeneous by constructions, that is, some nodes have larger batteries, farther-reaching communication devices, or more processing power.
- They can also be heterogenous by evolution, that is, all nodes started from an equal state, but because some nodes had to perform more tasks during the operation of the network, they have depleted their energy resources or other nodes had better opportunities to scavenge energy from the environment
- For e.g. nodes in shade are at a disadvantage when solar cells are used, , nodes with more memory or faster processors can be better suited for aggregation, nodes with more energy reserves for hierarchical coordination etc.

Gateway concepts– Need for gateways

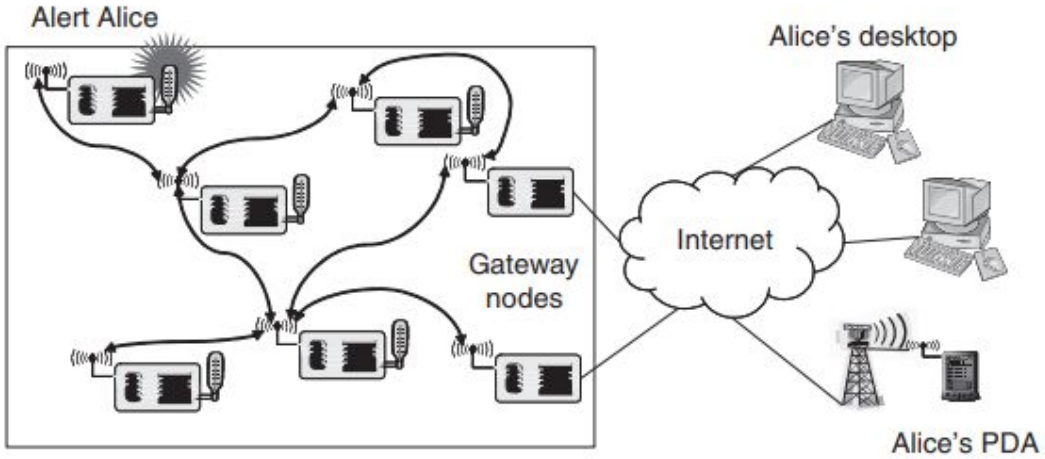
- For practical deployment, a sensor network only concerned with itself is insufficient.
- The network rather has to be able to interact with other information devices
- Eg.1- a user equipped with a PDA moving in the coverage area of the network
- Eg.2- with a remote user, trying to interact with the sensor network via the Internet (the standard example is to read the temperature sensors in one's home while traveling and accessing the Internet via a wireless connection)



- To this end, the WSN first of all has to be able to exchange data with such a mobile device or with some sort of gateway, which provides the physical connection to the Internet.
- This is relatively straightforward on the physical, MAC, and link layer – either the mobile device/the gateway is equipped with a radio transceiver as used in the WSN, or some (probably not all) nodes in the WSN support standard wireless communication technologies

WSN to Internet communication

- Assume that the initiator of a WSN–Internet communication resides in the WSN for example, a sensor node wants to deliver an alarm message to some Internet host.
- The first problem to solve is akin to ad hoc networks, namely, how to find the gateway from within the network.
- Basically, a routing problem to a node that offers a specific service has to be solved, integrating routing and service discovery



WSN to Internet communication

- If several such gateways are available, how to choose between them?
- In particular, if not all Internet hosts are reachable via each gateway or at least if some gateway should be preferred for a given destination host?
- How to handle several gateways, each capable of IP networking, and the communication among them?
- One option is to build an IP overlay network on top of the sensor network

Internet to WSN communication

- The idea is to build a larger, “virtual” WSN out of separate parts, transparently “tunneling” all protocol messages between these two networks and simply using the Internet as a transport network
- This can be attractive, but care has to be taken not to confuse the virtual link between two gateway nodes with a real link; otherwise, protocols that rely on physical properties of a communication link can get quite confused.

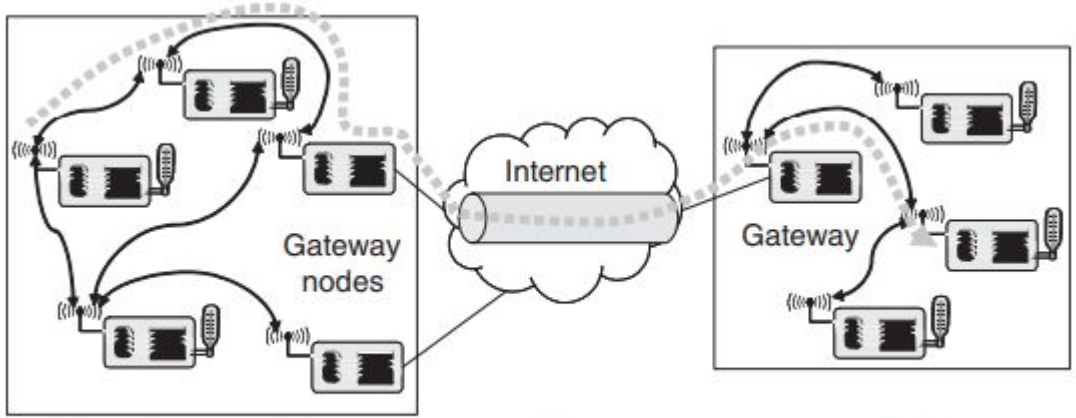


Figure 3.12 Connecting two WSNs with a tunnel over the Internet

Internet to WSN communication

- The idea is to build a larger, “virtual” WSN out of separate parts, transparently “tunneling” all protocol messages between these two networks and simply using the Internet as a transport network
- This can be attractive, but care has to be taken not to confuse the virtual link between two gateway nodes with a real link; otherwise, protocols that rely on physical properties of a communication link can get quite confused.

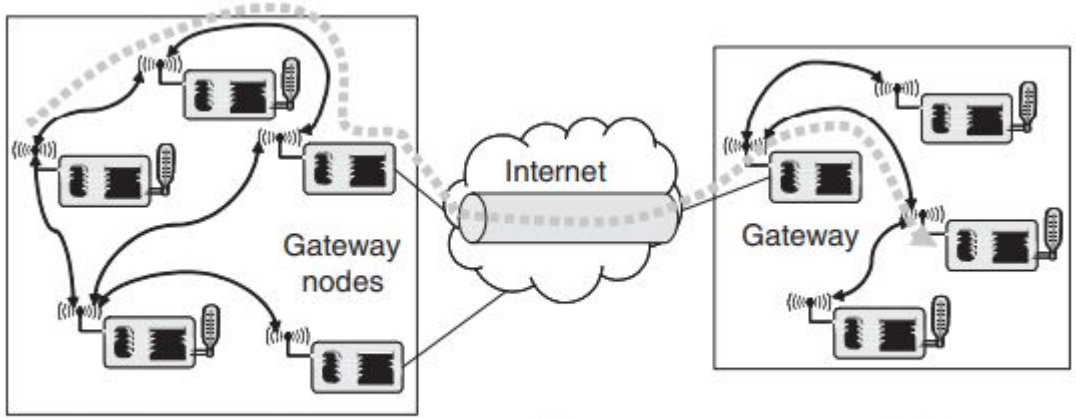


Figure 3.12 Connecting two WSNs with a tunnel over the Internet

Medium Access Control Protocols

UNIT 2

Fundamentals of MAC Protocols

1. The Medium Access Control (MAC) protocol is a set of guidelines that dictate how each node should transmit data over the shared wireless medium.
2. The primary objective of the MAC protocol is to minimize the occurrence of idle listening and collisions of data packets.
3. By efficiently managing access to the wireless medium, the MAC protocol helps to reduce energy consumption and optimize the use of network resources.
4. Delay, throughput, robustness, scalability, stability, and fairness have historically dominated MAC protocol design.

MAC Protocols-Issues

There are many issues that need to be addressed in order to design an efficient MAC protocol in a wireless ad hoc network environment.

1. **Wired networks schemes like CSMA/CD**

➔ Let us consider carrier sense multiple access with collision detection(CSMA/CD) which works as follow:

➔ A sender sense the medium (a wire) to see if it is free . If the medium is busy , the sender waits until it is free.

➔ If the medium is free sender starts transmitting data continues to listen into the medium.

➔ If sender now detects a collision while sending it stops at once and sends jamming signal .

➔ Why does this scheme fail in wireless networks?

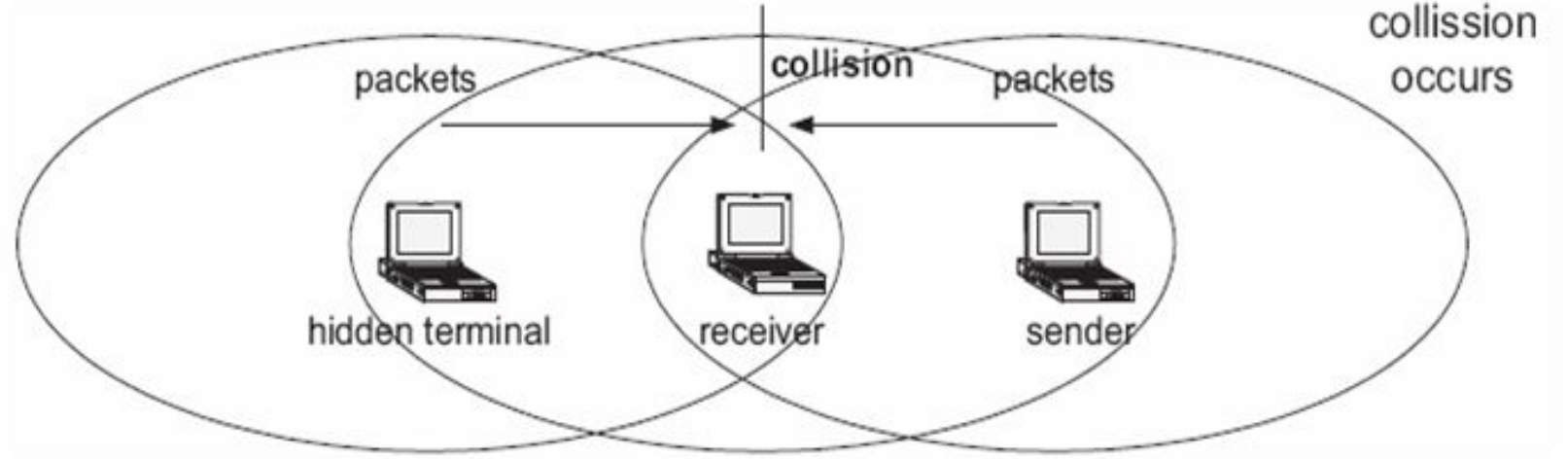
MAC Protocols-Issues

- ➔ CSMA/CD is not really interested in collision at sender, but rather in those at the receiver.
- ➔ The signal should reach the receiver without collisions.But the sender is one detecting the collisions.
- ➔ This is not problem using a wire , and if collision occurs somewhere in the wire , everybody will notice it.
- ➔ The situation is different in wireless networks.The strength of signal decreases proportionally to the square of the distance to a sender.
- ➔ The sender may now apply carrier senses and detect an idle medium.
- ➔ Thus the sender start sending but a collision happens at the receiver due to a second sender. This is hidden terminal problem.
- ➔ The sender detects no collision assumes that the data has been transmitted without errors, but actually a collision which destroyed the data at the receiver.

MAC Protocols-Issues

2. Hidden terminal problem

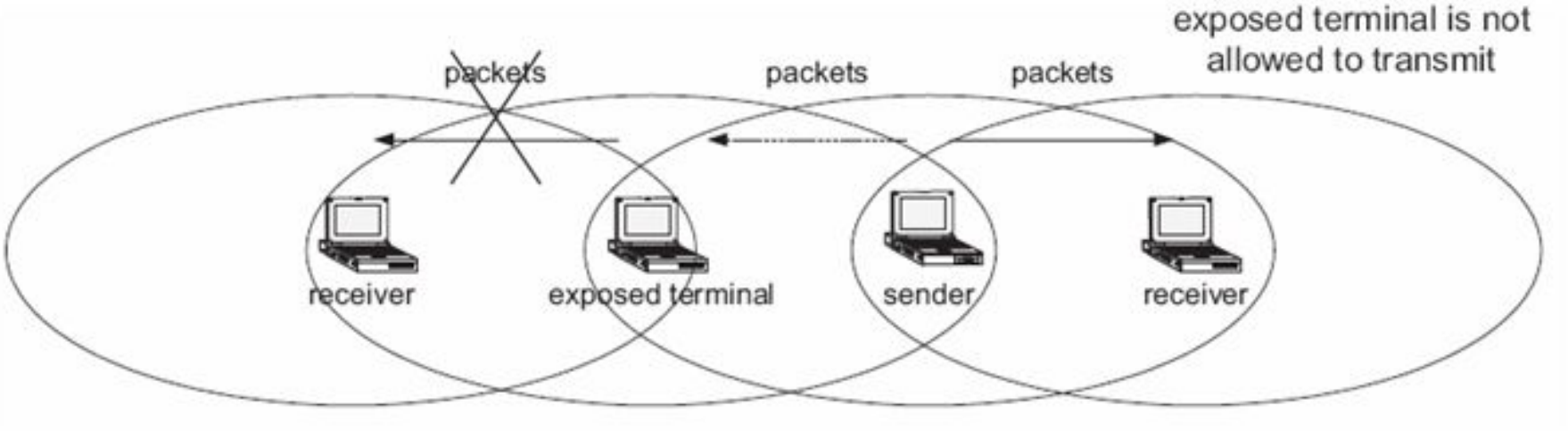
Two nodes that are outside each-other's range perform simultaneous transmission to a node that is within the range of each of them, hence, there is a packet collision.



MAC Protocols-Issues

3. Exposed terminal problem

The node is within the range of a node that is transmitting, and it cannot transmit to any node.



The exposed terminal problem is solved by the MAC (medium access control) layer protocol IEEE 802.11 RTS/CTS, with the condition that the stations are synchronized and frame sizes and data speed are the same. RTS stands for Request to Send and CTS stands for Clear to Send.

A transmitting station sends a RTS frame to the receiving station. The receiving station replies by sending a CTS frame. On receipt of CTS frame, the transmitting station begins transmission.

Any station hearing the RTS is close to the transmitting station and remains silent long enough for the CTS. Any station hearing the CTS is close to the receiving station and remains silent during the data transmission.

In the above example, station STC hears RTS from station STB, but does not hear CTS from station STA. So, it is free to transmit to station STD.

Common MAC Protocols

- The main determining factor in a WSN's success is the selection of the MAC technique.
- MAC Protocols can be roughly classified into the following classes

1) Fixed assignment protocols,
2) Demand assignment protocols
3) Random access protocols.

Common MAC Protocols

1) **Fixed assignment protocols**

- Each node is given a specified fixed quantity of the channel resources in fixed-assignment schemes
- The resource assignment is long term, and each node can use its resources exclusively without the risk of collisions.
- Frequency-division multiple access (FDMA), Time division multiple access (TDMA), and code-division multiple access(CDMA) are common protocols that fall under this category .

- **Time Division Multiple Access (TDMA)**
  - It facilitates many users to share the same frequency without interference tight time synchronization.
  - The individual mobile stations are cyclically assigned a frequency for the exclusive use of a time interval.
- **Frequency Division Multiple Access (FDMA)**
  - The frequency band is divided into channels of equal bandwidth so that each conversation is carried on a different frequency
  - Guard bands are used between the adjacent signal spectra to minimize crosstalk between the channels.
- **Code Division Multiple Access (CDMA)**
  - A sort of multiplexing that facilitates various signals to occupy a single transmission channel. It optimizes the use of available bandwidth.
  - In this system, a user has access to the whole bandwidth for the entire duration.
  - The basic principle is that different CDMA codes are used to distinguish among the different users.

Common MAC Protocols

2) Demand assignment protocols

- Primary goal is to increase channel usage by optimally or almost optimally assigning the channel's capacity to competing nodes.
- Unlike fixed-assignment schemes, where channel capacity is assigned exclusively to the network nodes in a predetermined fashion regardless of their current communication needs, demand assignment protocols ignore idle nodes and consider only nodes that are ready to transmit
- The access to the channel between competing nodes must typically be arbitrated by a network control mechanism
- Network control mechanism is divided into centralized and distributed types
- Token - and reservation-based systems employ distributed control, polling systems are an example of centralized control.

Polling

- A master control device queries, in some predetermined order, each slave node about whether it has data to transmit.
- If the node being polled has no data to transmit, it declines the controller's request.Then controller proceeds to query the next network node.
- If the polled node has data to transmit, it informs the controller of its intention to transmit.
- In response, the controller allocates the channel to the ready node which uses the full data rate to transmit its traffic.
- All nodes can receive equal access to the channel
- Drawback of polling is the substantial overhead caused by the large number of messages generated by the controller to query the communicating nodes.
- Efficiency of the polling scheme depends on the reliability of the controller.

Reservation

- The basic idea in a reservation-based scheme is to set some time slots for carrying reservation messages.
- Since these messages are usually smaller than data packets, they are called mini slots.
- When a station has data to send, it requests a data slot by sending a reservation message to the master in a reservation mini slot.
- Each station has its own reservation mini slot, collision can be avoided
- If reservation requests have a priority field, the master can schedule urgent data before delay-insensitive data.
- Packet collisions can happen only when stations contend for the mini slot, which use only a small fraction of the total bandwidth.
- Thus, the largest part of the bandwidth assigned to data packets is used efficiently

Common MAC Protocols

3) Random access protocols.

- These strategies do not assign any predictable or scheduled time for any node to transmit.
- All backlogged nodes must contend to access the transmission medium.
- Collision occurs when more than one node attempts to transmit simultaneously.
- To deal with collisions, the protocol must include a mechanism to detect collisions and a scheme to schedule colliding packets for subsequent retransmissions.
- Random access protocols were first developed for long radio links and for satellite communications. The ALOHA(Advocates of Linux Open-source Hawaii Association) protocol, also referred to as pure ALOHA, was one of the first such media access protocols.

Pure ALOHA:

- Pure ALOHA refers to the original ALOHA protocol. The idea is that each station sends a frame whenever one is available. Because there is only one channel to share, there is a chance that frames from different stations will collide.
- The pure ALOHA protocol utilizes acknowledgments from the receiver to ensure successful transmission. When a user sends a frame, it expects confirmation from the receiver.
- If no acknowledgment is received within a designated time period, the sender assumes that the frame was not received and retransmits the frame.
- When two frames attempt to occupy the channel simultaneously, a collision occurs and both frames become garbled.
- If the first bit of a new frame overlaps with the last bit of a frame that is almost finished, both frames will be completely destroyed and will need to be retransmitted.
- If all users retransmit their frames at the same time after a time-out, the frames will collide again.
- To prevent this, the pure ALOHA protocol dictates that each user waits a random amount of time, known as the back-off time, before retransmitting the frame. This randomness helps to avoid further collisions.

Slotted ALOHA:

- Slotted ALOHA is an improved version of the pure ALOHA protocol that aims to make communication networks more efficient.
- In this version, the channel is divided into small, fixed-length time slots and users are only allowed to transmit data at the beginning of each time slot.
- This synchronization of transmissions reduces the chances of collisions between devices, increasing the overall efficiency of the network.
- When a user wants to transmit a frame, it waits until the next time slot and then sends the frame. If the frame is received successfully, the receiver sends an acknowledgment.
- If the acknowledgment is not received within a time-out period, the sender assumes that the frame was not received and retransmits the frame in the next time slot.

MAC PROTOCOLS FOR WSN's

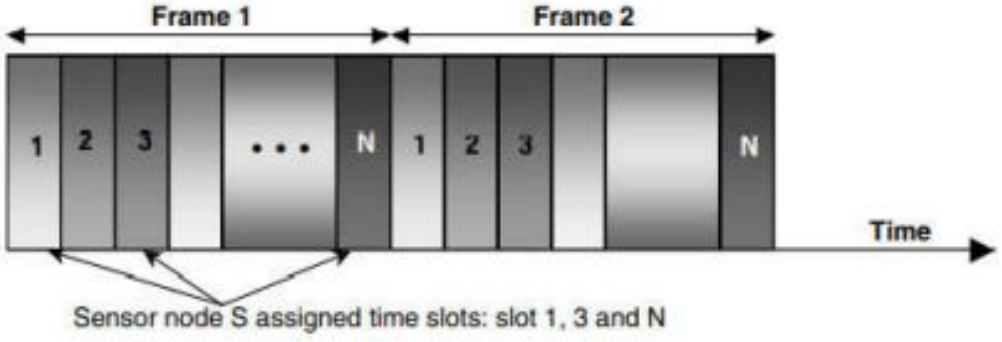
- The need to conserve energy is the most critical issue in the design of scalable and stable MAC layer protocols for WSNs.
- The main objective of most MAC-layer protocols is to reduce energy waste caused by collisions, idle listening, overhearing, and excessive overhead.
- These protocols can be categorized into two main groups:

1. Schedule- based protocols
2. Contention-based protocols/ Random access-based protocols.

MAC PROTOCOLS FOR WSN's

1. Schedule- based protocols

- It is class of deterministic MAC layer protocols in which access to the channel is based on a schedule.
- Channel access is limited to one sensor node at a time.
- This is achieved based on pre allocation of resources to individual sensor nodes.
- The majority of scheduled-based protocols for WSNs employ a kind of TDMA that divides the channel into time slots



igure 1: TDMA-based MAC protocols for wireless sensor networks

- Some schedule based protocols

1. **Self organizing medium access control for sensornets (SMACS):**

➔ SMACS essentially combines neighborhood discovery and assignment of TDMA schedules to nodes.

➔ The goal of SMACS is to detect neighboring nodes and to set up exclusive links or channels to these.

➔ A link is directional, that is on a given link all packets are transmitted in one direction. Furthermore, a link occupies a TDMA slot in either endpoint.

➔ When two nodes want bidirectional operation, two such links are needed; from the perspective of one node, there is a receive slot and a transmit slot to the other node.

➔ The assignment of links shall be such that no collisions occur at receivers.

➔ SMACS takes care that for a single node the time slots of different links do not overlap and for each link randomly one out of a large number of frequency channels/CDMA codes is picked and allocated to the link

➔ It is not required that a node and its neighbors transmit at entirely different times. In this case, however, they must transmit to different receivers and have to use different frequencies/codes.

2. Low-Energy Adaptive Clustering Hierarchy (LEACH)

- ➔ In LEACH, a TDMA-based MAC protocol is integrated with clustering and a simple “routing” protocol.
- ➔ Nodes take turns acting as cluster heads inside each cluster
- ➔ To establish communication between nodes and their cluster head, LEACH employs TDMA. Messages from the cluster head's cluster nodes are forwarded to the base station.
- ➔ A TDMA schedule is established by the cluster head node and sent to every other node in the cluster. Data message collisions are avoided by the scheduling.
- ➔ The nodes can utilize the schedule to identify the times when they need to be active as well. With the exception of the head cluster, this enables each cluster node to turn off its radio components until the designated time intervals.

SENSOR-MAC CASE STUDY

1. The sensor-MAC (S-MAC) protocol is designed explicitly to reduce energy waste caused by collision, idle listening, control overhead, and overhearing.
2. The goal is to increase energy efficiency while achieving a high level of stability and scalability.
3. S-MAC uses multiple techniques to reduce energy consumption, control overhead, and latency, in order to improve application-level performance.

SENSOR-MAC CASE STUDY-Protocol Overview

1. The protocol design assumes a large number of sensor nodes, with limited storage, communication, and processing capabilities.
2. The nodes are configured in an ad hoc, self organized, and self-managed wireless network.
3. Data generated by sensors are processed and communicated in a store-and-forward manner.
4. The applications supported by the network are assumed to alternate between long idle periods, during which no events occur, and bursty active periods, during which data flow toward the base station through message exchange among peer sensor nodes.
5. Typical applications that fall into this category include surveillance and monitoring of natural habitats and protection of critical infrastructure.

SENSOR-MAC CASE STUDY-Protocol Overview

6. In these applications the sensors must be vigilant over long periods of time, during which they remain inactive until some event occurs.S-MAC exploits the bursty profile of sensor applications to establish low-duty-cycle operation on nodes in a multihop network and to achieve significant energy savings.

7. During the long periods of time during which no sensing occurs, S-MAC nodes alternate periodically between listening and sleep modes.

Periodic Listen and Sleep Operations

- One of the S-MAC design objectives is to reduce energy consumption by avoiding idle listening.
- This is achieved by establishing low-duty-cycle operations for sensor nodes.
- Periodically, nodes move into a sleep state during which their radios are turned off completely.
- Nodes become active when there is traffic in the network.
- The basic periodic listen and sleep scheme is depicted in Figure 5.7. Based on this scheme, each node sets a wake-up timer and goes to sleep for the specified period of time.

Figure 2: S-MAC period listen and sleep modes of operations

Periodic Listen and Sleep Operations

- At the expiration of the timer, the node wakes up and listens to determine if it needs to communicate with other nodes.
- The complete listen- and-sleep cycle is referred to as a frame.
- Each frame is characterized by its duty cycle, defined as the listening interval-to-frame length ratio.
- Although the length of the listening interval can be selected independently by sensor nodes, for simplicity the protocol assumes the value to be the same for all nodes.
- Nodes are free to schedule their own sleep and listen intervals.
- It is preferable, however, that the schedules of neighboring nodes be coordinated in order to reduce the control overhead necessary to achieve communications between these nodes.

Periodic Listen and Sleep Operations

- Contrary to other protocols in which coordination is achieved through a master node such as a cluster head, S-MAC nodes form virtual clusters around schedules but communicate directly with their peers to exchange and synchronize their sleep and listen schedules.

Schedule Selection and Coordination

- The neighboring nodes coordinate their listen and sleep schedules such that they all listen at the same time and all sleep at the same time.
- To coordinate their sleeping and listening, each node selects a schedule and exchanges it with it neighbors during the synchronization period.
- Each node maintains a schedule table that contains the schedule of all its known neighbors.
- To select a schedule, a node first listens to the channel for a fixed amount of time, at least equal to the synchronization period.
- At the expiration of this waiting period, if the node does not hear a schedule from another node, it immediately chooses its own schedule.
- The node announces the schedule selected by broadcasting a SYNC packet to all its neighbors.

Schedule Selection and Coordination

- It is worth noting that the node must first perform physical carrier sensing before broadcasting the SYNC packet.
- This reduces the likelihood of SYNC packet collisions among competing nodes.
- If during the synchronization period the node receives a schedule from a neighbor before choosing and announcing its own schedule, the node sets its schedule to be the same as the schedule received.
- The node waits until the next synchronization period to announce the schedule to its neighboring nodes.
- A node may receive a different schedule after it chooses and announces its own schedule. This may occur if the SYNC packet is corrupted by either collision or channel interference.

Schedule Selection and Coordination

- If the node has no neighbor with whom it shares a schedule, the node simply discards its own schedule and adopts the new one.
- On the other hand, if the node is aware of other neighboring nodes that have already adopted its schedule, the node adopts both schedules.
- The node is then required to wake up at the listen intervals of the two schedules adopted.

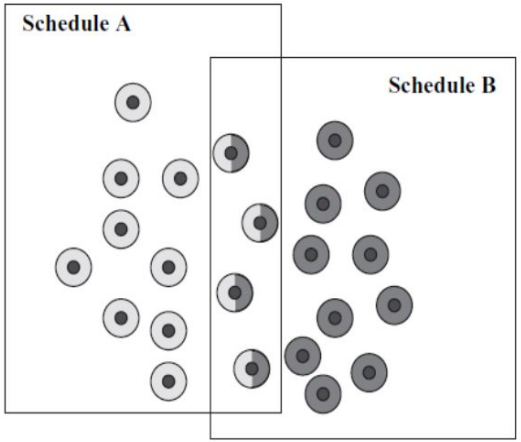


Figure 5.8 Border node schedule selection and synchronization.

Schedule Synchronization

- Neighboring nodes need to synchronize their schedules periodically to prevent long-term clock drift.
- Schedule updating is accomplished by sending a SYNC packet.
- For a node to receive both SYNC packets and data packets, the listen interval is divided into two subintervals as depicted in Figure 5.9.
- This figure illustrates three cases. In the first case the sender sends only a SYNC packet; in the second the sender sends only a data packet; and in the third the sender sends a SYNC packet in addition to the data packet.
- Access to the channel by contending nodes during these subintervals is regulated using a multi slotted contention window.

Adaptive Listening

- A closer look at the periodic listen and sleep scheme reveals that a message may incur increased latency as it is stored and forwarded between adjacent network nodes.
- If a sensor is to follow its sleep schedule strictly, data packets may be delayed at each hop.
- To address this shortcoming and improve latency performance, the protocol uses an aggressive technique referred to as adaptive listening.
- Based on this technique, a node that overhears, during its listen period, the exchange of a CTS or RTS packet between a neighboring node and another node assumes that it may be the next hop along the routing path of the overheard RTS/CTS packet, ignores its own wake-up schedule, and schedules an extra listening period around the time the transmission of the packet terminates.

Adaptive Listening

- The overhearing node determines the time necessary to complete the transmission of the packet from the duration field of the overheard CTS or RTS packet.
- Immediately upon receiving the data packet, the node issues an RTS packet to initiate an RTS/CTS handshake with the overhearing node.
- Ideally, the latter node is awake, in which case the packet forwarding process proceeds immediately between the two nodes.
- If the overhearing node does not receive an RTS packet during adaptive listening, it re-enters its sleep state until the next scheduled listen interval.

Access Control and Data Exchange

- To regulate access to the communication channel among contending sensor nodes, S-MAC uses a CSMA/CA-based procedure, including physical and virtual carrier sensing and the use of RTS/CTS handshake to reduce the impact of the hidden and exposed terminal problems.
- Virtual carrier sensing is achieved through use of the network allocation vector (NAV), a variable whose value contains the remaining time until the end of the current packet transmission.
- Initially, the NAV value is set to the value carried in the duration field of the packet transmitted.
- The value is decremented as time passes and eventually reaches zero.

Access Control and Data Exchange

- A node cannot initiate its own transmission until the NAV value reaches zero.
- Physical carrier sensing is performed by listening to the channel to detect ongoing transmission.
- Carrier sensing is randomized within a contention window to avoid collisions and starvation.
- A node is allowed to transmit if both virtual and physical carrier sensing indicate that the channel is free.
- To perform virtual carrier sensing effectively, nodes may be required to listen to all transmissions from their neighbors.
- As a result, nodes may be required to listen to packets that are destined for other nodes.

Access Control and Data Exchange

- Packet overhearing may lead to significant energy waste. To avoid overhearing, S-MAC allows nodes to move into sleep mode after they hear the exchange of an RTS or a CTS packet between two other nodes.
- The node initializes its NAV with the value contained in the duration field of the RTS or CTS packets and enters the sleep state until the NAV value reaches zero.
- Since data packets are typically larger than control packets, the overhearing avoidance process may lead to significant energy savings.
- The scheme used by S-MAC to avoid collisions is illustrated in figure

Access Control and Data Exchange

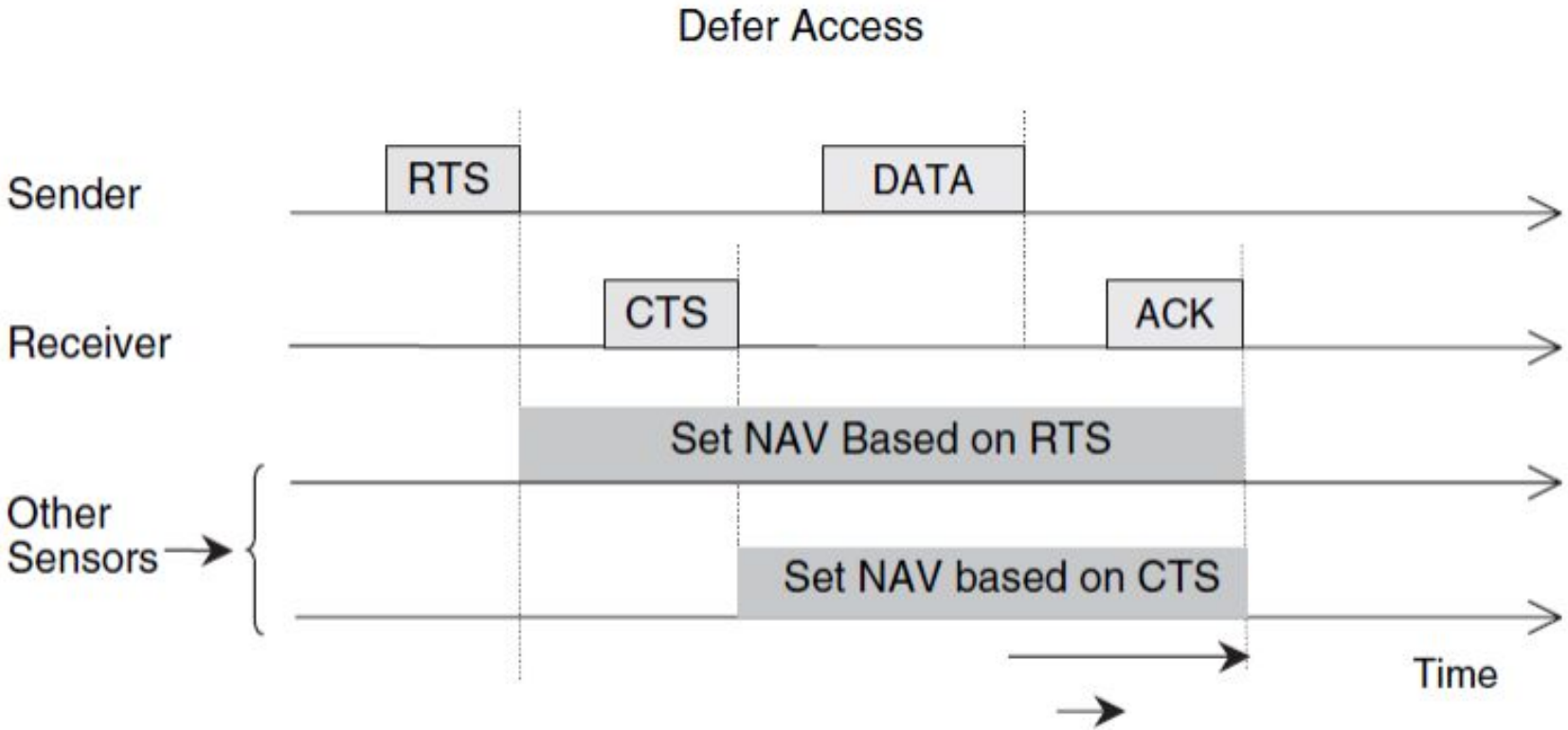


Figure 5.10 S-MAC collision avoidance scheme [5.46].

Access Control and Data Exchange

- A node attempting to transmit a message must first sense the channel. If the channel is busy, the node goes to sleep and wakes up when the channel becomes free again.
- If the channel is idle, a node, sending a data packet, first issues an RTS packet and waits for a CTS packet from the receiver.
- When it receives the CTS packet, the node sends its data packet. The transaction is completed when the node receives an acknowledgment from the receiver.
- After successful exchange of the RTS and CTS packets, the communicating nodes use their normal sleep time to exchange data packets.
- The nodes do not resume their regular sleep schedule until the data transmission is completed.
- Furthermore, the transmission of a broadcast packet, such as a SYNC packet, does not require the exchange of the RTS and CTS packet.

Message Passing

- To improve application-level performance, S-MAC introduces the concept of message passing, where a message is a meaningful unit of data that a node can process.
- Messages are divided into small fragments. These fragments are then transmitted in a single burst.
- The fragments of a message are transmitted using only one RTS/CTS exchange between the sending and receiving nodes.
- At the completion of this exchange, the medium is reserved for the time necessary to complete the transfer of the entire message successfully.
- Furthermore, each fragment carries in its duration field the time needed to transmit all the subsequent fragments and their corresponding acknowledgments. This procedure is depicted in Figure 5.11.

Message Passing

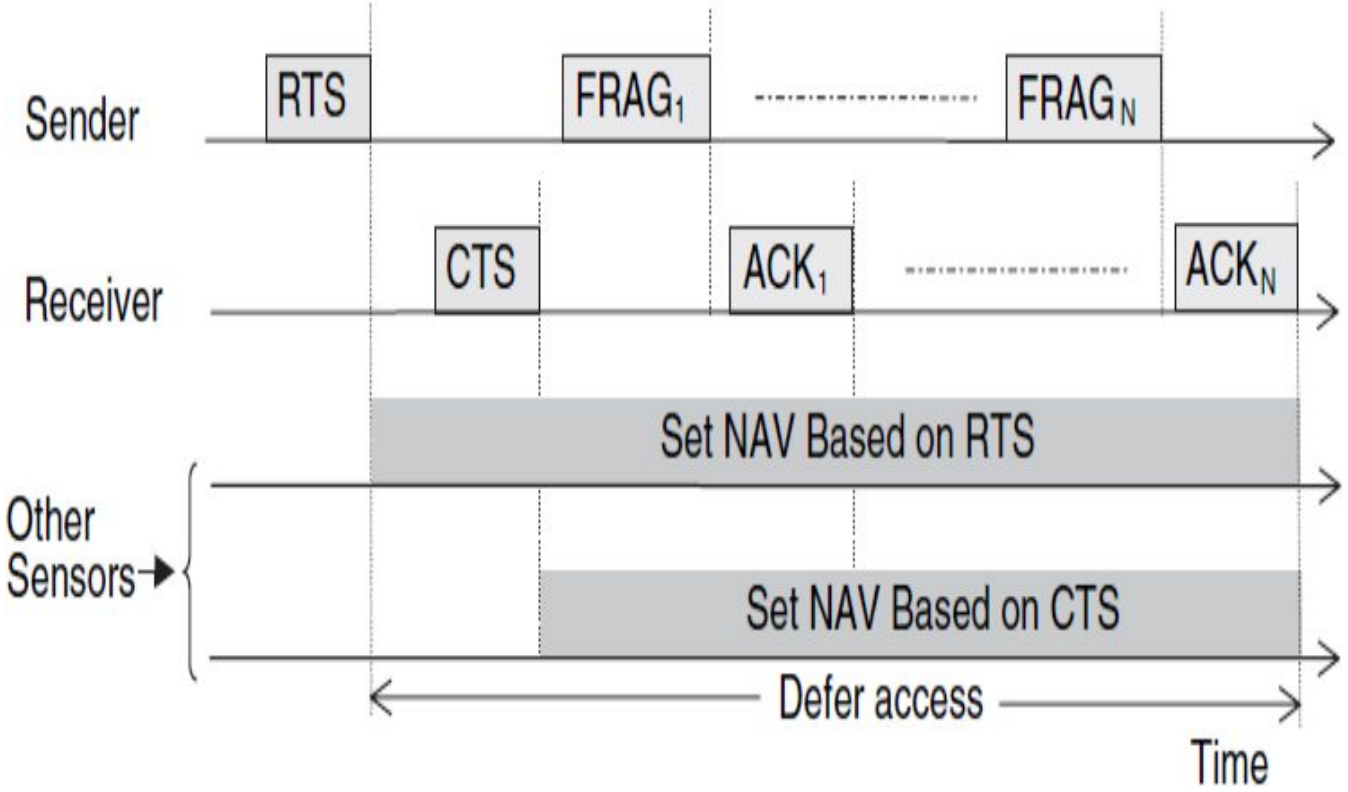


Figure 5.11 S-MAC message passing [5.46].

Message Passing

- Upon transmitting a fragment, the sender waits for an acknowledgment from the receiver.
- If it receives the acknowledgment, the sender proceeds with transmission of the next fragment.
- If it fails to receive the acknowledgment, however, the sender extends the time required to complete transmission of the segment to include the time to transmit one more fragment and its corresponding acknowledgment and retransmits the unacknowledged frame immediately.
- Sleeping nodes can hear about this extension only if they hear extended fragments or their corresponding acknowledgments.
- Nodes that only heard the initial RTS and CTS packet exchange remain unaware of the transmission extension.

Routing Protocols

Unit 2

Introduction

- The primary duty of wireless sensor nodes is to detect and gather data from a target domain, process the data, and communicate the information back to specified sites where the underlying application lives, notwithstanding the diversity in the goals of sensor applications.
- The process of data collection and forwarding is either triggered by the occurrence of specific events in the environment where the sensors are deployed or is initiated in response to a query issued by the application supported.
- It is worth noting that in many cases it is useful to aggregate data collected by various sensors before forwarding the data to the base station.
- Data aggregation reduces the number of messages transmitted, leading to a significant decrease in energy consumption due to communication
- The creation of an energy-efficient routing protocol is necessary to create pathways between sensor nodes and the data sink in order to do this operation effectively.
- The network lifetime must be maximized by the path selection process.

Data Dissemination and Gathering

- A data dissemination is a process by which data and queries for data are routed in the sensor network.
- A simple approach to accomplishing this task is for each sensor node to exchange data directly with the base station.
- A single-hop-based approach, however, is costly, as nodes that are farther away from the base station may deplete their energy reserves quickly, thereby severely limiting the lifetime of the network.
- This is the case particularly where the wireless sensors are deployed to cover a large geographical region or where the wireless sensors are mobile and may move away from the base station.
- Data sharing between the sensors and base stations is typically done utilizing multi-hop packet transmission over short communication distances to solve the drawbacks of the single-hop strategy.

Data Dissemination and Gathering

- In very dense WSNs, such an approach results in significant energy savings and lowers communication interference amongst sensor nodes competing for the channel.
- Data gathered by the sensors is transferred to the base station utilizing multi-hop pathways in response to requests made by the sinks or when particular events take place in the region being monitored.
- It is important to note that, depending on the application, sensor nodes may gather data that has been linked while travelling to the base station.
- Data packets between the source and the destination must be forwarded by intermediate nodes in a multi-hop WSN.
- The main purpose of the routing algorithm is to choose which group of intermediary nodes should be used to create a data-forwarding path between the source and the destination.

ROUTING CHALLENGES AND DESIGN ISSUES IN WIRELESS SENSOR NETWORKS

Challenges can be attributed to multiple factors, including severe energy constraints, limited computing and communication capabilities, the dynamically changing environment within which sensors are deployed, and unique data traffic models and application-level quality of service requirements.

Some challenger are:

1. Network Scale and Time-Varying Characteristics
2. Resource Constraints
3. Sensor Applications Data Models

ROUTING CHALLENGES AND DESIGN ISSUES IN WIRELESS SENSOR NETWORKS

1. Network Scale and Time-Varying Characteristics

- With severe energy constraints, sensor nodes can only operate with limited processing, storage, and communication capabilities.
- The densities of the WSNs may differ greatly, ranging from very sparse to very dense, due to the numerous possible sensor-based applications.
- However, in many applications, the sensor nodes which can sometimes number in the hundreds or even thousands are set up haphazardly and frequently without supervision over large coverage regions.
- As a result of the requirement to self-organize and conserve energy, sensor nodes in these networks behave in a dynamic and highly adaptive manner, continually adjusting to their level of activity or lack thereof.
- In order to avoid the severe performance deterioration of the supported application, sensor nodes may also be necessary to modify their behavior in response to the irregular and unpredictable behavior of wireless connections induced by excessive noise levels and radio-frequency interference.

ROUTING CHALLENGES AND DESIGN ISSUES IN WIRELESS SENSOR NETWORKS

2. Resource Constraints

- In order to deploy sensor nodes widely and cheaply, complexity is kept to a minimum.
- WSNs must operate on limited battery reserves while achieving a long lifetime, hence energy is a major challenge.
- A significant source of power consumption in wireless networks is multihop packet transmission. The duty cycle of the wireless sensors can be dynamically controlled to lower energy consumption.
- Yet, many mission-critical sensor applications make the energy management challenge particularly difficult. Due to the demands of these applications, it is necessary to concurrently maintain a set level of sensing and communication performance limits.
- Hence, the topic of how to create scalable routing algorithms that can function well under a variety of performance limitations and design requirements arises.The development of these protocols is fundamental to the future of WSNs.

ROUTING CHALLENGES AND DESIGN ISSUES IN WIRELESS SENSOR NETWORKS

3. Sensor Applications Data Models

- The information flow between the sensor nodes and the data sink is described by the data model. How data are sought and used depends a lot on the type of application used in these models.
- Data collection models for a class of sensor applications must be based on periodic sampling or be triggered by the occurrence of particular events. Before being sent to the data sink, data can be collected, stored, and possibly processed by a sensor node in other applications.
- A third category of sensor applications, however, necessitates bidirectional data models since they call for two-way communication between sensors and data sinks.The complexity of the route design challenge is increased by the requirement to support numerous data models.
- It becomes a massive design and engineering challenge to tailor the routing protocol to the particular data needs of an application while also supporting a wide range of data models and providing the best possible performance in terms of scalability, reliability, responsiveness, and power efficiency.

ROUTING STRATEGIES IN WIRELESS SENSOR NETWORKS

- Routing algorithms for ad hoc networks can be classified according to the manner in which information is acquired and maintained and the manner in which this information is used to compute paths based on the acquired information.
- Three different strategies can be identified:

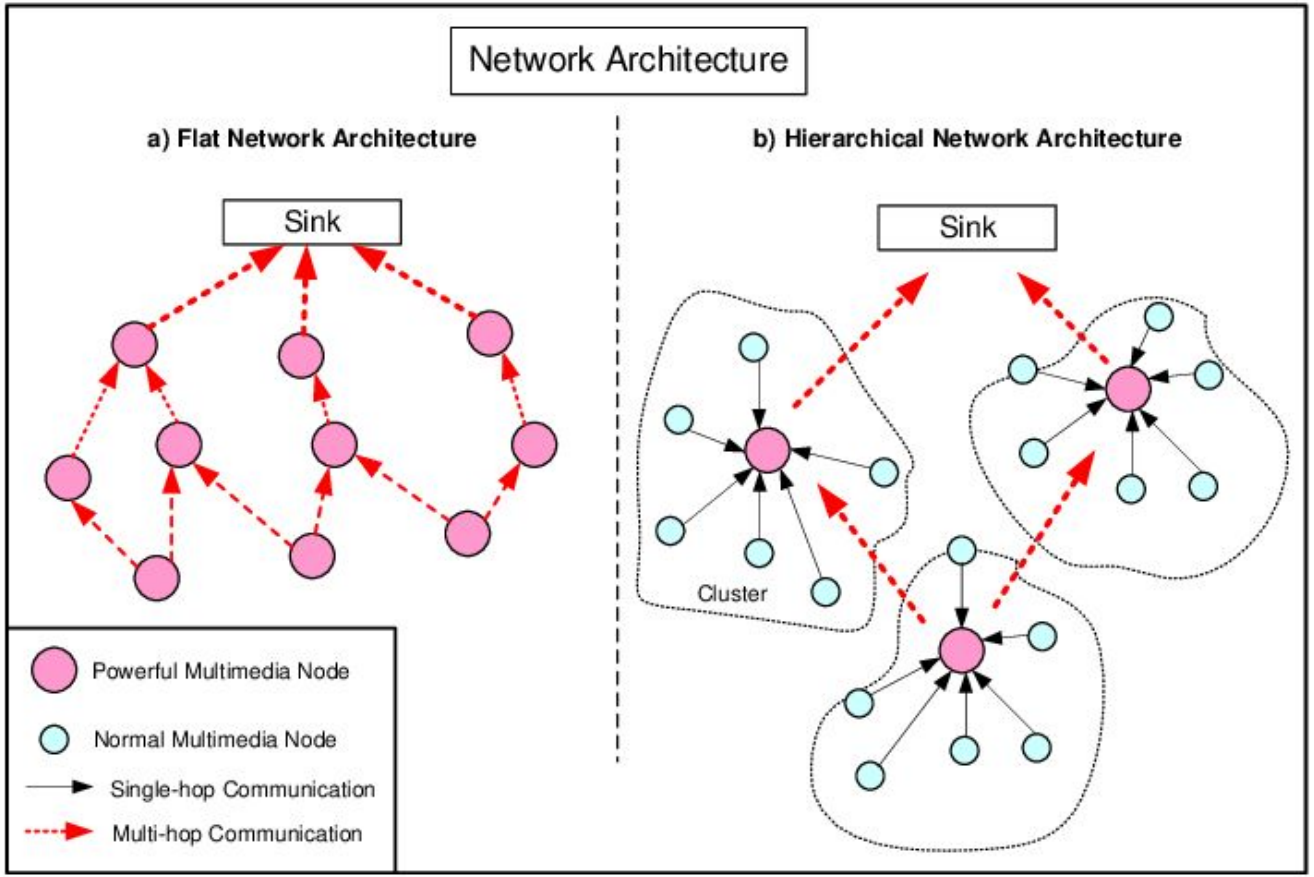
1. Proactive
2. Reactive
3. Hybrid

ROUTING STRATEGIES IN WIRELESS SENSOR NETWORKS

1. **Proactive**

- The proactive approach, also known as table-driven, focuses on redistributing routing data(contains the information of the routes to all the possible destination mobile nodes) on a regular basis to keep routing tables accurate and consistent across all network nodes.
- The network's structure can either be flat or hierarchical.
- Flat proactive routing strategies have the potential to compute optimal paths.
- The overhead required to compute these paths may be prohibitive in a dynamically changing environment.
- Hierarchical routing is better suited to meet the routing demands of large ad hoc networks

ROUTING STRATEGIES IN WIRELESS SENSOR NETWORKS



ROUTING STRATEGIES IN WIRELESS SENSOR NETWORKS

2. Reactive

- Reactive routing strategies establish routes to a limited set of destinations on demand. The route is discovered only when it is required/needed
- These strategies do not typically maintain global information across all nodes of the network.
- They must therefore, rely on a dynamic route search to establish paths between a source and a destination. It consists of two major phases namely, route discovery and route maintenance
- Route discovery typically involves flooding a route discovery query, with the replies traveling back along the reverse path.
- Route Maintenance performs the maintenance work of the route as the topology in the mobile ad-hoc network is dynamic in nature and hence, there are many cases of link breakage resulting in the network failure between the mobile nodes
- The reactive routing strategies vary in the way they control the flooding process to reduce communication overhead and the way in which routes are computed and re established when failure occurs.

ROUTING STRATEGIES IN WIRELESS SENSOR NETWORKS

3. Hybrid

- Hybrid strategies rely on the existence of network structure to achieve stability and scalability in large networks.
- In these strategies the network is organized into mutually adjacent clusters, which are maintained dynamically as nodes join and leave their assigned clusters.
- Clustering provides a structure that can be leveraged to limit the scope of the routing algorithm reaction to changes in the network environment.
- A hybrid routing strategy can be adopted whereby proactive routing is used within a cluster and reactive routing is used across clusters.
- The main challenge is to reduce the overhead required to maintain the clusters

ROUTING STRATEGIES IN WIRELESS SENSOR NETWORKS

In summary, traditional routing algorithms for ad hoc networks tend to exhibit their least desirable behavior under highly dynamic conditions. Routing protocol overhead typically increases dramatically with increased network size and dynamics. A large overhead can easily overwhelm network resources. Consequently, although they are well adapted to operate in environments where the computation and communications capabilities of the network nodes are relatively high compared to sensor nodes, the efficiency of these techniques conflict with routing requirements in WSNs.New routing strategies are therefore required for sensor networks that are capable of effectively managing the trade-off between optimality and efficiency.

WSN Routing Techniques

1. One class of routing systems uses a flat network topology where all nodes are viewed as peers. Low infrastructure maintenance costs and the possibility of finding several pathways between communication nodes for fault tolerance are two advantages of a flat network architecture.Eg. Flooding and Gossiping

2. A second class of routing protocols imposes a structure on the network. Network nodes are arranged in clusters in this family of protocols, with a node having a greater residual energy, for example, acting as the cluster head. Coordination of group activities and information sharing amongst clusters fall within the purview of the cluster head. It is possible to lower energy usage and increase network lifetime by clustering.Eg. Low-Energy Adaptive Clustering Hierarchy (LEACH)

WSN Routing Techniques

3. A third class of routing protocols uses a data-centric approach to disseminate interest within the network. The approach uses attribute-based naming, whereby a source node queries an attribute for the phenomenon rather than an individual sensor node. Different strategies can be used to communicate interests to the sensor nodes, including broadcasting, attribute-based multicasting, geo-casting, and anycasting.Eg. Sensor Protocols for Information via Negotiation(SPIN)

4. A fourth class of routing protocols uses location to address a sensor node. Location-based routing is useful in applications where the position of the node within the geographical coverage of the network is relevant to the query issued by the source node. Such a query may specify a specific area where a phenomenon of interest may occur or the vicinity to a specific point in the network environment.

1. Flooding

It is the simplest design. In this method, each node receiving data repeats it by broadcasting the data to every neighbor unless the maximum hop lifetime of the data has been reached.

- **Advantages**

1. Simple to setup and implement.
2. Data and queries reach all the nodes in the network.

- **Disadvantages of Flooding**

1. Implosion- No restriction on multiple nodes sending same packets to the same destination.
2. Overlapping- Neighbor nodes may receive the same message if the nodes access the same event.
3. Resource Blindness- Flooding does not care about energy efficiency of the nodes.

2. Gossiping

It is the enhancement of Flooding. In this, when a node receives data, it randomly chooses a neighbor and sends the data to it. Unlike Flooding, we does not need to bother about duplicate data packets being sent to the same location. It also contributes to the latency of network.

- **Advantages**

1. This protocol is easily scalable.
2. It eliminates some of the shortcomings of Flooding.
3. This protocol sends data in autonomous and decentralized manner

- **Disadvantages of Flooding**

The destination is selected randomly so it may lead to starvation for some nodes as they may not be selected to send data at all.

3. Sensor Protocols for Information via Negotiation - SPIN

Sensor Protocols for Information via Negotiation (SPIN) has the required features which can overcome the shortcomings of flooding. When interested nodes send a request, SPIN will send the data to the corresponding node otherwise it will not on its own. SPIN messages can be distinguished into three types:

- ➔ ADV- ADV message is used to signal that the sensor has data to send and describes the data by the help of a sensor
- ➔ REQ- REQ message is used when a node is ready to receive data from neighboring node
- ➔ DATA- The information to be sent is contained here

Advantages

SPIN is more efficient than flooding since the negotiation reduces the implosion and overlap.

Disadvantages

SPIN-2 is more effective than SPIN-1 as it uses energy or resource threshold so that limited number of nodes can participate in data transmission.

TRANSPORT CONTROL PROTOCOLS

Traditional Transport Control Protocols-TCP, UDP & Mobile IP

TCP

- On the Internet, TCP is the most widely used connection-oriented transport control protocol.
- In order to provide dependable, orderly, controllable, and elastic transmission, TCP makes advantage of network services offered by the IP layer.
- TCP operation is divided into three stages:

1. Connection establishment
2. Data transmission
3. Disconnect

TCP

1. **Connection establishment**

- During this stage, a logical connection for TCP is formed. A logical connection is an association between a TCP sender and recipient that can be uniquely identified by their IP addresses and TCP port numbers.
- There could be many connections active between endpoints at once. These connections will have separate TCP port numbers despite sharing the same IP address.
- A three-way handshake is used by TCP to establish a connection.
- The TCP sender and receiver will negotiate parameters like the beginning sequence number, window size, and others during the handshake and will let each other know when data transmission can start

2. Data transmission

- TCP enables dependable and well-organized information transfer between the sender and the recipient.
- When a segment is lost, TCP utilizes (accumulative) ACK to find it. The segment header's sequence number allows for an ordered transmission.
- TCP also supports flow control and congestion control with sender adjustable transmission rates.
- This task is carried out via TCP using a window-based approach, in which the sender manages a variable called cwnd (congestion window).
- The maximum number of segments that the TCP sender can send is cwnd. After receiving an ACK from the receiver or following a timeout, cwnd is updated.
- Both flow control and delivery notification are performed using ACK, therefore the two tasks are somewhat intertwined.

- Via cwnd, TCP controls its flow and congestion. The procedure is divided into three stages:
  - **Slow start**: All transmissions begin slowly by default. For each ACK that is received during this phase for a segment that was transmitted, the cwnd rises by one.If ACK is not received because of segment loss, cwnd consequently rises.
  - **Congestion avoidance:** The system enters the congestion avoidance state when cwnd reaches a maximum value (threshold). The system enters the slow start phase once more, the threshold is set to half of the current cwnd, the segment timer is doubled, and the cwnd is reset if the timer ends before an ACK corresponding to the segment is received. Round-trip time (RTT), which is determined by the ACK, is used to update the timer
  - **FRFT:** The same technique that updates cwnd in the congestion avoidance state is also employed in the rapid recovery and fast retransmission state.

3. Disconnect

The connection will be cut off and the relevant resource released once the data transmission is finished.

UDP

- User Datagram Protocol (UDP) is a Transport Layer protocol.Unlike TCP, it is an unreliable and connectionless protocol.
- This protocol does not have any techniques for recovering lost information because it exchanges datagrams without a sequence number between the transmitter and the receiver.
- It cannot guarantee ordered transmission because the datagrams do not include a sequence number.
- Moreover, it lacks features for flow control or congestion.
- As UDP does not perform congestion or flow control, it can end up outperforming TCP in situations where both protocols are active.
- A TCP-friendly rate control (TFRC) for UDP has been suggested in recent years to implement a certain amount of control in this protocol.
- When TCP and UDP are available on a connection, the fundamental principle underlying TFRC is to deliver almost comparable throughput to both protocols.

Protocol	TCP	UDP
Connection	connection-oriented	connectionless
Usage	high reliability, critical-less transmission time	fast, efficient transmission, small queries, huge numbers of clients
Ordering of data packets	rearranges packets in order	no inherent order
Reliability	yes	no
Streaming of data	read as a byte stream	sent and read individually
Error checking	error checking and recovery	simply error checking, no error recovery
Acknowledgement	acknowledgement segments	no acknowledgment

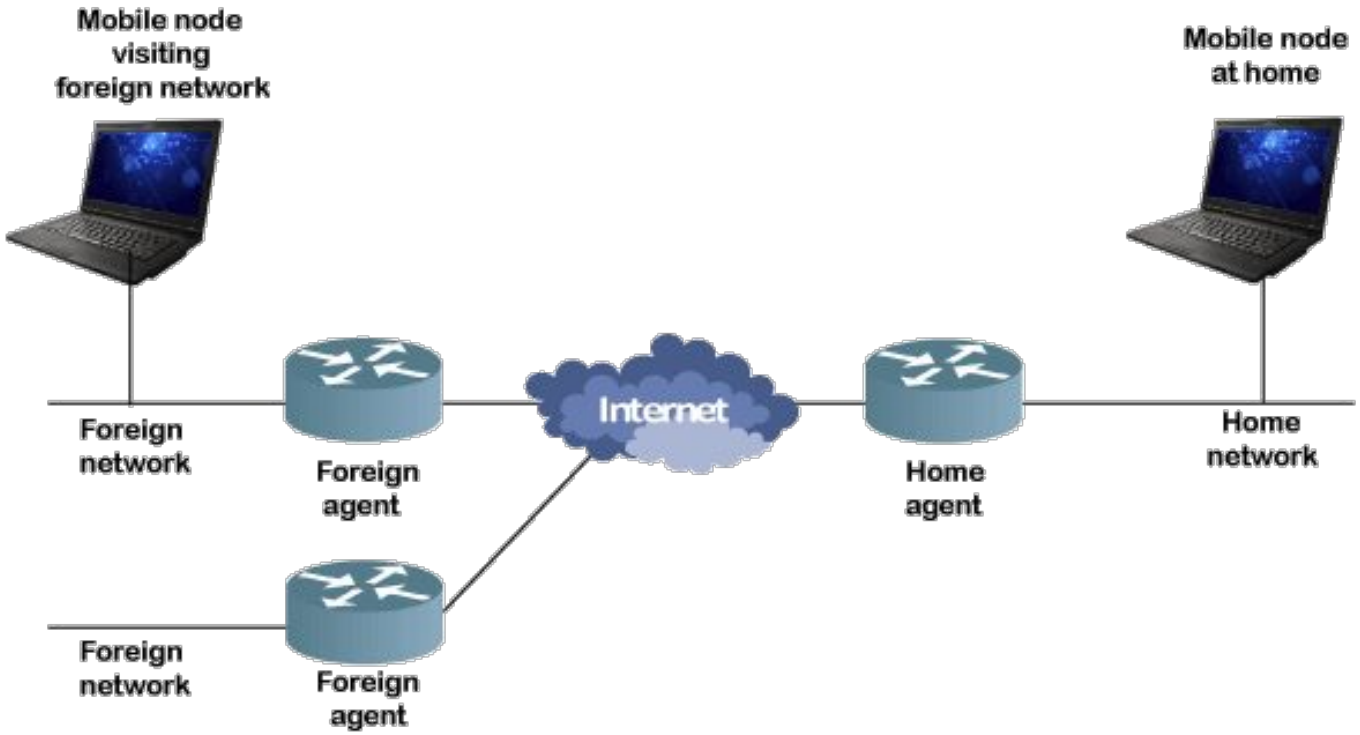
Mobile IP

- Mobile IP is a communication protocol (created by extending Internet Protocol, IP) that allows the users to move from one network to another with the same IP address.
- In IP networks, when a device is within its home network, the routing is based on the static IP addresses. The device within a network is connected through normal IP routing by the IP address assigned on the network.
- It is the same as how a postal letter is delivered to the fixed address on the envelope. The problem occurs when a device goes away from its home network and is no longer reachable using normal IP routing.
- In this condition, the active sessions of the device are terminated. The idea of Mobile IP was introduced to resolve this issue.
- It facilitates users to keep the same IP address while going to a different network or a different wireless operator without being communication disrupted or without sessions or connections being dropped.
- Remote login, remote printing, and file transfers are some examples of applications where it is undesirable to interrupt communications while an individual roams across network boundaries.

Mobile IP architecture

Mobile IP has the following three components, as shown in Figure:

- Mobile Node
- Home Agent
- Foreign Agent



Mobile IP architecture

- The **Mobile Node** is a device such as a cell phone, personal digital assistant, or laptop whose software enables network roaming capabilities.
- The **Home Agent** is a router on the home network serving as the anchor point for communication with the Mobile Node; it tunnels packets from a device on the Internet, called a Correspondent Node, to the roaming Mobile Node. (A tunnel is established between the Home Agent and a reachable point for the Mobile Node in the foreign network.)
- The **Foreign Agent** is a router that may function as the point of attachment for the Mobile Node when it roams to a foreign network, delivering packets from the Home Agent to the Mobile Node.
- The **care-of address(COA)** is the termination point of the tunnel toward the Mobile Node when it is on a foreign network. The Home Agent maintains an association between the home IP address of the Mobile Node and its care-of address, which is the current location of the Mobile Node on the foreign or visited network

Working of Mobile IP

The working of Mobile IP can be described in 3 phases:

- **Agent Discovery**

In the Agent Discovery phase, the mobile nodes discover their Foreign and Home Agents. The Home Agent and Foreign Agent advertise their services on the network using the ICMP Router Discovery Protocol (IRDP).

- **Registration**

The registration phase is responsible for informing the current location of the home agent and foreign agent for the correct forwarding of packets.

- **Tunneling**

This phase is used to establish a virtual connection as a pipe for moving the data packets between a tunnel entry and a tunnel endpoint.

Transport Protocol Design Issues

The following elements must to be taken into account while designing transport protocols for WSNs:

- **Reliability**: Since the majority of the data come from the sensor nodes to the sink, there may be congestion around the sink.Although the MAC protocol can recover packets that have been lost due to bit errors, it is unable to handle packets that have been lost due to buffer overflow. A packet loss recovery mechanism, similar to the ACK and selective ACK used in TCP, is required for WSNs.
- **Efficient use of network resources:** The majority of WSN applications are reactive, which means they observe passively and hold data until an event occurs before delivering it to the sink. Due to an occurrence, these programs might only need to send a few packets.

Transport Protocol Design Issues

- **Adaptability**: A transport protocol should, if at all possible, be created with cross layer optimization in mind. For instance, if a routing algorithm alerts the transport protocol of a route failure, the protocol can infer that packet loss is not due to congestion but rather to the failure of the route rather than congestion. In this scenario, the sender is free to stick with its current rate
- **Compatibility**: Fairness for a range of sensor nodes should be ensured by the transport control protocols

Examples of Existing Transport Control Protocols for WSN

Most examples can be grouped in one of the four groups:

- Upstream congestion control,(Eg. CODA)
- Downstream congestion control, (Eg. PSFQ)
- Upstream reliability guarantee, (Eg. ESRT)
- Downstream reliability guarantee. (Eg. GARUDA)

Examples of TCP for WSN:

1. CODA (Congestion Detection and Avoidance)
2. ESRT (Event-to-Sink Reliable Transport)
3. PSFQ (Pump Slowly, Fetch Quickly)
4. GARUDA

1. **CODA (Congestion Detection and Avoidance)**

- Event-driven sensor networks operate under an idle or light load and then suddenly become active in response to a detected or monitored event.
- It is during these periods of event impulses that the likelihood of congestion is greatest and the information in transit of most importance to users.
- To address the challenge an energy efficient congestion control scheme for sensor networks called CODA (COngestion Detection and Avoidance) is proposed that comprises three mechanisms:

(i) receiver-based congestion detection;

(ii) open-loop hop-by-hop backpressure;

(iii) closed-loop multi-source regulation

1. **CODA (Congestion Detection and Avoidance)**

- CODA, detects the congestion by observing the buffer size of sensor nodes and the load of the wireless channel.
- If these two characteristic exceed from a pre-defined threshold, using an open-loop hop-by-hop backpressure, the node that has detected congestion will then alert its upstream neighbor to lower its rate
- CODA regulates a multisource rate through a closed-loop end-to-end approach, as follows:

(1) When a sensor node exceeds its theoretical rate, it sets a “regulation” bit in the “event” packet;

(2) If the event packet received by the sink has a “regulation” bit set, the sink sends an ACK message to the sensor nodes and informs them to reduce their rate;

1. CODA (Congestion Detection and Avoidance)

(3) if the congestion is cleared, the sink will send an immediate ACK control message to the sensor nodes, informing them that they can increase their rate.

- **Disadvantages**
  - unidirectional control, only from the sensors to the sink
  - no reliability consideration;
  - The response time of its closed-loop multisource control increases under heavy congestion since the ACK issued from the sink will probably be lost

Examples of Existing Transport Control Protocols for WSN

2. ESRT (Event-to-Sink Reliable Transport)

- ESRT is a novel transport solution developed to achieve reliable event detection in WSN with minimum energy expenditure.
- It includes a congestion control component that serves the dual purpose of achieving reliability and conserving energy.
- It periodically computes a reliability figure (r), representing the rate of packets received successfully in a given time interval.
- ESRT then deduces the required sensor reporting frequency (f) from the reliability figure (r) using an expression such as f = G(r).
- Finally, ESRT informs all sensors of the values of (f) through an assumed channel with high power.

2. ESRT (Event-to-Sink Reliable Transport)

- ESRT uses an end-to-end approach to guarantee a desired reliability figure through adjusting the sensors' reporting frequency.
- It provides overall reliability for the application. The additional benefit of ESRT is energy conservation through control of reporting frequency.
- ESRT addresses event reliability but not congestion control explicitly.
- **Disadvantages of ESRT**

- advertises the same reporting frequency to all sensors (since different nodes may have contributed differently to congestion, applying different frequencies would be more appropriate)
- considers mainly reliability and energy conservation as performance measures.

Examples of Existing Transport Control Protocols for WSN

3. PSFQ (Pump Slowly, Fetch Quickly)

- By pacing data at a reasonably moderate rate and allowing sensor nodes that suffer from data loss to recover any missing segments from close neighbors, PSFQ distributes data from sink to sensors.
- The goal is to localize data recovery among close neighbors in order to minimize loss recovery and achieve loose delay bounds.
- Pump, fetch, and report are the three processes that make up PSFQ.
- Until all of the data fragments have been sent, Sink broadcasts a packet to its neighbors every T time units.
- The sensor node enters fetch mode when a sequence number gap is found and sends a NACK in the reverse path to retrieve the lost fragment

3. PSFQ (Pump Slowly, Fetch Quickly)

- Lastly, using a straightforward and scalable hop-by-hop report method, the sink can request information from sensors regarding the status of data delivery.
- **Disadvantages:**

- Its slow pump causes a significant delay
- It cannot detect packet loss for single packet transmission
- Its hop-by-hop recovery with cache requires bigger buffer sizes.

Examples of Existing Transport Control Protocols for WSN

4. GARUDA

- Garuda is a transport control protocol designed specifically for wireless sensor networks (WSNs).
- Garuda aims to provide reliable and efficient data transfer in low-power and lossy networks, while minimizing energy consumption and network overhead.
- It is built on a two-tier node design, and core sensor nodes are chosen from nodes that are 3i hops away from the sink (i is an integer). Second-tier nodes are the noncore nodes that are still present.
- A nearby core node is selected by each noncore sensor node to serve as its core node. Core nodes are used by noncore nodes to recover lost packets

4. GARUDA

- GARUDA uses two stage loss recovery processes. This is done with the help of pulsing-based approach(Wait for first packet- WFP).
- In first stage, packet recovery is done by the core nodes. When a packet arrives at central node which is not according the sequence, it notifies to a central node in upstream direction that some packet is missing.
- Another is called non core recovery phase in which non central nodes requests to transmit packets again from core nodes.
- The WFP signal serves two purposes apart from aiding in loss recovery:

  1) It allows the sink to inform the sensors about an impending message that has reliability requirements, and

  2) It enables sensors to request for retransmissions when they do not receive the first packet successfully

Performance of Transport Control Protocols: Congestion & packet loss

Energy consumption, which is calculated for end-to-end and hop-by-hop situations, is the parameter used to compare congestion. Loss performance is a different metric that is dependent on cache and nocache methods.

1. **Congestion**

- End to end and hop by hop are two common methods for reducing congestion.
- The source node must identify congestion in either the receiver-assisted (ACK-based loss detection) mode or the network assisted mode in an end-to-end protocol like standard TCP (using explicit congestion notification)
- Rate modifications therefore only take place at the source node.

- In hop-by-hop congestion control, intermediate nodes alert the originating connection node when there is congestion.
- Hop-by-hop control may be able to clear congestion more quickly than the end-to-end method while also lowering packet loss and energy usage in sensor nodes
- End-to-end mechanisms energy efficiency depends on the path length, but hop-by-hop control is independent of the path length and yields a better efficiency ratio.
- The ratio of all packets discarded in the sensor network to all packets received at the sink for hop-by-hop congestion control is the energy tax, according to CODA.

2. Packet loss recovery

- Cache and non cache recovery are typically the two approaches that are available for packet loss recovery
- A similar end-to-end ARQ (automatic repeat request) to the conventional TCP is non cache recovery.
- Using a hop-by-hop methodology, cache-based recovery relies on retransmissions between two nearby nodes and caching at the intermediate nodes.
- Therefore, retransmissions may happen in h hops in the non cache situation, necessitating greater overall energy.
- The node that replicates transmitted packets locally for a predetermined amount of time is referred to as the cache point, while the node that packets are dropped due to congestion is referred to as the loss point.

2. Packet loss recovery

- The length of the retransmission path, or lp, will be defined as the quantity of hops between the caching node and the node where the loss occurs.
- lp = h1, where h1 is the number of hops from the loss point to the source node, in the non cache case.
- If lost packets are located on nearby nodes in the cache example, lp can be 1.
- Packet copies can only be stored for a finite amount of time because sensor nodes have a finite amount of buffer space. Because of this, lp in the cache scenario may be greater than 1 but lower than h1(1<lp<h1).
- By using cache-based recovery, each packet is kept at each intermediate node it passes through until it is successfully received by its neighboring node or until a timeout occurs (whichever is sooner)

2. Packet loss recovery

- Distributing caching is a different technique that would distribute packet copies among intermediate nodes.
- Only one or a few intermediate nodes store each packet.In addition to using less buffer space than traditional caching, distributed caching may have a longer lp than traditional caching (but still be smaller than in the non cache situation).

INTRODUCTION, WIRELESS TRANSMISSION AND MEDIUM ACCESS CONTROL

UNIT 3

Introduction

- User mobility and device portability are two different types of mobility.
- The term "user mobility" describes a user who has access to the same or equivalent telecommunication services at various locations; in other words, the user is mobile and the services follow them.
- Simple call-forwarding solutions from the telephone or computer desktops that enable roaming are examples of systems that support user mobility.
- Device portability - When a communication device is portable, it can be moved (with or without a user).
- The mobile phone system is a common illustration of a system that supports device portability, as the system automatically switches the device from one radio transmitter (also known as a base station) to another if the signal deteriorates.

Introduction

- The word "wireless" is applied to gadgets. This only explains how to connect to a network or other communication partners without using a wire.
- Transmission of electromagnetic waves through "the air" takes the role of the cable (although wireless transmission does not need any medium).

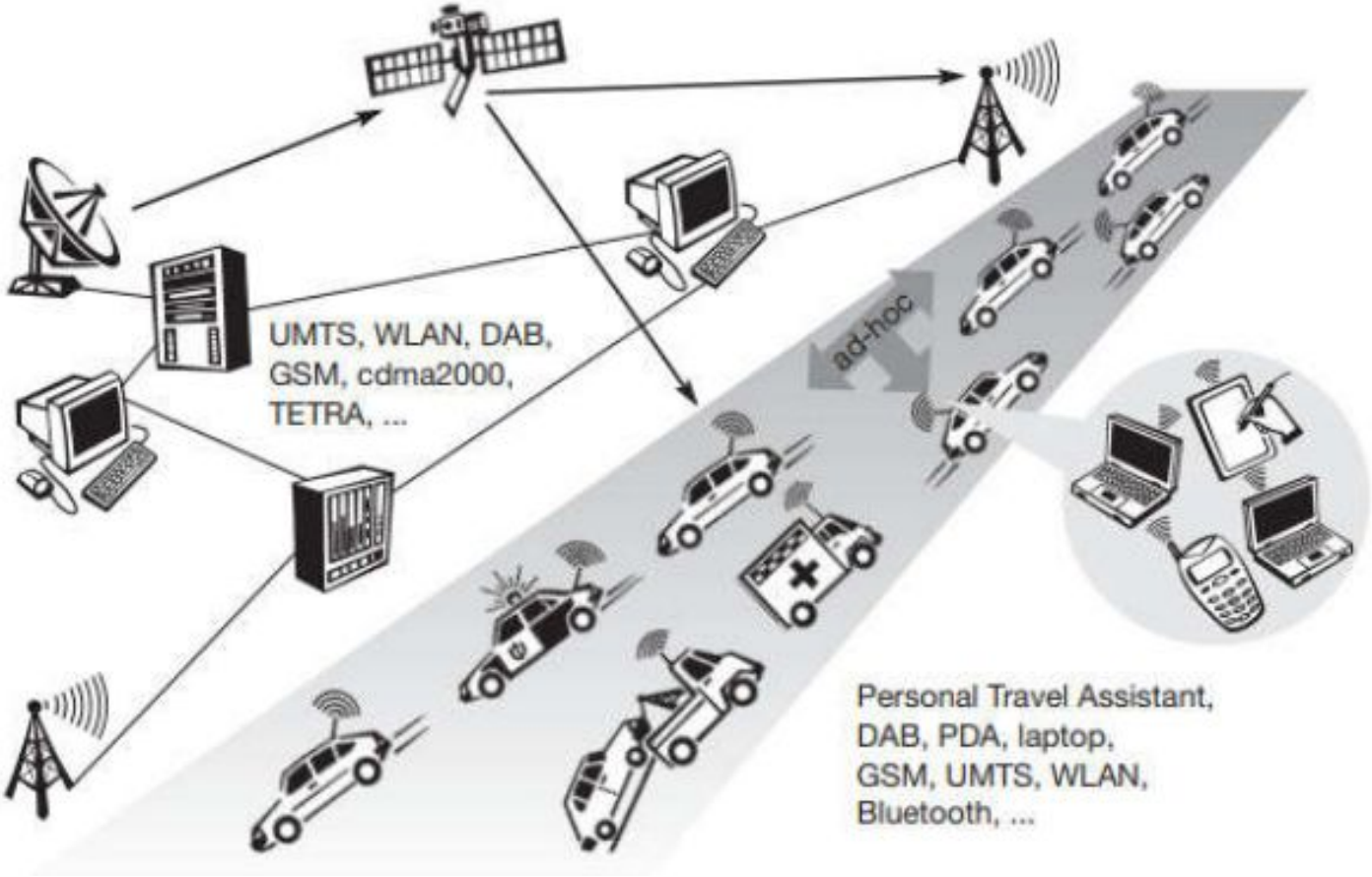
Applications of wireless networks and mobile communications

1. **Vehicles**

- While some already exist in today's cars, there will be many more wireless communication systems and mobility-aware applications in cars of the future.
- Digital audio broadcasting (DAB) with 1.5 Mbit/s allows for the reception of music, news, traffic updates, weather forecasts, and other broadcast information.
- A universal mobile telecommunications system (UMTS) phone with 384 kbit/s voice and data connectivity might be available for personal communication.
- Satellite communication can be employed in remote places, and the global positioning system is used to identify the car's current location. (GPS)
- To transmit information quickly in an emergency or to preserve a safe distance from one another, nearby cars create a small ad hoc network.

- In the event of an accident, not only will the airbag deploy, but a provider will also receive an emergency call alerting the police and ambulance service
- Vehicles will communicate with one another via an ad-hoc network in order to alert them about accidents and help them slowdown in time, even before a driver is aware of one.
- A typical setup for mobile communications including numerous wireless devices is shown in Figure 1. Mobile phone networks (GSM, UMTS) and trunked radio systems (TETRA) will connect to networks with a fixed infrastructure to form wireless LANs. (WLAN).
- Personal digital assistants (PDA), computers, and mobile phones, such as those connected via Bluetooth, can all be a part of wireless pico networks within a car.

Figure 1.
A typical application of mobile communications: road traffic



2. Infotainment and more

- Internet everywhere? Not if wireless networks are absent! Think about a city's travel guide. Static data can be downloaded from a CD-ROM, DVD,or even the Internet at home.
- However, wireless networks can deliver current information at any suitable location. By determining your location via GPS, a local base station, or triangulation, the tour guide may provide you with information on a building's past while simultaneously downloading details about a concert that will be taking place there that same night over a local wireless network
- You can select a seat, pay using electronic money, and email these details to a service provider.
- In order to enable, for instance, ad-hoc gaming networks as soon as players meet to play together, entertainment and games are a rising area of wireless network applications.

2. Infotainment and more

- Internet everywhere? Not if wireless networks are absent! Think about a city's travel guide. Static data can be downloaded from a CD-ROM, DVD,or even the Internet at home.
- However, wireless networks can deliver current information at any suitable location. By determining your location via GPS, a local base station, or triangulation, the tour guide may provide you with information on a building's past while simultaneously downloading details about a concert that will be taking place there that same night over a local wireless network
- You can select a seat, pay using electronic money, and email these details to a service provider.
- In order to enable, for instance, ad-hoc gaming networks as soon as players meet to play together, entertainment and games are a rising area of wireless network applications.

3. Business

- Today's travelling salesperson requires immediate access to the company's database to make sure that the files on his or her laptop represent the current situation, to allow the business to monitor all of its travelling employees' activities, to maintain consistent databases, etc.
- The laptop can become a truly mobile office with wireless connection, but effective and strong synchronization techniques are required to guarantee data consistency.
- Mobile communications should always provide the best access to the internet, the company's intranet, or the telephone network, regardless of the time and location.

4. Replacement of wired networks

- In some circumstances, such as with remote sensors, at trade exhibitions, or in old buildings, wireless networks can also be used in place of wired networks.
- It is frequently impractical to link remote sensors for weather forecasts, earthquake detection, or to give environmental data due to financial considerations. In this case, wireless connections, like those provided by satellite, can be useful
- Computers, sensors, or information displays in historic structures are additional uses for wireless networks since excessive cabling could damage priceless walls or flooring. The use of wireless access points in a room corner may be the answer.

Wireless Transmission : Frequency for radio transmission

- Numerous frequency bands are available for radio transmission. There are pros and downsides to each frequency band.
- The frequency range that can be employed for data transmission is depicted in rough detail in Figure 3.In the graph, frequencies from 300 Hz to over 300 THz are depicted.
- The wavelength λ is directly connected to the frequency by the following equation:
  λ= c/f
  where c ≅ 3·108 m/s (the speed of light in vacuum) and f the frequency.

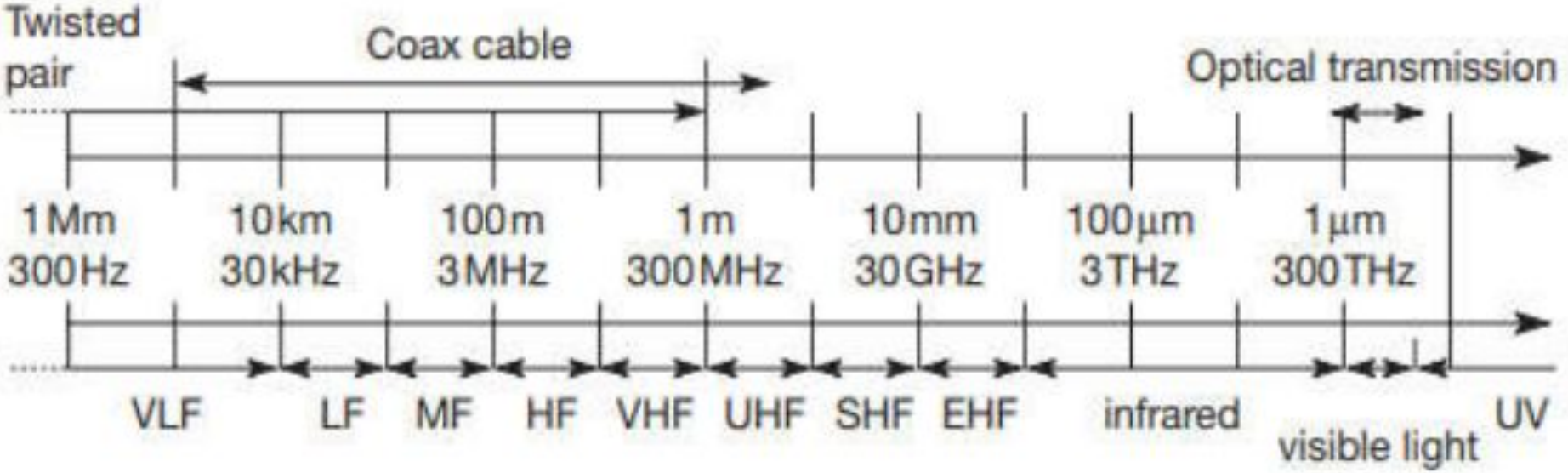


Figure 3: Frequency spectrum

Traditional Wired Networks

- For traditional wired networks, frequencies of up to several hundred kHz are used for distances up to some km with twisted pair copper wires.
- frequencies of several hundred MHz are used with coaxial cable.
- Fiber optics are used for frequency ranges of several hundred THz.

Very Low Frequency(VLF)

- These are very long waves.
- Waves in the **low frequency (LF)** range are used by submarines, because they can penetrate water and can follow the earth's surface.
- Some radio stations still use these frequencies, e.g., between 148.5 kHz and 283.5 kHz in Germany.

medium frequency (MF) and **high frequency (HF)**

- The **medium frequency (MF)** and **high frequency (HF)** ranges are typical for transmission of hundreds of radio stations either as amplitude modulation **(AM)** between 520 kHz and 1605.5 kHz.

very high frequency (VHF) and ultra high frequency (UHF) bands

- As we move to higher frequencies, the TV stations follow. Conventional analog TV is transmitted in ranges of 174–230 MHz and 470–790 MHz using the very high frequency **(VHF)** and ultra high frequency **(UHF)** bands.
- In this range, digital audio broadcasting (DAB) takes place as well .
- UHF is also used for mobile phones with analog technology (450–465 MHz)
- the digital GSM (890–960 MHz, 1710–1880 MHz), digital cordless telephones following the DECT standard (1880–1900 MHz), 3G cellular systems following the UMTS standard (1900–1980 MHz, 2020–2025 MHz, 2110–2190 MHz) and many more.
- VHF and especially UHF allow for small antennas and relatively reliable connections for mobile telephony.

Super high frequencies (SHF)

- **Super high frequencies (SHF)** are typically used for directed microwave links (approx. 2–40 GHz) and fixed satellite services in the C-band (4 and 6 GHz), Ku-band (11 and 14 GHz), or Ka-band (19 and 29 GHz).
- Some systems are planned in the **extremely high frequency (EHF)** range which comes close to infra red. All radio frequencies are regulated to avoid interference, e.g., the German regulation covers 9 kHz–275 GHz.

- The next step into higher frequencies involves optical transmission, which is not only used for fiber optical links but also for wireless communications.
- **Infra red (IR)** transmission is used for directed links, e.g., to connect different buildings via laser links. The most widespread IR technology, infra red data association (IrDA), uses wavelengths of approximately 850–900 nm to connect laptops, PDAs etc.
- Finally, visible light has been used for wireless transmission for thousands of years. While light is not very reliable due to interference, but it is nevertheless useful due to built-in human receivers.

Signals

- Data is physically represented by signals.
- Users of a communication system exchange data through the transmission of signals
- Data, or bits, are converted into signals and back again via Layer 1 of the ISO/OSI basic reference model.
- Signals are functions of time and location. Signal parameters represent the data values.
- The most interesting types of signals for radio transmission are periodic signals, especially sine waves as carriers.
- The general function of a sine wave is:
  $g(t) = At \sin(2 \pi ft t + \varphi t)$
  Signal parameters are the amplitude A, the frequency f, and the phase shift φ

Signal Representation

- Time domain representation of a signal
- Frequency domain representation of a Signal
- Phase domain representation of a signal

Time Domain Representation of signal

- A typical way to represent signals is the time domain see Figure 2.2).
- Here the amplitude A of a signal is shown versus time (time is mostly measured in seconds s, amplitudes can be measured in, e.g., volt V).
- This is also the typical representation known from an oscilloscope. A phase shift can also be shown in this representation.
- Representations in the time domain are problematic if a signal consists of many different frequencies . In this case, a better representation of a signal is the frequency domain (see Figure 2.3)

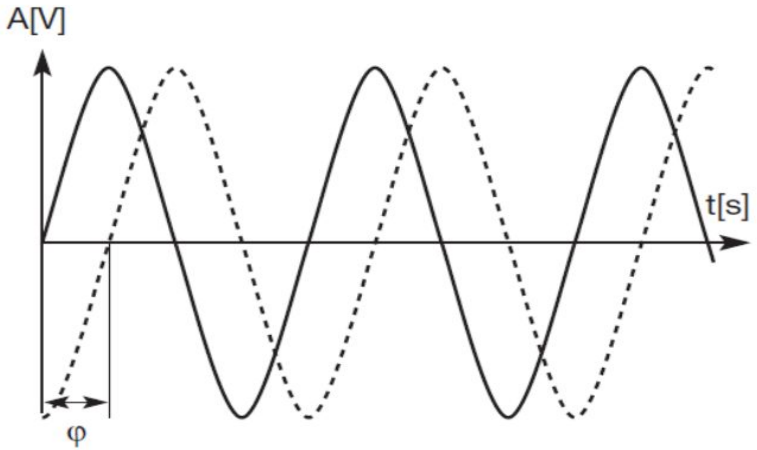


Figure 2.2
Time domain representation of a signal

Frequency Domain Representation of signal

- Here the amplitude of a certain frequency part of the signal is shown versus the frequency.
- Figure 2.3 only shows one peak and the signal consists only of a single frequency part.
- Arbitrary periodic functions would have many peaks, known as the frequency spectrum of a signal. A tool to display frequencies is a spectrum analyzer.

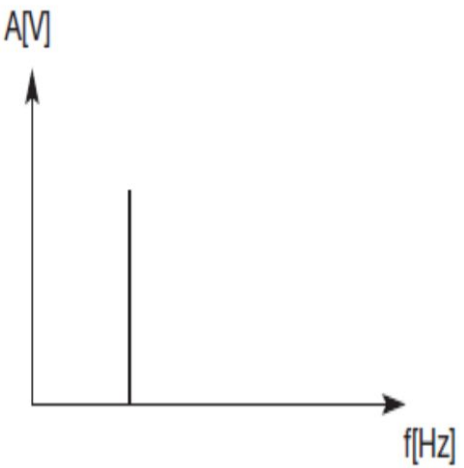
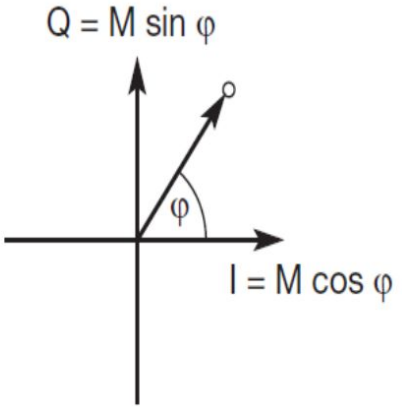


Figure 2.3
Frequency domain representation of a signal

Phase Domain Representation of signal

- A third way to represent signals is the phase domain shown in Figure 2.4.
- This representation, also called phase state or signal constellation diagram, shows the amplitude M of a signal and its phase φ in polar coordinates. (The length of the vector represents the amplitude, the angle the phase shift.)
- The x-axis represents a phase of 0 and is also called In-Phase (I).
- A phase shift of 90° or π/2 would be a point on the y-axis, called Quadrature (Q).

Figure 2.4
Phase domain representation of a signal



Antennas

- As the name wireless already indicates, this communication mode involves ‘getting rid’ of wires and transmitting signals through space without guidance.
- We do not need any ‘medium’ (such as an ether) for the transport of electromagnetic waves.
- Somehow, we have to couple the energy from the transmitter to the out-side world and, in reverse, from the outside world to the receiver.
- This is exactly what antennas do. Antennas transmit and receive electromagnetic radiation from space through a wire or coaxial cable(or any other appropriate conductor)

Types of Antennas

1. **Isotropic radiator:**

➔ An isotropic antenna is defined as a hypothetical antenna having the same radiation in all directions
➔ But in reality, such an antenna does not exist.
➔ Real antennas all have directional effects, which means that the radiation intensity varies depending on the direction the antenna is facing.

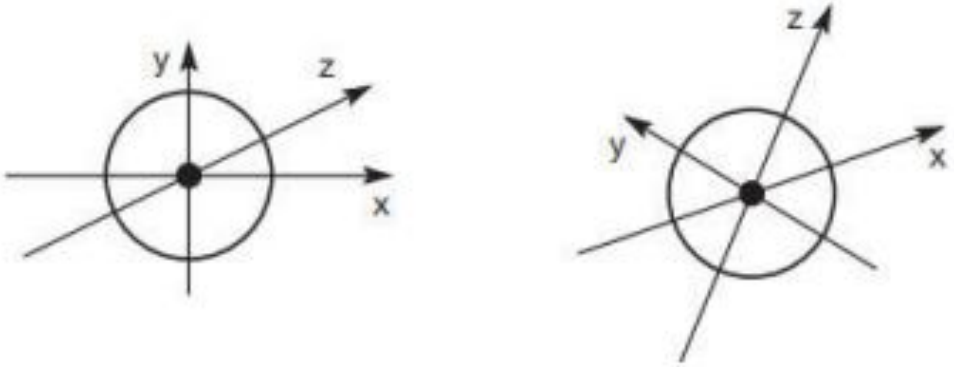


Figure 7: Radiation pattern of an isotropic radiator

2.Center-fed dipole

- A dipole an antenna with a center-fed driven element for transmitting or receiving radio frequency energy.
- A thin, center-fed dipole, also known as a Hertzian dipole, is the most basic actual antenna and is depicted
- The dipole consists of two collinear conductors of equal length, separated by a small feeding gap.
- The length of the dipole is not arbitrary, but, for example, half the wavelength λ of the signal to transmit results in a very efficient radiation of the energy.
- If mounted on the roof of a car, the length of λ/4 is efficient. This is also known as Marconi antenna.

Figure 8: Simple antennas

2.Center-fed dipole

- A λ/2 dipole has a uniform or omnidirectional radiation pattern in one plane and a figure eight pattern in the other two planes as shown in Figure

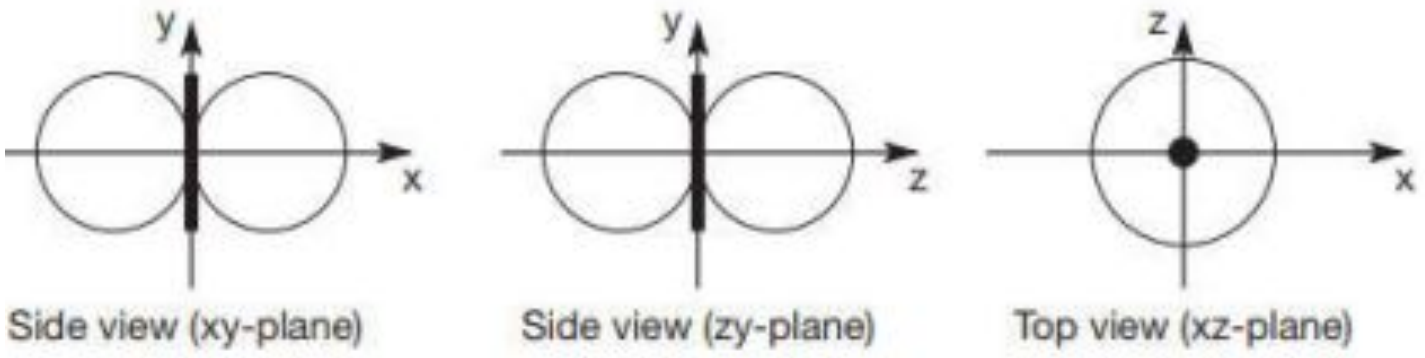
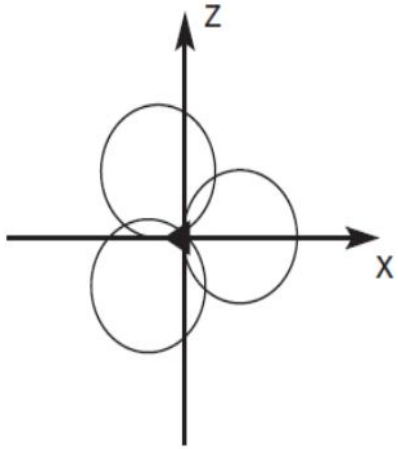


Figure 9: Radiation pattern of a simple dipole

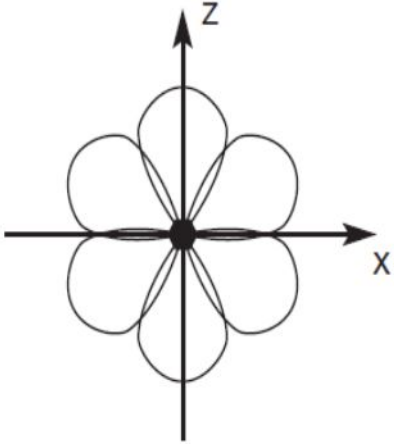
- The only way this kind of antenna can overcome environmental obstacles is by increasing the signal's power. Challenges may include hills, valleys, structures, etc.

2.Center-fed dipole

- An omnidirectional radiation pattern is not very useful if an antenna is installed, for instance, in a valley or between two structures.
- Directional antennas with predetermined set favored transmission and reception directions can be employed in this situation.
- Satellite dishes are an exceptional type of directional antenna.
- Directed antennas are typically applied in cellular systems. Several directed antennas can be combined on a single pole to construct a sectorized antenna.
- A cell can be sectorized into, for example, three or six sectors, thus enabling frequency reuse



Top view, 3 sector



Top view, 6 sector

2.Center-fed dipole

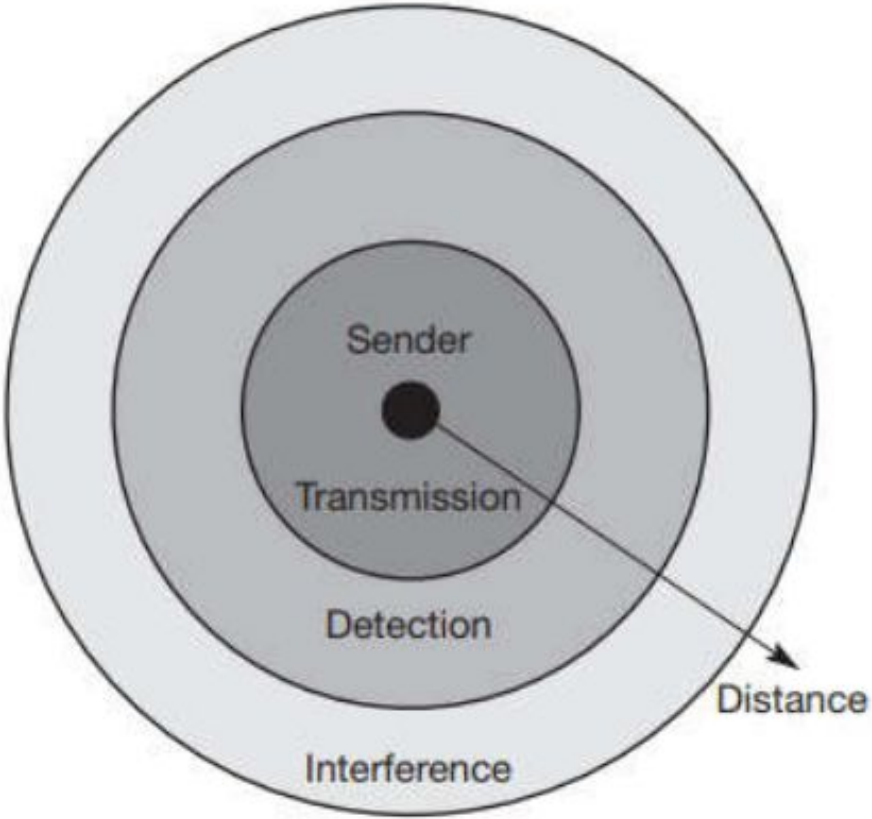
- A more advanced solution is provided by smart antennas which combine multiple antenna elements (also called antenna array) with signal processing to optimize the radiation/reception pattern in response to the signal environment.
- These antennas can adapt to changes in reception power, transmission conditions and many signal propagation effects
- Antenna arrays can also be used for beam forming.

(Beamforming is a kind of radio frequency (RF) management in which an access point makes use of various antennas to transmit the exact same signal.)

Signal propagation

- Wireless communication networks have signal senders and receivers much like wired networks do.
- In contrast to wired networks, where a signal may only travel via a wire, wireless networks do not have a wire for the signal to use to establish the direction of propagation
- The wire usually displays the same traits at each place as long as it is not cut or damaged.
- Depending on the length, one can exactly predict how a signal will behave while passing over this wire, for example, received power.

- For wireless transmission, this predictable behavior is only valid in a vacuum, i.e., without matter between the sender and the receiver.



Ranges for transmission, detection, and interference of signals

- **Transmission range:** Within a certain radius of the sender transmission is possible, i.e., a receiver receives the signals with an error rate low enough to be able to communicate and can also act as sender.
- **Detection range:** Within a second radius, detection of the transmission is possible, i.e., the transmitted power is large enough to differ from background noise. However, the error rate is too high to establish communication.
- **Interference range:** Within a third even larger radius, the sender may interfere with other transmission by adding to the background noise. A receiver will not be able to detect the signals, but the signals may disturb other signals.
- However, real life does not happen in a vacuum, radio transmission has to contend with our atmosphere, mountains, buildings, moving senders and receivers etc.
- In reality, the three circles referred to above will be bizarrely-shaped polygons with their shape being time and frequency dependent.

Problems in signal propagation wireless

1. Path loss of radio signals/Attenuation
2. Additional signal propagation effects
3. Multi-path propagation

Path loss of radio signals/Attenuation

- **Attenuation** is a general term that refers to any reduction in the strength of a signal.
- Sometimes called loss, attenuation is a natural consequence of signal transmission over long distances.
- In free space radio signals propagate as light does (independently of their frequency), i.e., they follow a straight line (besides gravitational effects).
- If such a straight line exists between a sender and a receiver it is called **line-of-sight (LOS)**.
- Even if no matter exists between the sender and the receiver (i.e., if there is a vacuum), the signal still experiences the free space loss.
- The received power *Pr* is proportional to $1/d^2$ with d being the distance between sender and receiver **(inverse square law)**.

Path loss of radio signals/Attenuation

- As soon as there is any matter between sender and receiver, the situation becomes more complex. Most radio transmission takes place through the atmosphere – signals travel through air, rain, snow, fog, dust-particles, smog etc.
- While the path loss or attenuation does not cause too much trouble for short distances, e.g., for LANs, the atmosphere heavily influences transmission over long distances, e.g., satellite transmission.
- Even mobile phone systems are influenced by weather conditions such as heavy rain. Rain can absorb much of the radiated energy of the antenna (this effect is used in a microwave oven to cook), so communication links may break down as soon as the rain sets in.

Radio waves can exhibit three fundamental propagation behaviors depending on their frequency:

- **Ground wave** (<2 MHz): Waves with low frequencies follow the earth's surface and can propagate long distances. These waves are used for, e.g., submarine communication or AM radio.
- **Sky wave** (2–30 MHz): Many international broadcasts and amateur radio use these short waves that are reflected2 at the ionosphere. This way the waves can bounce back and forth between the ionosphere and the earth's surface, travelling around the world.
- **Line-of-sight** (>30 MHz): Mobile phone systems, satellite systems cordless telephones etc. use even higher frequencies. The emitted waves follow a (more or less) straight line of sight. This enables direct communication with satellites (no reflection at the ionosphere) or microwave links on the ground. However, an additional consideration for ground-based communication is that the waves are bent by the atmosphere due to refraction

Additional signal propagation effects

- An extreme form of attenuation is **blocking or shadowing** of radio signals due to large obstacles
- The higher the frequency of a signal, the more it behaves like light. Even small obstacles like a simple wall, a truck on the street, or trees in an alley may block the signal.
- Another effect is the **reflection** of signals.If an object is large compared to the wavelength of the signal, e.g., huge buildings, mountains, or the surface of the earth, the signal is reflected.
- The reflected signal is not as strong as the original, as objects can absorb some of the signal's power.
- Signals transmitted from a sender may bounce off the walls of buildings several times before they reach the receiver. The more often the signal is reflected, the weaker it becomes.

Additional signal propagation effects

- Another effect is the **refraction** of signals. This effect occurs because the velocity of the electromagnetic waves depends on the density of the medium through which it travels.
- Only in vacuum does it equal c. As the figure shows,waves that travel into a denser medium are bent towards the medium.
- This is the reason for LOS radio waves being bent towards the earth. The density of the atmosphere is higher closer to the ground.
- While shadowing and reflection are caused by objects much larger than the wavelength of the signals If the size of an obstacle is in the order of the wavelength or less, then waves can be **scattered** (see Figure 2.13, left side). An incoming signal is scattered into several weaker outgoing signals.

Figure 2.12

Blocking (shadowing), reflection and refraction of waves

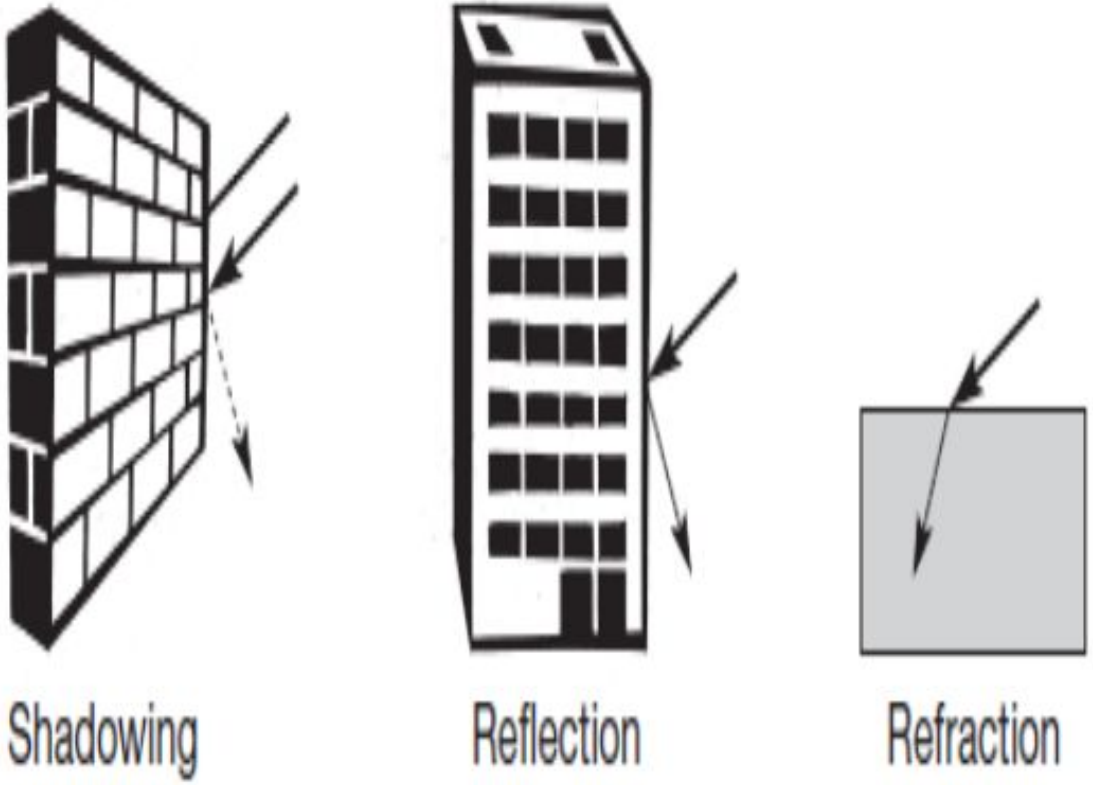
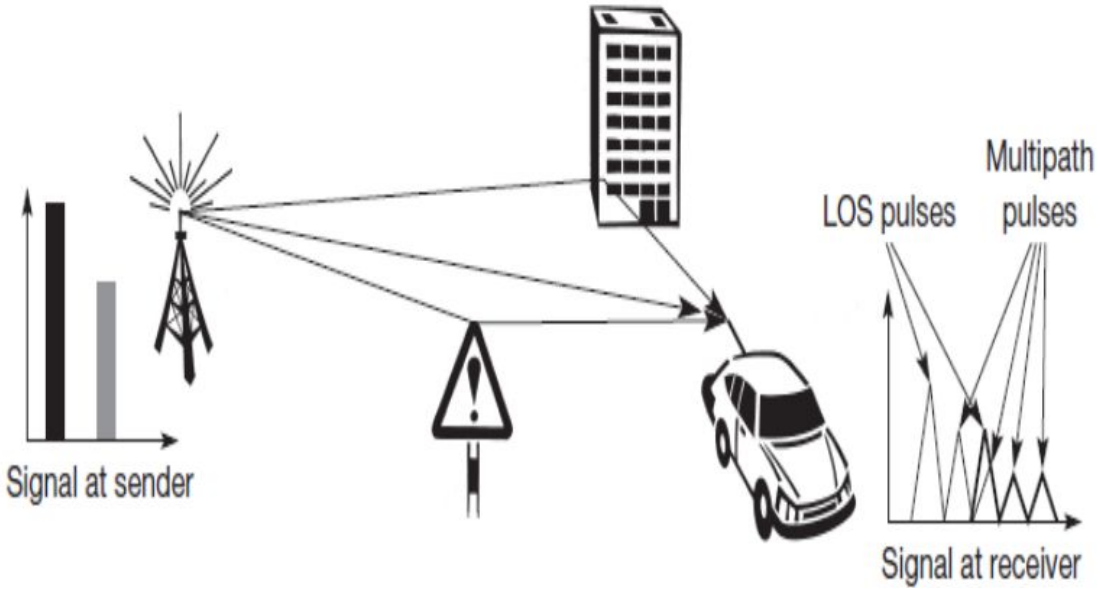


Figure 2.13
Scattering and diffraction of waves

Multi-path propagation

- Radio waves emitted by the sender can either travel along a straight line, or they may be reflected at a large building, or scattered at smaller obstacles. This simplified figure only shows three possible paths for the signal.



MULTIPLEXING

- Multiplexing describes how several users can share a medium with minimum or no interference.
- One example, is highways with several lanes. Many users (car drivers) use the same medium (the highways) with hopefully no interference (i.e., accidents).
- This is possible due to the provision of several lanes (space division multiplexing) separating the traffic. In addition, different cars use the same medium (i.e., the same lane) at different points in time (time division multiplexing).

- There are numerous options for granting access to the channel. These primarily comprise the following:

1. Frequency Division Multiple Access (FDMA) **(Refer unit 2)**
2. Time Division Multiple Access (TDMA) **(Refer unit 2)**
3. Code Division Multiple Access (CDMA) **(Refer unit 2)**
4. Space Division Multiple Access (SDMA)

Space Division Multiple Access (SDMA)

- Space division multiple access, also known as spatial division multiple access, is a MIMO (multiple-input multiple-output) architecture technique that is commonly used in wireless and satellite communication.
- It has the following characteristics:

1. Using the same channel, all users can communicate at the same time.
2. SDMA is fully interference-free.
3. A single satellite can communicate with many satellites using the same
4. frequency receiver.
5. The base station in SDMA can monitor a moving user thanks to the use of directional spot-beam antennas.
6. Controls the amount of energy radiated by each user in space.

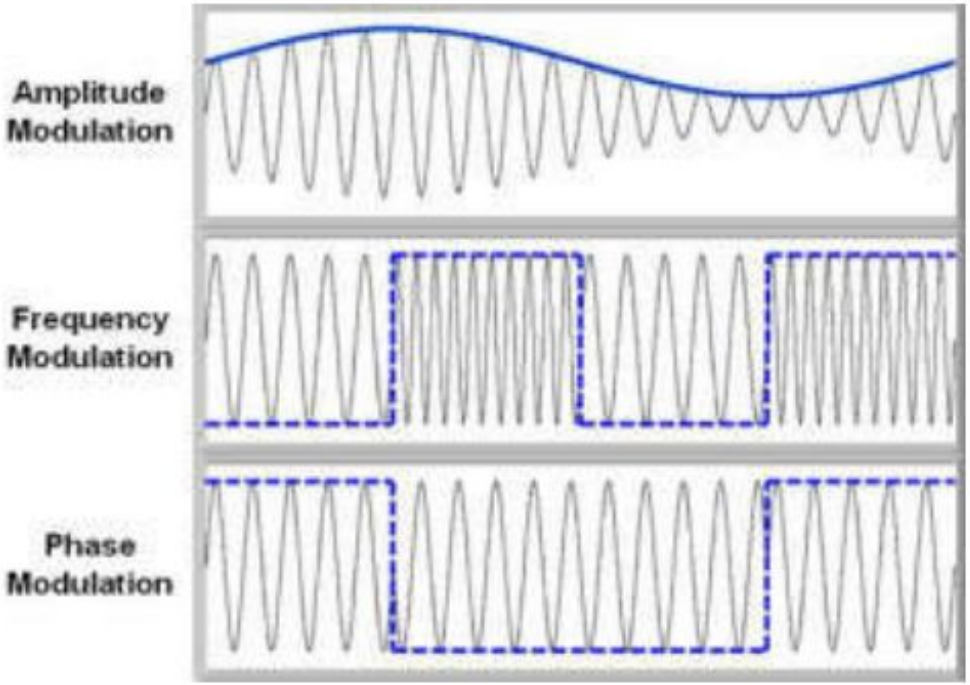
- Space Division Multiple Access (SDMA) is used for allocating a **separated space** to users in wireless networks.
- A typical application involves assigning an optimal base station to a mobile phone user.
- The mobile phone may receive several base stations with different quality.
- A MAC algorithm could now decide which base station is best, taking into account which frequencies (FDM), time slots (TDM) or code (CDM) are still available (depending on the technology).
- Typically, SDMA is never used in isolation but always in combination with one or more other schemes.
- The basis for the SDMA algorithm is formed by **cells** and **sectorized** antennas which constitute the infrastructure implementing space division multiplexing (SDM).

MODULATION

- Modulation is a process of changing the characteristics of the wave to be transmitted by superimposing the message signal on the high-frequency signal
- In this process video, voice and other data signals modify high frequency signals – also known as the carrier wave
- These modulation techniques are classified into two major types: analog and digital or pulse modulation.
- In both the techniques, the baseband information is converted to Radio Frequency signals, but in analog modulation, these RF communication signals are a continuous range of values, whereas in digital modulation these are prearranged discrete states.

- Types of analog modulation are:

1. **Amplitude modulation (AM):** the amplitude of the carrier wave is varied in proportion to the message signal, and the other factors like frequency and phase remain constant.
2. **Frequency modulation (FM):** Frequency modulation (FM) varies the frequency of the carrier in proportion to the message or data signal while maintaining other parameters constant.The advantage of FM over AM is the greater suppression of noise at the expense of bandwidth in FM.
3. **Phase modulation (PM):** In phase modulation, the carrier phase is varied in accordance with the data signal. In this type of modulation, when the phase is changed it also affects the frequency, so this modulation also comes under frequency modulation.



Types of Analog Modulation

Analog modulation (AM, FM, and PM) is more sensitive to noise. If noise enters into a system, it persists and gets carried till the end receiver. Therefore, this drawback can be overcome by the digital modulation technique.

Digital Modulation

- In digital modulation, a message signal is converted from analog to digital message and then modulated by using a carrier wave.
- The carrier wave is keyed or switched on and off to create pulses such that the signal is modulated. Similar to the analog, here the parameters like amplitude, frequency, and phase variation of the carrier wave decides the type of digital modulation.
- The types of digital modulation are based on the type of signal and application used such as Amplitude Shift Keying, Frequency Shift Keying, Phase Shift Keying

CELLULAR SYSTEMS

- SDM is used in cellular networks for mobile communications.
- Each transmitter, commonly referred to as a base station, serves a specific cell.
- Cell radii can be as small as a few metres in a building, hundreds of metres in a city, or even tens of kilometres across the nation.
- Cells are never perfectly round or hexagonal; instead, the shape of a cell depends on its surroundings (buildings, mountains, valleys, etc.), the weather, and occasionally even the strain on the system.
- This strategy is frequently used in mobile telecommunication systems, where a mobile station inside a base station's cell can connect with that base station and vice versa.

Why mobile network operators do not employ strong transmitters with large cells like, for example, radio stations do, instead of installing thousands of base stations across a nation (which is rather expensive)?

Advantages of Small cell cellular systems

1. **Higher capacity**

Smaller the size of the cell more the number of concurrent users i.e. huge cells do not allow for more concurrent users.

2. **Less transmission power**

Huge cells require a greater transmission power than small cells.

3. **Local interference only**

For huge cells there are a number of interfering signals, while for small cells there is limited interference only.

4. **Robustness**

As cellular systems are decentralized, they are more robust against the failure of single components.

Disadvantages of Small cell cellular systems

1. Infrastructure needed

Small cells require a complex infrastructure to connect all base station. The infrastructure required includes switches for call forwarding, location registers etc.

2. Handover needed

The mobile station has to perform a handover when changing from one cell to another very frequently.

3. Frequency planning

To avoid interference, frequency spectrum should be distributed properly with a very less range of frequency spectrum.

Telecommunication, Satellite and Broadcast Systems

UNIT 3

GSM

Global System for Mobile Communications

- GSM is the most successful digital mobile telecommunication system in the world today.
- It is used for transmitting mobile voice and data services.
- It is used by over 800 million people in more than 190 countries.
- It was founded in 1984 as Groupe Spéciale Mobile Association.

Mobile services in GSM

- GSM makes it possible to combine various voice and data services and communicate with current networks.
- GSM has defined three different categories of services:

1. Bearer services / Data services
2. Tele services / Telephony services
3. Supplementary services.

1. **Bearer services / Data Services**

- Data services or Bearer Services are used through a GSM phone. to receive and send data is the essential building block leading to widespread mobile Internet access and mobile data transfer.
- GSM currently has a data transfer rate of 9.6k.
- New developments that will push up data transfer rates for GSM users are HSCSD (high speed circuit switched data) and GPRS (general packet radio service) are now available.
- Bearer services permit transparent and nontransparent, synchronous or asynchronous data transmission.
- **Transparent bearer** services only use the functions of the physical layer (layer 1) to transmit data.
- **Non-transparent bearer** services use protocols of layers two and three to implement error correction and flow control.

2. Tele services / Telephony services

- The abilities of a Bearer Service are used by a Teleservice to transport data.
- As telephony is the principal service, the fundamental objective of GSM was to provide high-quality digital voice transmission, at least providing the normal bandwidth of 3.1 kHz of analogue phone networks.
- Special codecs (coder/decoder) are used for voice transmission,while other codecs are used for the transmission of analog data for communication with traditional computer modems used in, e.g.,fax machines.
- A useful service for very simple message transfer is the short message service (SMS), which offers transmission of messages of up to 160 characters.

2. Tele services / Telephony services

- The successor of SMS, the enhanced message service (EMS), offers a larger message size (e.g., 760 characters, concatenating several SMs), formatted text, and the transmission of animated pictures, small images and ringtones in a standardized way.
- Another service offered by GSM is the emergency number. The same number can be used throughout Europe. This service is mandatory for all providers and free of charge. This connection also has the highest priority, possibly pre-empting other connections, and will automatically be set up with the closest emergency center.
- MMS offers the transmission of larger pictures (GIF, JPG, WBMP), short video clips etc. and comes with mobile phones that integrate small cameras.
- Another non-voice tele service is group 3 fax, which is available worldwide. In this service, fax data is transmitted as digital data over the analog telephone network.

3. Supplementary services.

- Supplementary services are additional services that are provided in addition to teleservices and bearer services.
- These services include :

1. caller identification,
2. call forwarding,
3. call waiting,
4. Conferencing,
5. barring of outgoing (international) calls

- Companies are particularly interested in closed user groups because they enable features like a company-specific GSM sub-network that is only accessible to group members.

GSM SYSTEM ARCHITECTURE

A GSM system consists of three subsystems:

1. The Radio Subsystem (RSS)
2. The Network and Switching Subsystem(NSS)
3. The Operation Subsystem (OSS)

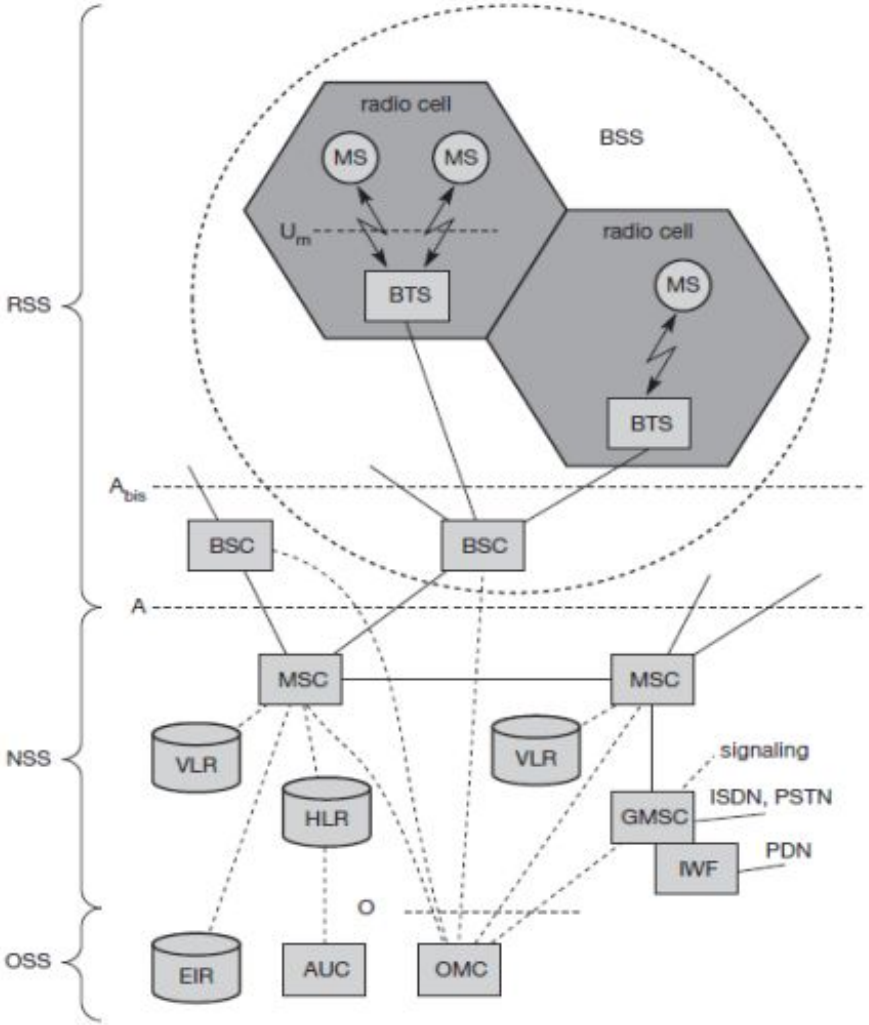


Figure 4.4
Functional architecture of a GSM system

1. **Radio subsystem**

- The radio subsystem (RSS) comprises all radio specific entities, i.e., the mobile stations (MS) and the base station subsystem (BSS).
- The RSS and NSS are connected via the A interface (solid lines) while the OSS is connected via the O interface (dashed lines) in the diagram
- Base station subsystem (BSS):

1. A GSM network has numerous base station subsystems, each controlled by a base station controller (BSC).
2. The BSS is responsible for maintaining radio connections to an MS, voice coding and decoding, and rate adaption to and from the wireless network.
3. The BSS contains many BTSs in addition to a BSC.

- Base transceiver station (BTS)

1. A BTS comprises all radio equipment, i.e., antennas, signal processing, amplifiers necessary for radio transmission.
2. A BTS can form a radio cell or, using sectorized antennas, several cells (see section 2.8), and is connected to MS via the Um interface (ISDN U interface for mobile use), and to the BSC via the Abis interface.

- Base station controller (BSC)

1. The BSC basically manages the BTSs.
2. It reserves radio frequencies, handles the handover from one BTS to another within the BSS, and performs paging of the MS.
3. The BSC also multiplexes the radio channels onto the fixed network connections at the A interface.

- Mobile station (MS)

1. The MS comprises all user equipment and software needed for communication with a GSM network.
2. 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

2. Network and switching subsystem

1. The network and switching subsystems are the "heart" of the GSM system (NSS).
2. The NSS connects the wireless network to conventional public networks, handles handovers between multiple BSSs, includes capabilities for global user localisation, and facilitates charging, accounting, and roaming of users across different providers and countries.
3. The NSS consists of the following switches and databases:

    1) Mobile services switching center (MSC)

    2) Home location register (HLR)

    3) Visitor location register (VLR)

- Mobile services switching center (MSC)

1. MSCs are high-speed digital ISDN switches. They use the A interface to link to other MSCs and BSCs, forming the GSM system's fixed backbone network.
2. An MSC is in charge of all signals for connection setup, connection release, and connection handover to other MSCs
3. Typically, an MSC is in charge of numerous BSCs in a certain geographic area.
4. Other fixed networks, such as PSTN and ISDN, are connected to a gateway MSC (GMSC).

- Home location register (HLR)

1. The HLR is the most important database in a GSM system as it stores all user-relevant information.
2. 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).
3. 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

- Visitor location register (VLR)

1. 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)
2. If a new MS comes into an LA the VLR is responsible for, it copies all relevant information for this user from the HLR.
3. This hierarchy of VLR and HLR avoids frequent HLR updates and long-distance signaling of user information.

3. Operation subsystem

- The operating subsystem (OSS), the third component of a GSM system, contains the functions required for network operation and maintenance.
- The OSS has its own network entities and communicates with others via SS7 signalling
- The following entities have been defined:

1. Operation and maintenance center (OMC)
2. Authentication centre (AuC)
3. Equipment identity register (EIR)

- The operation and maintenance centre (OMC)

1. It uses the O interface to monitor and control all other network elements (SS7 with X.25).
2. Traffic monitoring, network entity status reports, subscriber and security management, and accounting and billing are all common OMC administration responsibilities.
3. Telecommunication management networks (TMNs), as defined by the ITU-T, are used by OMCs.

- Authentication Centre (AuC)

1. As the radio interface and mobile stations are particularly vulnerable, a separate AuC has been defined to protect user identity and data transmission.
2. 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 Equipment Identity Register (EIR)

1. The EIR is a database for all IMEIs, i.e., it stores all device identifications registered for this network.
2. As MSs are mobile, they can be easily stolen. With a valid SIM, anyone could use the stolen MS.
3. 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.
4. The EIR also contains a list of valid IMEIs (white list), and a list of malfunctioning devices (gray list).

Localization And Calling in GSM

- The automatic, worldwide localization of users is a key aspect of the GSM system.
- The system always knows where a user is, and the same phone number can be used anywhere in the world.
- Even if a user does not use the mobile station, GSM performs periodic location updates to provide this service (assuming the MS is still logged into the GSM network and is not totally switched off).
- The current location (just the location area, not the specific geographical location) is always stored in the HLR, and the VLR in charge of the MS tells the HLR when the location changes.
- The HLR transfers all user data to the new VLR as soon as an MS enters the range of a new VLR (a new location region).

- Roaming is the process of switching VLRs while maintaining continuous availability of all services.
- Roaming can take place within a single provider's network, between two providers in the same country (national roaming is frequently not supported owing to operator rivalry), or between different carriers in other countries (international roaming).
- People usually identify the term roaming with international roaming because it is this form of roaming that makes GSM so appealing: one device, 190 countries

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

1. **Mobile station international ISDN number (MSISDN)**
2. **International mobile subscriber identity (IMSI)**
3. **Temporary mobile subscriber identity (TMSI)**
4. **Mobile station7 roaming number (MSRN)**

1. **Mobile subscriber international ISDN number (MSISDN)**

- For a GSM customer, the phone number is the most important number.
- The phone number is not linked to a specific device, but rather to the SIM, which is unique to each user.
- The exact length varies per country. It is used to identify a mobile subscriber across the globe uniquely.
- For addresses, the MSISDN uses the ITU-T standard E.164, which is also used in fixed ISDN networks. The country code (CC) (e.g., +49 179 1234567 with 49 for Germany), the national destination code (NDC) (i.e., the network provider's address, e.g., 179), and the subscriber number make up this number (SN).

MSISDN Format :

Mobile Country Code(CC) is 1 to 3 digits long.CC Indicates the country to which the number belongs. When a mobile subscriber dials a local mobile phone number, the country code should be omitted. A zero may be substituted for the country code.

For dialing outside of the country, the national format is used.

NDC or National Destination code: Identifies the area of a network operator within a country.

Subscriber Number (SN): Also referred to as subscriber digits. The mobile operator chooses these. The SN is appended following the CC + NDC.

2. International mobile subscriber identity (IMSI)

- GSM employs the IMSI to identify subscribers internally.
- A mobile country code (MCC)(e.g., 240 for Sweden, 208 for France), a mobile network code (MNC) (i.e., the network provider's code), and lastly a mobile subscriber identity number (MSIN) make up an IMSI (MSIN).
- The IMSI is stored in the Subscriber Identity Module (SIM) inside mobile devices and is sent by the mobile device to the appropriate network in which the SIM is associated.

Example:

Country – 234– UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Carrier - 15 – VODAFONE LTD

SIM ID – 23415

3. Temporary mobile subscriber identity (TMSI)

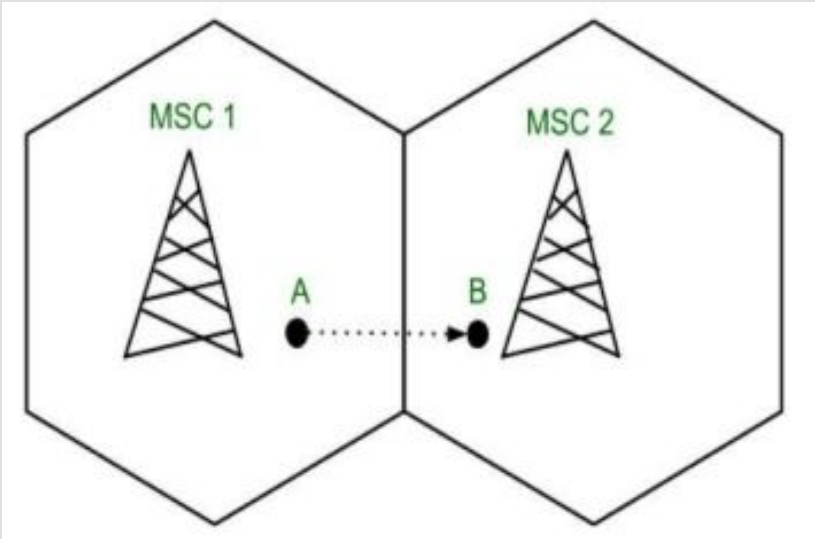
- GSM employs the 4 byte TMSI for local subscriber identification to disguise the IMSI, which would reveal the actual identity of the user signalling over the air interface.
- TMSI is chosen by the current VLR and is only valid for a limited time and within the VLR's location region (for ongoing communication, TMSI and LAI are sufficient; the IMSI is not required).
- A VLR may also modify the TMSI on a regular basis

4. Mobile station roaming number (MSRN)

- Another temporary address that hides the identity and location of a subscriber is MSRN.
- The VLR generates this address on request from the MSC, and the address is also stored in the HLR.
- MSRN contains the current visitor country code (VCC), the visitor national destination code (VNDC), the identification of the current MSC together with the subscriber number.
- The MSRN helps the HLR to find a subscriber for an incoming call.

HANDOVER

- Handover or handoff refers to the procedure of transferring ongoing call or data connectivity from one Base Station to another in cellular telecommunications.
- When a phone goes to a different cell while a call is in progress, the MSC (Mobile Switching Center) transfers the call to a new channel associated with the new Base Station.
- When a mobile user A moves from one cell to another, the signal strength of BSC 1 decreases while the signal strength of BSC 2 improves, allowing the mobile user to continue making calls or accessing data without interruption.

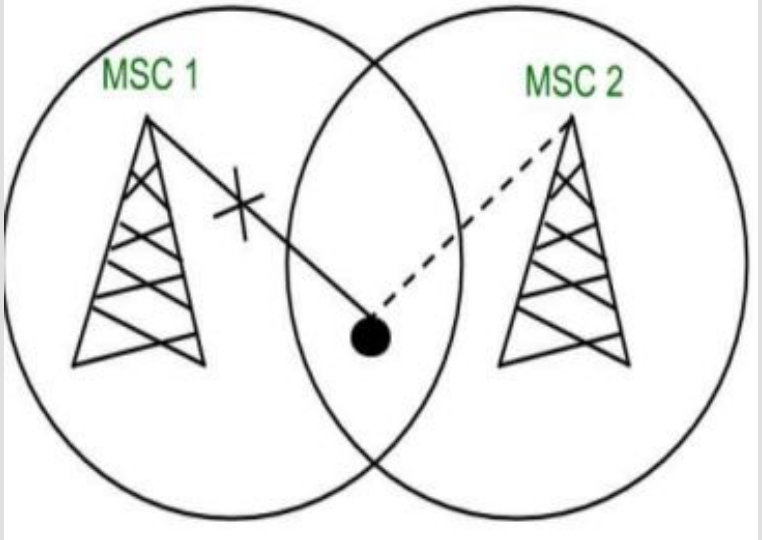


Types of Handoff

- **Hard handoff**

1. When transitioning from one Base Station to another Base Station, there is an actual interruption in connectivity.
2. The Base Station and MSC are not burdened since the changeover occurs so swiftly that the users are barely aware of it.
3. The quality of the connection is poor.

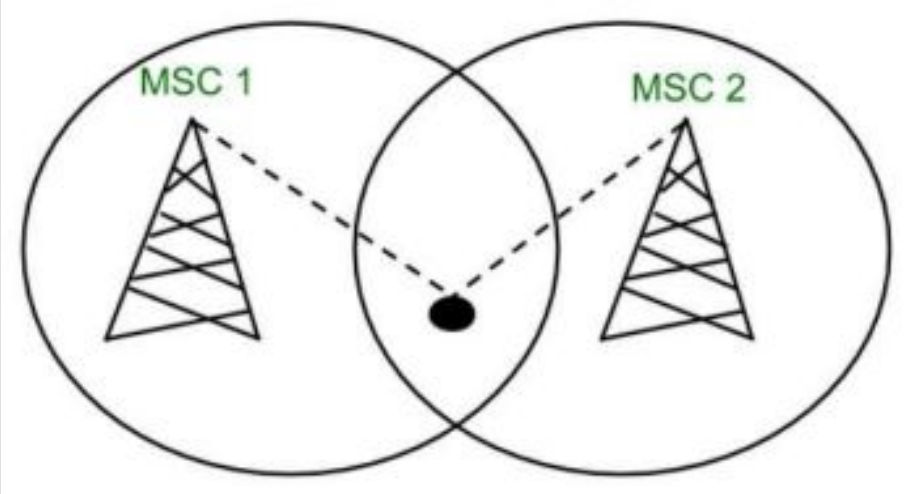
The 'break before make' policy was implemented by Hard Handoff.



Types of Handoff

- **Soft handoff**

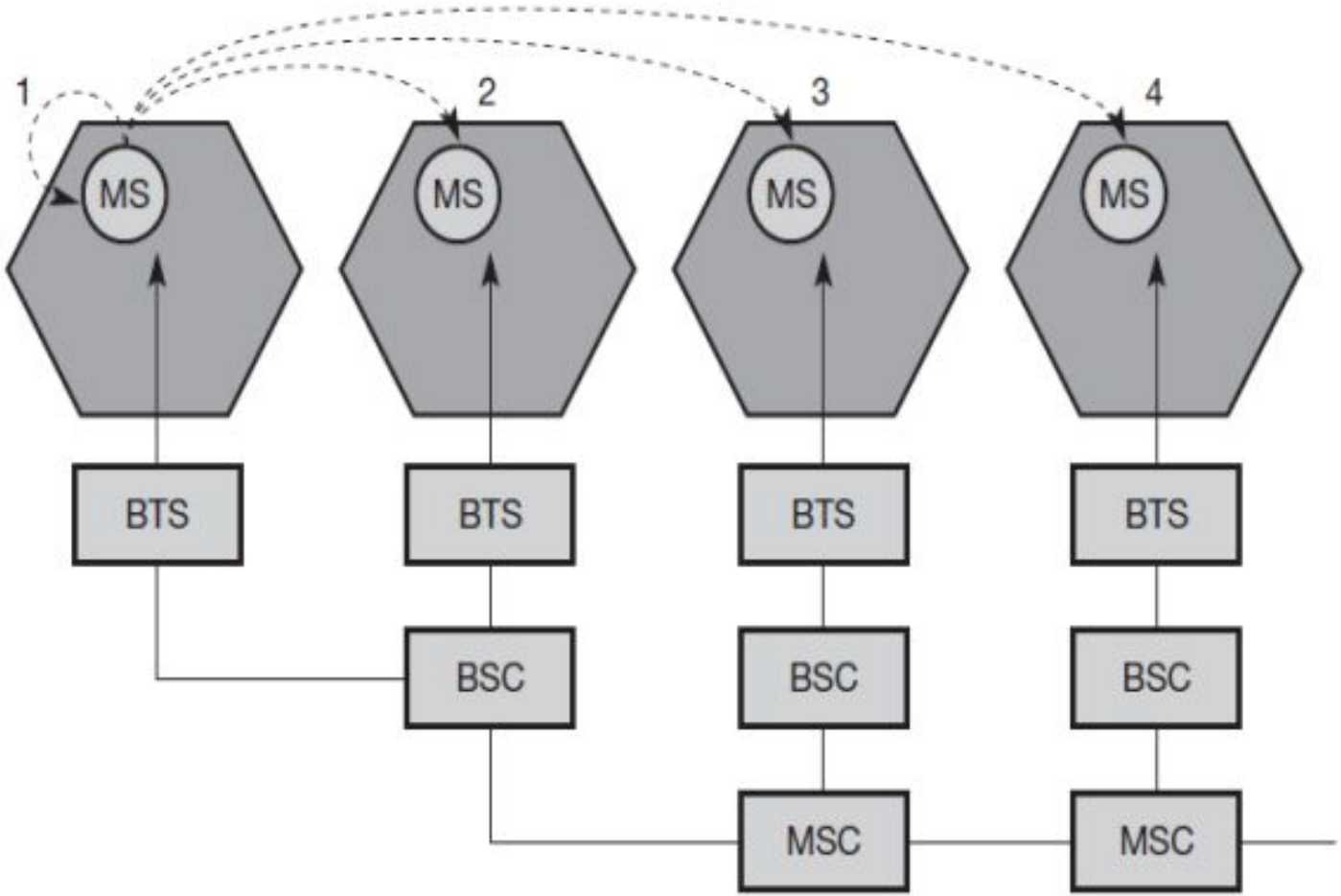
1. When radio signals are added or removed from the Base Station, at least one of the links is retained in Soft Handoff.
2. The'make before break' principle was implemented by Soft Handoff. Hard Handoff is more expensive than Soft Handoff.



Four possible handover scenarios in GSM:

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

Figure 4.11
Types of handover in GSM



Security in GSM

- GSM provides a variety of security services based on information saved in the AuC and individual SIMs (which is plugged into an arbitrary MS).
- The SIM card stores personal and confidential information and is secured with a PIN to prevent unwanted access.
- The secret key Ki, for example, is saved in the SIM and is used for authentication and encryption procedures.

Security services offered by GSM:

1. Access control and authentication

The first step is to verify that the SIM user is legitimate. To utilise the SIM, the user must enter a secret PIN.The subscriber authentication is the next stage. A challenge response method is used in this step.

2. Confidentiality

All user-related data is encrypted for privacy. Following authentication, the BTS and MS encrypt speech, data, and signalling. This level of confidentiality occurs just between MS and BTS, not from end to end or throughout the entire fixed GSM/telephone network.

Security services offered by GSM:

3. 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(Visitor Location Register) after each location update. Additionally, the VLR can change the TMSI at any time.

DECT

Digital Enhanced Cordless Telecommunications

- It is a wireless standard that is very often used for landline phones.
- A DECT system always contains two components that constantly communicate with each other.
- The two components are a base station, also called the fixed part, and at least one handset, or portable part.

DECT System architecture

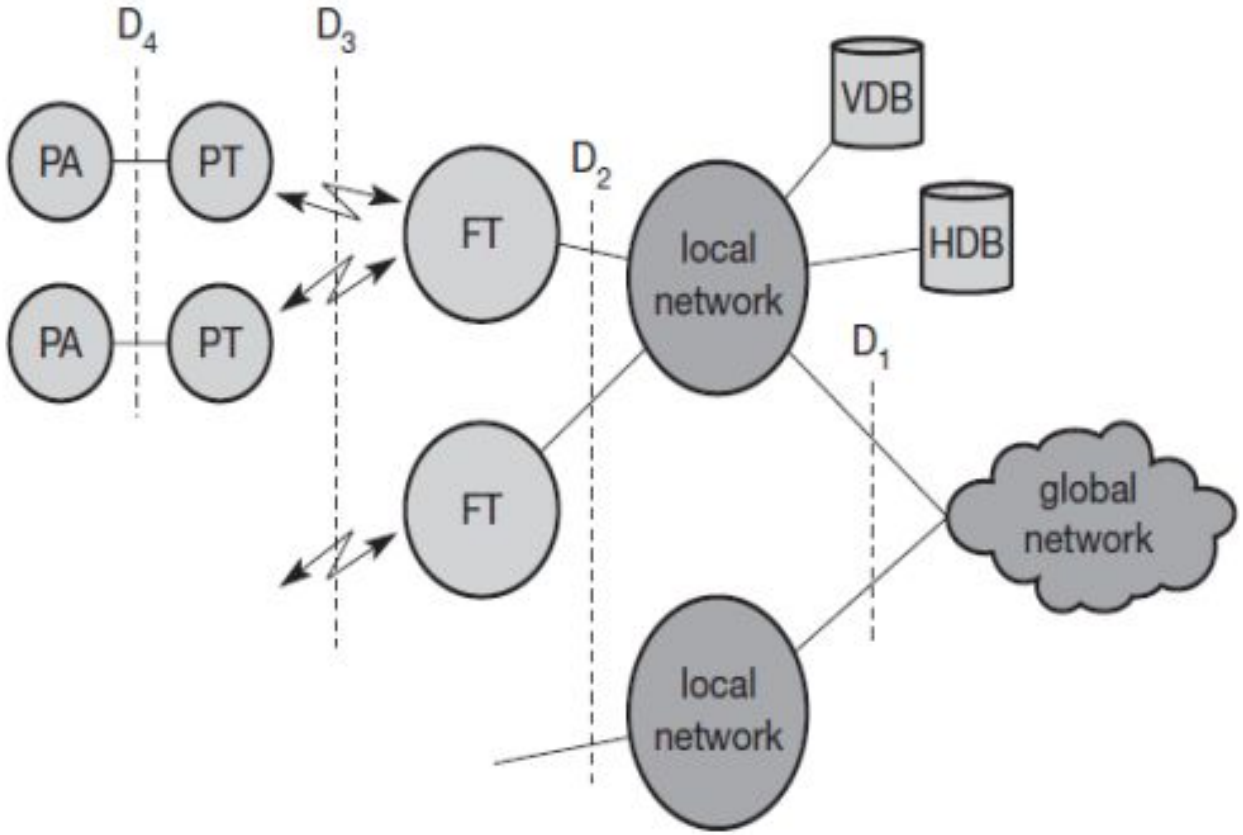


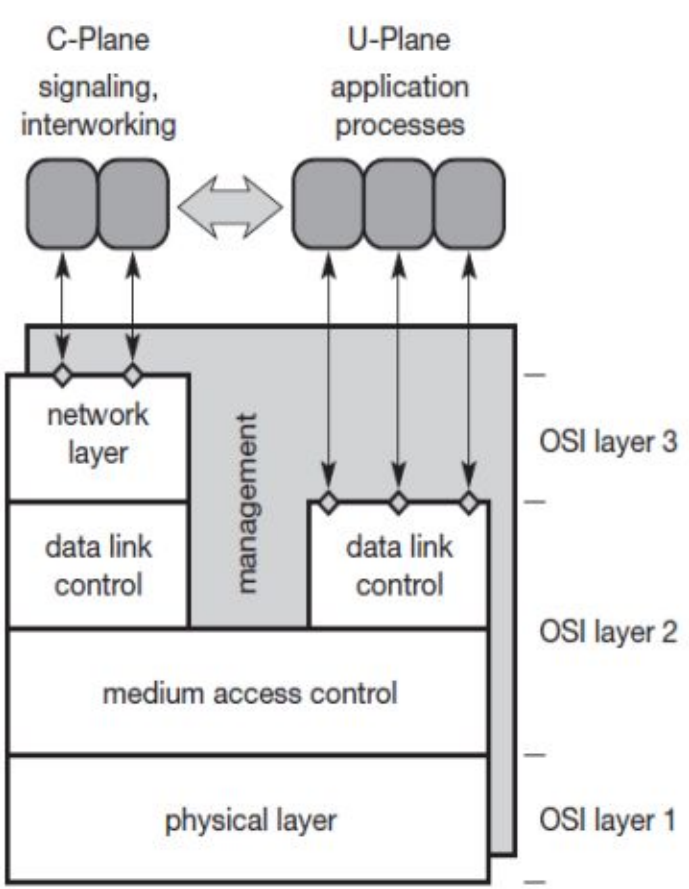
Figure 4.18
DECT system architecture reference model

- The local communication system is linked to the outside world via a global network, which provides its services over the D1 interface.
- Public switched telephone networks (PSTN), public land mobile networks (PLMN), such as GSM, or packet switched public data networks are examples of global networks. (PSPDN).
- These networks provide a variety of services, such as data transportation, address translation, and data routing between local networks.
- Local networks provide local telecommunication services that can range from straightforward switching to clever call forwarding, address translation, etc

- All normal network tasks must be integrated in the local or global network, where the databases home data base (HDB) and visitor data base (VDB) are also located, despite the DECT system's relatively simple core.
- Both databases support mobility with functions that are similar to those in the HLR and VLR in GSM systems.
- Incoming calls are automatically forwarded to the current subsystem responsible for the DECT user, and the current VDB informs the HDB about changes in location.
- The fixed radio termination (FT) and portable radio termination (PT) make up the DECT core network, which essentially just offers a multiplexing service.

DECT Protocol Architecture

Figure 4.19
DECT protocol layers



- The DECT protocol reference architecture follows the OSI reference model. Figure shows the layers covered by the standard: the physical layer, medium access control, and data link control for both the control plane (CPlane) and the user plane (U-Plane).
- An additional network layer has been specified for the C-Plane, so that user data from layer two is directly forwarded to the UPlane.
- A management plane vertically covers all lower layers of a DECT system.

- Physical layer:

1. The physical layer of every wireless network includes all operations for modulation and demodulation, incoming signal detection, sender/receiver synchronization, and gathering status data for the management plane.
2. The physical channel structure is generated by this layer with a predetermined, guaranteed throughput. The physical layer assigns a channel for data transmission in response to a request from the MAC layer.

- Medium access control layer

1. By activating and deactivating physical channels, the media access control (MAC) layer creates, maintains, and releases channels for higher layers.
2. Multiple logical channels are multiplexed onto physical channels using MAC.
3. There are logical channels for broadcast messages, user data transmission, paging, and signalling network control.
4. The segmentation and reassembly of packets as well as error control and error correction are additional services provided.

- Data link control layer

1. The data link control (DLC) layer creates and maintains reliable connections between the mobile terminal and the base station.
2. Two services have been defined for the C-Plane: a connectionless broadcast service for paging (called Lb) and a point-to-point protocol similar to LAPD in ISDN, but adapted to the underlying MAC (called LAPC+Lc).

- Network layer

1. The network layer of DECT is similar to those in ISDN and GSM and only exists for the C-Plane.
2. This layer provides services to request, check, reserve, control, and release resources at the fixed station (connection to the fixed network, wireless connection) and the mobile terminal (wireless connection).
3. The mobility management (MM) within the network layer is responsible for identity management, authentication, and the management of the location data bases. Call control (CC) handles connection setup, release, and negotiation.
4. Two message services, the connection oriented message service (COMS) and the connectionless message service (CLMS) transfer data to and from the interworking unit that connects the DECT system with the outside world.

TETRA
Terrestrial Trunked Radio

- Trunked radio systems constitute another method of wireless data transmission.
- These systems use many different radio carriers but only assign a specific carrier to a certain user for a short period of time according to demand.
- While, for example, taxi services, transport companies with fleet management systems and rescue teams all have their own unique carrier frequency in traditional systems, they can share a whole group of frequencies in trunked radio systems for better frequency reuse via FDM and TDM techniques.

- TETRA offers two standards: the Voice+Data (V+D) service (ETSI, 1998l) and the packet data optimized (PDO) service (ETSI, 1998m).
- While V+D offers circuit-switched voice and data transmission, PDO only offers packet data transmission, either connection-oriented to connect to X.25 or connectionless for the ISO CLNS (connectionless network service)
- The latter service can be point-to point or point-to-multipoint, the typical delay for a short message (128 byte) being less than 100 ms.
- V+D connection modes comprise unicast and broadcast connections, group communication within a certain protected group, and a direct ad hoc mode without a base station.
- However, delays for short messages can be up to 500 ms or higher depending on the priority.

- TETRA also offers bearer services of up to 28.8 kbit/s for unprotected data transmission and 9.6 kbit/s for protected transmission.
- Examples for end-to-end services are call forwarding, call barring, identification, call hold, call priorities, emergency calls and group joins
- While V+D uses up to four TDMA voice or data channels per carrier, PDO performs statistical multiplexing. For accessing a channel, slotted Aloha is used.
-

UMTS
Universal Mobile Telecommunications System

- Universal Mobile Telecommunications Service (UMTS) refers to a group of radio technologies associated with the 3G cellular networks.
- It is based on the Code Division Multiple Access (CDMA) network standard.
- CDMA essentially enables the same wireless signal to be shared by many different devices, so the capacity for CDMA networks is up to five times that of 2G GSM-based services.
- This significantly reduces the likelihood of service drop-off or issues with connectivity in the event of heavy system traffic.
- UMTS is a packet-switched system.

UMTS Applications

- Streaming / Download (Video, Audio)
- Videoconferences.
- Fast Internet / Intranet.
- Mobile E-Commerce (M-Commerce)
- Remote Login
- Background Class applications
- Multimedia-Messaging, E-Mail
- FTP Access
- Mobile Entertainment (Games)

Features	GSM	UMTS
Network Architecture	Circuit-switched	Circuit-switched and packet-switched
Radio Access Technology	FDMA and TDMA	Wideband CDMA (W-CDMA)
Bandwidth	200 kHz	5 MHz
Data Rate	Up to 384 kbps	Up to 2 Mbps for HSDPA; up to 7.2 Mbps for HSDPA
Applications	Voice and SMS	Multimedia applications
Roaming Support	Limited	Automatic international roaming
Video Quality	Poor	Improved compared to GSM
Cost	Affordable	More expensive than GSM
Broadband	Not broadband	Offers broadband capabilities

SATELLITE SYSTEMS

- Satelites, which are far away from the surface of the earth, can cover large area, with several satellite beams being controlled and operated by one satellite.
- Large areas can be covered due to the rotation of satellites around the earth.

APPLICATIONS of Satellite Systems:

Satellites have historically been applied in the following fields:

- **Weather prediction**: Several satellites send images of the world using infrared or visible light, for example. It would be impossible to forecast hurricanes without the assistance of satellites.
- **Radio and TV broadcast satellites:** Satellites used for radio and TV broadcasts make thousands of radio and television programmes accessible. Since it costs less to install and typically requires no additional fees, this technology competes with cable in many locations. In central Europe, modern satellite dishes have diameters of 30 to 40 cm. (the diameters in northern countries are slightly larger).

- **Military satellites:** One of the earliest applications of satellites was their use for carrying out espionage. Many communication links are managed via satellite because they are much safer from attack by enemies.
- **Satellites for navigation**: The global positioning system (GPS), which was once primarily utilised for military purposes, is today well-known and accessible to everybody. Worldwide exact localization is possible with the technology, and with some additional techniques, the precision can reach a few metres. The majority of ships and aircraft use GPS in addition to more conventional navigation systems.

BASICS: GEO, LEO, MEO & HEO

- **Geostationary (or geosynchronous) earth orbit (GEO)**

GEO satellites have a distance of almost 36,000 km to the earth. Examples are almost all TV and radio broadcast satellites, many weather satellites and satellites operating as backbones for the telephone network

- **Medium earth orbit (MEO)**

MEOs operate at a distance of about 5,000–12,000 km. Up to now there have not been many satellites in this class, but some upcoming systems (e.g., ICO) use this class for various reasons

- **Low earth orbit (LEO)**

While some time ago LEO satellites were mainly used for espionage, several of the new satellite systems now rely on this class using altitudes of 500–1,500 km

- **Highly elliptical orbit (HEO)**

This class comprises all satellites with noncircular orbits. Currently, only a few commercial communication systems using satellites with elliptical orbits are planned. These systems have their perigee over large cities to improve communication quality.