LOW POWER RADIO NETWORKS

By

Monish Chaudhary

Supervisor: Peter Baxendale

Centre for Communication Systems

School of Engineering

University of Durham

A THESIS UNDERTAKEN IN ENGINEERING SCHOOL OF THE UNIVERSITY OF DURHAM IS SUBMITTED FOR THE DEGREE OF MASTER OF SCIENCE (MSc).

MARCH 2005

Declaration

I hereby declare that this thesis is a record of work undertaken by myself and supervised by Peter Baxendale, that has not been the subject of any previous application for a degree, and all the sources of information have been duly acknowledged.

Abstract

Low power radio networks are the networks which depend upon wireless radio links and consume very low energy for their operation. These networks suit best for applications where frequent renewal of power supply is not possible.

Power supply has always remained a major concern in radio networks. An efficient low power consuming network is always recommended for greater mobility and lifetime of the network. This thesis introduces low power radio networks, their features and applications. Energy concerns and various techniques that can be used for energy conservation are discussed, along with the security techniques that can be used to make the system reliable. Different technologies available in the market and their features and applications are considered. Included is a detailed study of the IEEE 802.15.4 standard. A simulation study of the CSMA/CA algorithm and topology discovery algorithms is presented.

Acknowledgements

I would like to express my sincere thanks to my supervisor Peter Baxendale for its support, help and patience. He has been a major source of inspiration throughout the degree. I would also like to thank Jim Swift for his help. Thanks must be expressed to David Clark and Yu Liye for being nice to me all this time and sharing their knowledge about the subject with me. Finally, thanks to my parents and my friends; without their support and motivation it was difficult to accomplish this degree.

Contents

Acronyms and abbreviations

ACK	acknowledgement
ACL	access control list
ACL	asynchronous connectionless
AODV	ad-hoc on-demand distance vector
BER	bit error rate
BPSK	binary phase-shift keying
CAP	contention access period
CCA	clear channel assessment
CFP	contention free period
CSMA-CA	carrier sense multiple access with collision avoidance
CSMA-CD	carrier sense multiple access with collision detection
CTS	clear to send
DCF	distributed coordination function
DLL	data link layer
DSDV	destination-sequenced distance vector
DSR	dynamic source routing
DSSS	direct sequence spread spectrum
ED	energy detection
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission
FCS	frame check sequence
FFD	full function device
FHSS	frequency hopping spread spectrum
GTS	guaranteed time slot
IEEE	Institute of Electrical and Electronics Engineers
ISM	industrial, scientific and medical
LR-WPAN	low-rate wireless personal area network
LQI	link quality indication
LSB	least significant bit
MAC	medium access control
MCPS	MAC common part sublayer
MFR	MAC footer
MHR	MAC header
MLME	MAC layer management entity
MMAC	multimedia mobile access communication
MPDU	MAC protocol data unit
MSDU	MAC service data unit
NAV	network allocation vector
NIC	network interface card
OBEX	object exchange protocol
OFDM	orthogonal frequency division multiplexing
O-QPSK	offset quadrature phase-shift keying
OSI	open systems interconnection
PAN	personal area network
PCF	point coordination function

PHR	PHY header
PHY	physical
PIB	pan information base
PLME	PHY layer management entity
POS	personal operating space
PPDU	PHY protocol data unit
PRN	pseudo random noise
PSDU	PHY service data unit
RF	radio frequency
RFD	reduced function device
RREP	request reply
RREQ	route request
RTS	request to send
SCO	synchronous connection oriented
SFD	start-of-frame delimiter
SHR	synchronization header
SINR	signal-to-interference-plus-noise ratio
SSID	service set identifier
TCP/IP	transmission control protocol/internet protocol
TDD	time division duplex
TDMA	time division multiple access
WECA	Wireless Ethernet Compatibility Alliance
WLAN	wireless personal area network

Chapter 1

Low Power Radio Networks

1.1 Introduction

Low power radio networks are radio networks which are linked by radio links and which are driven by extremely low power. Low power radio networks are mainly driven by batteries instead of mains power supply which further intensifies the need of low and efficient consumption of available power in order to maximize the lifetime of the network without renewal of the energy source and without human intervention. Because of radio connectivity they are mobile and have no wires attached to restrict their mobility. Low power radio networking is the best option for the applications where there is shortage of power and frequent renewal of energy source is not possible. Due to the installation of radio devices at remote sites, it sometimes becomes difficult to renew the energy source, which simply indicates the end of network lifetime. In these circumstances a low power network with greater lifetime will always be preferred over the one with short lifetime. An efficient and low power consuming network will provide greater mobility and lifetime.

1.2 License Free Frequency Bands

A major factor responsible for making radio networks extremely popular is the availability of some license free bands around the world which are regulated by the government of the respective countries. It simply means that under some regulated manner any one can use these frequency bands in their homes and commercial installation to achieve certain degree of automation and monitoring without any sort of charges towards the government or any other authority. There are different frequency bands available around the world and they are divided region wise.

Following are some of the unlicensed frequency bands which can be used for low power radio networking [Gut03a] [IEEE03a].

- 868.0-868.6 MHz band: Used in most of the European countries
- 902-928 MHz band: Used in North America
- 2.40-2.48 GHz: Used in most of the countries around the world
- 5.7-5.89 GHz: Used in most of the countries worldwide.

Networks working in these unlicensed bands must comply with the local regulations of the country they are operating in. In the USA, the Federal Communications Commission (FCCI) is the institution in charge of the regulation of the networks. In European countries the regulatory body is the European Telecommunications Standards Institute (ETSI). Other countries around the world use their own regulatory agencies; many of them accept either FCC or ETSI acceptance as proof of compliance [Gut03a]. These regulatory bodies make sure that the manufacturers and sellers don't sell any transmitters which exceed the maximum limit of the radiation that they can use while operating in these license free frequency bands. Transmitters must have a FCC ID label on them to be sold which proves that they have undergone testing to verify that the radiation is below the limit. Under Part 15 of the federal regulations, users may employ low-powered transmitters as long as they don't interfere with licensed users and in addition they must be willing to accept interference from other low-powered users [FCC15]. Recently this part was relegated to simple applications such as cordless phones, garage door openers, walkie-talkies and key–chain gadgets that lock car doors. Because of the short distance involved in the node-to-node transmissions, networks can take advantage of the flexibility of low-powered transmissions. No license is needed for operation; anyone can use them as long as the equipment is FCC certified.

These frequency bands have multiple channels associated with them. For example, according to IEEE 802.15.4 standard the 860.0-868.0 MHz frequency band has only one channel associated with it while other frequency bands such as 915 MHz and 2.4 GHz are having 10 and 16 channels respectively associated with them [IEEE03b].

1.3 Features [Call04a]

1.3.1 Low Range

Low power radio networks are also effective only to a restricted area which means that they have a low coverage area. Because these networks are operating in license free bands, there are regulations which restrict the transmission power so as to avoid it from interfering with other radio signals which are working in licensed bands and also to make sure that networks using an unlicensed band don't interfere with the privacy of other networks. Another reason for limiting transmission power is the requirement for low power consumption, which can only support short range connectivity. It is seen that the typical range of low power radio networks is some 10s of meters, which makes it suitable for indoor applications or in outdoor applications with limited area by using mesh networks [Gut03b].

1.3.2 Low Cost

Cost of a network plays an important role to ensure the popularity and wide scale acceptance of the network. The idea is to avoid all high cost components and reduce silicon area in chips used by minimizing complexity and memory requirements. Use of these low power radio systems in common life has compelled companies to be involved in mass production of these systems, which has make the technology very inexpensive, less complex and easily accessible. Because of the widespread use of these systems they need to be more reliable, less complex, and capable of self configuration and self maintenance, which means that the system should be adaptive to, for example, topological changes. By self configuration is meant the sensitivity towards other nodes in the vicinity, whilst self maintenance is defined as the ability to recover from fault without any sort of human intervention. It makes sense because one can not afford a low cost network with high maintenance cost.

1.3.3 Low Power Consumption

Applications in low power radio networks require wireless connectivity and mobility at the same time which makes it difficult to have external power leading to a dependency on batteries. These batteries should last a long time because applications such as monitoring and control of industrial equipment can not afford to change batteries frequently in order that the maintenance schedule of the equipment is not disturbed. Frequent change of batteries would violate the low cost and low maintenance intention behind the development of these networks. There are also certain applications which have to rely on energy scavenged from the environment, such as sensing tyre pressure in automobiles. Hence, sensors obtain necessary energy from mechanical or thermal energy present in the tyres. Low power networks adopt certain techniques to conserve energy such as introducing sleep cycles, which means that the device is not working at all times but periodically.

1.3.4 Low Data Rate

Low power radio networks typically have low data rate output. In most of the application scenarios, these networks are used in remote sensing and monitoring, where a high data rate is not needed. Different frequency bands give different data rates, ranging from 20kb/s to 250 kb/s. Moreover, because of extremely low power consumption, the network can not afford to give higher data rates.

1.3.5 Relaxed Latency

Low power radio networks support relaxed latency. By latency we mean the time taken to deliver a message from source to destination. Efficient power consumption and longer battery life being the main aims of these networks, most of the nodes remain in sleep mode most of the times and wake up periodically for a short period to deliver data packets. This mechanism adds to the latency. Therefore low power radio networks best suit those applications which are not latency critical.

1.4 Applications

Low power radio networks have a wide area of application. All those applications which have relaxed throughput requirements and in which data rate is a few bits per day can use low power radio networks for greater control and automation. Some of the applications which come under this are industrial control and monitoring, home automation, consumer electronics, security and military sensing, asset tracking, intelligent agriculture, health monitoring, automotive sensing, toys and games etc [Call04a] [Gut03c].

1.4.1 Industrial Control and Monitoring

Low power networking can be successfully used in industries for control and monitoring where an application can afford longer latencies and data throughput is low. Such networks can be used, for example, for controlling systems involving certain moving parts of machines to monitor their temperature, vibration etc where wired sensors can not be used. In such case the life time of the network is critical; otherwise it will add undue maintenance to the machine while changing the sensors. In such cases sensors usually don't use any batteries as their energy source but work by using thermal or mechanical energy from the moving parts. This is called energy scavenging and ensures longer lifetime to the network. For controlling heat ventilation and air conditioning, sensors can be installed at various sites which gather information and send it to the control room periodically. Low power networks can be use for monitoring purposes in industry. By setting up such networks in the entire campus one can monitor what is going on in every part by just sitting in the control room. One example of such an application is controlling the lights of a large installation, which if it had been done manually can be quite cumbersome. It can also be used to detect the presence of certain harmful gases or leakage of certain chemicals in the premises to avoid any accident.

1.4.2 Home Automation

With the introduction of wireless networks the home automation industry has faced a tremendous change and the number of consumers has increased manifold. Installation of such a system does not need any major change in existing set up because of the wireless nature of the system, which makes it quite easy to install with minimal configuration by the user. One can control most of the household electrical and electronic devices with a universal remote control and can program them to work at pre-programmed times or on demand. These applications involve controlling television, CD players, stereo, washing machine, turning the lights on or off, pulling curtains at night, locking doors and windows etc. Another major application of low power networks can be in a personal computer and its peripherals, such as wireless keyboards, mice, printers etc.

1.4.3 Home Security and Military Sensing

Low power radio networks can be used to secure homes by the introduction of sensors that support magnetic doors, detect broken glass or smoke and sensors for the detection of human intrusion. Such sensors, upon detection of any such activity can send a response to the owner or can set an alarm. Similarly in the military area also the low power networks can be deployed to sense the activity of the enemy without any personal attendance at the remote sites and can also be use for surveillance purpose in the defence perimeter. These can be also be used in case of a potential attack where they can locate targets and enemy movement. These can be very difficult to destroy because of their wide distribution and low probability of detection; aided by the use of spread spectrum techniques.

1.4.4 Intelligent Agriculture

Low power radio networks can be used for agriculture purposes also, which can make a positive impact on the growth of crops by precise knowledge of the various key factors influencing crop health [You04]. These networks can be used to increase the quality and quantity of the product without much cost and human intervention. In this case sensors can be deployed in the required area to gather field information regarding moisture content in the soil, the amount of rainfall, temperature, humidity etc. These sensors gather information and pass it via some form of gateway to a central control unit. Thus any sort of deficiency can be reported at the earliest and can be dealt with immediately. As the data throughput is low and higher latencies can be tolerated, the lifetime of the system can be from months to years without renewal of energy source. Several such systems have already been deployed world-wide.

1.4.5 Health Monitoring

Non life-critical health information can be gathered with the help of radio networks, such as performance of athletes by monitoring their pulse rate and respiration rate and analysing their muscular movement, sending information to a computer for further analysis. One application could be automatic scaling of body weight at the door step. These developments are further expected to boom with the introduction of certain biological sensors which can sense harmful enzymes, nucleic acids etc.

1.4.6 Automotive Sensing

Wireless monitoring of automotive parts can be quite beneficial where sometimes wired networking is not possible because of motion of certain parts. A simple and most common application is remote for keyless entry and for central locking of doors and windows from one point. These days’ modern cars are deployed with sensors to check the performance of the vehicle regarding fuel consumption, wear and tear, vibration etc. Another application is to mount one sensor on each tyre which can send information regarding the tyre pressure to the central unit in the car. Here also the sensors derive their energy from thermal and mechanical energy of the tyre, and don’t need any additional energy source, leading to a life as long as the life of the tyre.

Chapter 2

Issues involved in Low Power Radio Networks

2.1 Energy concerns

Although wireless networks have existed for many years already, explicit concern about their energy efficient operation has emerged only recently. It is quite evident when the power source is either costly or in short supply; energy efficiency is of paramount importance.

Wireless networks can be roughly categorised into four major classes [Eph02]: cellular, wireless LAN, satellite and ad-hoc. Although cellular networks use wireless connectivity in the first and last segment of the communication path there are concerns about longevity of the battery life (which can usually be recharged). The second class, wireless LAN is completely wire less, but most of the devices work over a single hop transmission. Typical examples are laptops with Bluetooth or 802.11 network cards. No doubt there are concerns about energy in such networks but they are not as pressing. The third class of wireless networks are the networks which use satellite links for communication. Despite the possibility of recharging the batteries after regular intervals they face a serious problem of energy drain. The fourth class is relevant to this thesis: the ad-hoc networks. These networks communicate through multiple hop transmission, have no infrastructure and need maximum battery lifetime without renewal.

The energy radiated by the antenna of the transmitting node does not follow a strict rule regarding its coverage of the area because of the physical obstructions in its path and environmental conditions. Therefore, it attenuates as it travels away from the transmitter. However, in ideal conditions, the signal level decays non-linearly following the equation:

$$S(r) = S\, r^{-\alpha}$$

where S is the amplitude of the transmitted signal,

r is the distance from the transmitter, and

α is a parameter whose value ranges from 2-4.

For a successful reception of a signal, it is required to establish a certain level of quality of service (QoS) which can be in terms of a maximum acceptable value of bit error rate (BER). This should be 10^{-9} or lower in case of data signals and can be higher for speech.

Successful transmission also depends upon other factors like noise that the channel introduces the bandwidth and the rate of transmission. These factors are summarised by the equation [Eph02]:

$$\text{SINR} > \phi$$

where SINR is received signal-to-interference-plus-noise ratio detected at the receiver. ϕ is the threshold that depends upon the detecting structure, modulation/demodulation and coding/decoding.

The left hand side of the equation depends upon the channel, other user signals, transmitting and receiving antenna, RF transmission power (P) and transmission rate (R). ‘P’ and ‘R’ are highly adjustable and determine the amount of signal energy associated with a symbol. A link’s feasibility depends upon various factors starting from desired BER and ending with the chosen values of ‘P’ and ‘Q’ while its strength and quality can be adjusted with the right combination of ‘P’ and ‘Q’. A feasible link will go down if the value of SINR goes below the threshold value ϕ.

Energy is consumed in a wireless network in three major modes i.e. transmitting, receiving or idle mode [PCS02].

Transmitting mode: In the transmitting mode energy is spent in two ways. The first is in the amplifier which is responsible for generation of the RF transmission. This includes internal heat loss and radiated energy and depends on the efficiency of the amplifier. The second form of energy consumption is through the node processor which implements all the signal generation, coding, modulation and other signal processing functions.

Receiving mode: In the receiving mode, energy is consumed by the processor which includes the low-noise amplifier which boosts the output of the receiving antenna to a suitable level for demodulation, decoding, buffering and so on.

Idle mode: In the idle mode even if the device is not transmitting or receiving it may consume power since a voltage controlled oscillator may be in operation, ready to demodulate an incoming signal. The device keeps listening to the wireless medium for the packets it should receive, so that it can come to receiving mode. Though energy is consumed, it is very much less than the transmitting and receiving mode.

The introduction of a sleep mode in wireless networks has greatly improved energy efficiency. In this case, the device does not even listen to the wireless medium, so it switches off the radio, which brings a lot of energy saving. But the energy saving comes at the cost of increased latency. As wireless network nodes are meant to work without energy renewal for a long period and the data rate associated with them is very low, the node can remain in sleep mode for most of the time and wake up only for a few moments to listen for data.

Besides this there are several other factors which can contribute towards the energy saving criteria. Some on them are as follows [Eph02]:

Proper selection of batteries can make a difference. It is not the amount of energy in the battery which matters but the pattern in which energy is drained out. The energy consumed while using a network with a TDMA protocol can give more lifetime than one using other MAC protocols because in TDMA, every device has a fixed slot for transmission for each device while in other MAC protocols, for example, in some contention based protocols a device might end up consuming more energy while contending for the medium.

Energy can also be saved by the proper selection of the hardware. For example, power amplifiers usually have a non linear curve when they are driven to saturation. So power amplifiers with better energy curves can save energy. Proper circuit layout can also help as it can reduce thermal effects and excessive heating of the device.

Antennas play a major role and the type of antenna and its shape can make a significant difference. The correct antenna with the right shape and the right material can cover a larger distance while consuming less power. The radiation pattern, which depends upon the type of antenna, also affects the energy consumption and coverage area.

The correct combination of modulation/demodulation and coding/decoding determine the spectral effects which include spread spectrum and susceptibility to interference. Therefore they determine the correct value of signal strength required for successful reception and hence affects energy consumption.

The RF signal attenuates when it travels away from the transmitter so the transmission of the data over a single hop can also result in excessive power consumption because of large transmission power. This problem can be solved by sending the data over multiple hops. The transmission cost decreases in this case, but the energy used in processing increases, as does the latency. The transmission energy saved by multi-hop transmission may exceed the energy consumed in processing by the selection of a proper routing algorithm.

2.2 Network layer concerns

2.2.1 Topology

Low power radio networks are implemented by three main types of topology architectures: star, mesh and star-mesh hybrid topologies [Reid04]. The use of a particular network topology in a network depends upon various factors such as the operating frequency of the network, data rate associated with the network or required by the network, transmission distance, battery life and the mobility. The star topology is also called a point-to-multipoint network which is the simplest form of network configuration and established between two radios which are in direct communication with each other [Gero05]. Each node can only communicate with the coordinator. The wireless link between the identical radio devices and the base station, or "coordinator", is not shared which means that the link remains dedicated to coordinator and device and no other device can use that link until the current transmission is complete. So once established it will remain established until the device is in working condition. This network topology uses low energy, but because of direct linking between the base station and the device the overall communication range between the device and the base station is small (30-100 meters).

The second topology architecture, mesh network, is a multipoint-to-multipoint network, where a device can communicate with multiple devices. Mesh networks are multi-hop networks where the data passes through a series of nodes before reaching the destination. It is different from the point-to-point network in that in these networks the network devices can communicate with each other, which is not possible in a star network. Since a single node is well connected with all its neighbours, it makes the network highly fault tolerant. In case of any node failure a sending device can opt for a different route to pass the message to the destination node. Since the node can make a direct link with more than one neighbouring device, the network becomes more complex. Mesh networks can cover a large area but have a higher energy consumption rate because of increase in control overhead resulting from greater connectivity. Because of the multi-hop transmission it has also a higher latency.

The third and the newest type of topology is the star-mesh hybrid, which has the characteristics of both star network and mesh network. It is a low power consuming, simple, self healing wireless communication network with an extended communication range. These are ad-hoc, self organising and self healing networks. If the topology of the network changes then the network reorganises the topology structure and informs all the nodes about the change. Similarly if a node goes down or fails, then again the network coordinator sends the relevant information to all the nodes. It is self-healing in that if any route involved in communication between the two devices fails, then the network starts the discovery of a new route without human intervention. Nodes in the mesh network play multiple roles of sender, receiver and router, to route messages between two devices which are not in direct contact with each other. A cluster tree falls under the category of star-mesh hybrid networks employing larger coverage of area and consuming low power. It employs mobility and flexibility for the networks which may stretch well beyond an area of 30-100 meters. Fig 2.1 shows star and mesh topology.

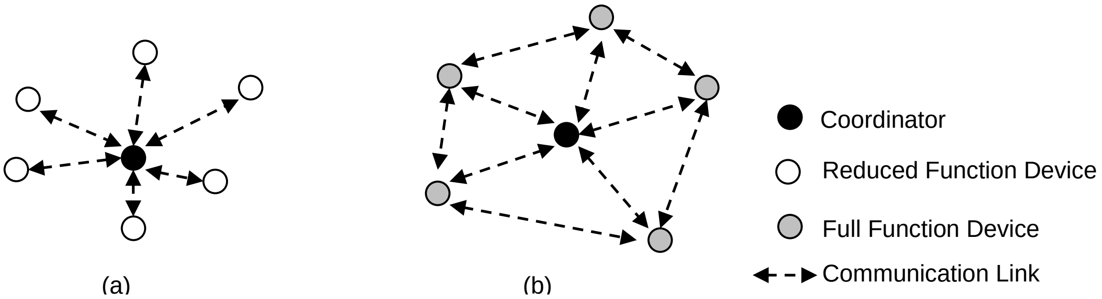


Fig 2.1 (a) Star topology (b) Mesh topology

There are a lot of benefits of using a wireless mesh network which often make it the prime choice for the networking of lower power sensor networks. Because the number of sensors used in a sensor network is quite high and is spread over a wide area, it is beneficial to use such a mesh network which is flexible and self healing. As the mesh network is an approved network system by the regulation authorities it can be used freely without breaking any law, and covers a range of monitoring applications.

2.2.2 Routing

Routing plays a major role in the case of networks based on mesh topology architecture. Routing is used to convey a message from source to destination. The problem of routing is quite significant in the case of an ad hoc network, where the nodes are mobile. The proper selection of a routing algorithm is important to balance the energy spent in updating changes in the topology and the paths and a low duty cycle for longer battery life. There are a lot of routing protocols fulfilling these needs which are used in such networks. Some of these are destination-sequenced distance vector protocol (DSDV), temporally ordered routing algorithm (TORA), the ad-hoc on demand distance vector protocol (AODV) and others [Call04b] [RVB01]**.** Some of these are described in section 3.2. There are different issues involved while we deal with routing protocols so our main goals while choosing a routing algorithm for a sensor network are described below [Stoj01]**.**

2.2.2.1 Minimizing the energy consumption per routing task

There are various metrics used when a network chooses a particular path for routing a message between two points. Hop count is usually used to find a cheaper path, this being a constant metric. However, if transmission power is used as a metric then power may be saved, as it would directly involve the distance between the two nodes, inferred with the help of received signal strength. Hence, knowing the distance the network can pick single hop nodes at greater distance resulting in a lower hop count path than picking closer nodes, which would result in using more intermediate nodes.

2.2.2.2 Loop-freedom

The routing protocol should be loop free so that route requests do not keep circulating between nodes after timeout and cause energy wastage. The network can be made loop free by adding feature such as “freshness level” of the request or update. This will indicate the nodes to drop updates or ignore them if they have received some other routing update with greater freshness level in the past.

2.2.2.3 Minimizing the overhead

A large amount of energy is spent in proactive routing algorithms in the form of communication overhead, where routing tables are maintained in advance. This needs to be minimized, which can be done with the help of reactive routing protocols, in case of networks which are energy constrained.

2.2.2.4 Memory usage

The use of memory should be minimised as the energy used in memorising an old route and processing it can use a considerable part of the battery life. Therefore, routing protocols which select a fresh route with every task are preferred in the case of low power radio networks.

2.2.2.5 Multiple routes for a single task

Sometimes the algorithms follow a single route to forward the data to the destination, which raises the problem of increased energy loss by the nodes which are involved in the route and consume more energy as compared to nodes which are not in route, resulting in their premature failure. So the algorithm should ensure multiple routes for a single task and also ensure involvement of a greater number of nodes to pass the message, helping them to share the burden of energy consumption, resulting in longer network life.

2.3 Data link and Physical layer concerns

2.3.1 Medium Access

Medium access is a significant problem of low power radio networks because of the large size of the networks and lack of time synchronization between the nodes. Most of the nodes are inactive most of the times which makes synchronization difficult. Besides this there are various other factors which a good medium access control system should posses, including scalability, channel utilization and others [Heid01] [Tils].

2.3.1.1 Scalability and adaptivity

A good MAC protocol needs to be highly scalable because sensor networks usually need the addition of nodes as per the needs of the application. It should also be adaptive to changes in the topology, and to the various energy saving modes the nodes are operating in. In sensor networks new nodes may be introduced and old nodes may die, so a medium access protocol should be capable of adjusting itself according to the conditions.

2.3.1.2 Channel utilization

Although the data rate in low power radio networks is quite low as compared to other radio networks, it is still necessary that the channels are properly used. For example, in the case on an event where the associated nodes try to forward data to the base station, there can be a lot of nodes trying to forward data at the same time which can trigger the problem of collisions. Good Medium Access Control (MAC) should try to maintain the maximum possible data rate with the help of maximum possible nodes using the same channel while avoiding collisions.

2.3.1.3 Throughput

Throughput is the amount of data successfully transferred between the sender and the receiver in a given time. The throughput depends upon the specific application. In sensor networks the throughput rate is usually low but the there are some other factors like collision avoidance, channel utilization and control overhead which can affect the throughput. A good MAC protocol should be efficient in handling the problem of collisions and ensure less control overhead, which can otherwise consume a lot of energy at the cost of data throughput.

2.3.1.4 Latency

Latency is the time taken in delivering the data between two devices. Sensor networks have a high latency because the primary goal is to maximize the network life. However this latency can further increase if there are frequent collisions and nodes have to contend for a long time before sending data. A good MAC protocol will ensure a collision free network with fast channel access to the nodes waiting to forward data.

2.3.2 Security and Robustness

As previously stated sensor networks or low power radio networks often operate in 2.4 GHz ISM band, which is used all over the world. It is a license free band and can be used by anyone for networking. Because of this there is always a possibility of user data being tapped or manipulated by other users. This issue is the most important issue while using such low power networks. Also, there is a need to make sure that the transmitted data has reached the destination and received data is uncorrupted. There are a series of steps being taken and techniques incorporated in the network systems which are rendering such networks more secure.

Along with this, robustness of a network also depends upon the successful transmission of data over the medium. A few steps for making a network secure and robust are described below [IEEE03c].

2.3.2.1 Frame acknowledgement

An acknowledgement frame following a data or MAC command transmission can confirm its successful delivery. And if the acknowledgement is not received within a specified time interval then the transmission is considered unsuccessful. In such a case the sender will re-transmit the data packet to complete the transaction. Hence the use of an acknowledgement frame can make the network more reliable.

2.3.2.2 Data verification

Once the data has been received by the destination node it has no means to know whether the data received is free of errors or not. To counter this problem some sort of error detection code can be used which can check the validity of the received data. The most common and successful error detection code used in low power radio networks and approved by the International Telecommunication Union is cyclic redundancy check (CRC).

2.3.2.3 Access control

With many users using the same frequency band, the data can be accessible by a third party which sometimes proves fatal. Access control is the process through which the user can restrict the number of people without due permission to transmit. This can be done with the help of creating an access control list (ACL) of devices which are allowed to communicate with the particular device.

2.3.2.4 Data encryption

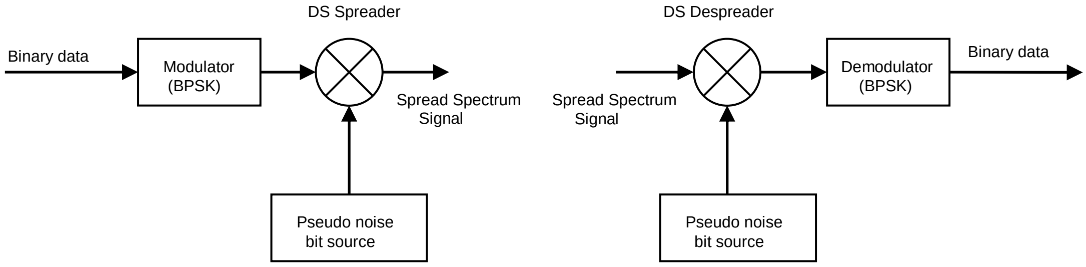
Data sent over the medium can be tapped by a third party and can also be modified. To avoid such a problem the data can be encrypted before sending. The device having the cryptographic key can read the data. Thus a significant level of security can be achieved.

2.3.3 Power

The efficient use of available power is a big issue in low power radio networks. Low power radio networks are employed with the help of sensors which are energy constrained and can not maintain a long life if energy saving measures are not taken. Sensors are usually deployed in regions where frequent human access is not possible and sensors can not be mains powered also. Periodical energy renewal is also not possible so a good network needs to implement techniques to extend the life of the network. Some sensors scavenge energy from some thermal effects or from solar energy. With all these considerations, the aim is to optimize battery life with efficient use of resources. Besides this, power transmission limits are imposed by the regulatory bodies so as to avoid any sort of interference with other networks. The maximum transmit power should conform to the local regulations and should not exceed the limit. The limits however are different for the networks operating in different frequency bands. Hence such transmit power considerations put further pressure to employ accurate and sensitive transceivers in the networks.

2.3.4 Modulation

According to IEEE 802.15.4 standard, two types of modulation techniques are used in low power radio networks which depend upon the frequency band in which the network is operating. 868/915 MHz frequency bands use Binary Phase Shift Keying (BPSK) while the 2.4 GHz frequency band uses Offset Quadrature Phase Shift Keying (O-QPSK) [Gut03d]. But this is not enough if the data has to be secured and the energy has to be saved. Further there is always a problem of noise and interference from other users. For this, both frequency bands employ a spread spectrum technique on top their respective modulation schemes. This spread spectrum technique is called Direct Sequence Spread Spectrum (DSSS) [DM03] [ECE4321].



(a) Transmitter (b) Receiver

Fig 2.2 Block diagram of DSSS transmitter and receiver

DSSS is the technique which is used to increase the bandwidth of a transmitted signal. Therefore the resultant signal after the DSSS is a signal with much greater bandwidth and lower spectral density. In DSSS the pseudo-random number (PRN) is multiplied directly by the data entering the carrier modulator. This increases the bit rate of the entering data because of the chip rate of the PRN sequence. Therefore the resultant of the modulator is a spread spectrum with wider bandwidth in ratio with the PRN bit rate. When the signal is received it is demodulated with the help of same PRN sequence. Hence the original signal can be recovered. The spread of signal in a wider bandwidth can be seen as wastage of bandwidth but this loss is compensated with the use of same bandwidth by multiple users. The transmission is interference free because of the unique PRN sequences used by different stations, allowing them security and privacy. The other advantage of DSSS is that when the signal is spread then the spectral power density of the signal decreases hence it appears as noise on the receiver. This characteristic of DSSS is widely used in defence communication systems where information is prone to be tapped.

Chapter 3

Development in Low Power Radio Networks

3.1 Energy Saving Techniques

The invasion of sensor network technology in different spheres of life and its use as a solution for a wide range of data gathering applications has immensely increased the challenges for research. Sometimes these networks may consist of a few hundred to thousands of sensor nodes gathering data. As the basics of the sensor network lies in its low power consumption and long battery life, there has been an increasing amount of pressure on the researchers to devise and implement more and more power saving techniques to realise the aim in the best possible way. As there is no scope for energy renewal for long periods, these energy constrained networks need to be modified to get sufficient data without making too much compromise with the energy consumption, event frequency and desired output quality [Hac03a]. As some sensors used today are even extracting energy from environmental sources by means of energy scavenging, so there is an urgent need for efficient use of available resources. There are a lot of schemes used today to achieve the target of low power consumption, including coordination to minimize duty cycle by using an adaptive MAC, adaptive topology and routing and in network processing by using some data centric routing procedure [Estrin02]. Therefore predominantly there are two major areas through which significant energy conservation can be realised: using the MAC layer to turn off the radios while they are idle, and in-network processing through the use of appropriate routing.

It has been noted that much of the energy in sensor nodes is wasted if they remain in idle mode without reception of any data or while waiting for data. This power consumption is approximately the same as that consumed during transmission and reception, typically 10-15 mW. The power consumed during sleep mode is approximately equal to the energy consumed by the sensor or it's CPU, which is typically below 5 mW or less [Estrin02]. So a major wastage of energy lies in idle listening when no event happens, along with other factors such as collisions, control overhead and overhearing which are common to all the wireless networks. Some energy saving techniques are described below.

3.1.1 Limiting Radio Operation

Switching off the radio when idle or while there is no data transmission can save a lot of energy in multi-hop networks. This is done by introducing periodic sleep cycles in the frame structure. In this case, the sensor listens for the data for a short interval and goes to sleep after that. However this mode can leave certain areas in the network which would become inaccessible, creating a disruption in the smooth functioning of the network. To counter this problem, a technique called *adaptive duty-cycling* [Gan03] selects the nodes to be powered down after careful consideration of their available energy level and their coverage area, while ensuring that the continuity of the network is maintained. Another technique to solve this problem lies in a procedure called *Wake on demand*, in which the nodes use multiple radios. A low power radio is used to wake a high power radio whenever there is a need for data transfer. Fig 3.1 shows the superframe where sleep cycles are introduced. The device wakes up periodically and listens for data and sleeps again. This saves significant amount of energy because radios are off during sleep cycles.

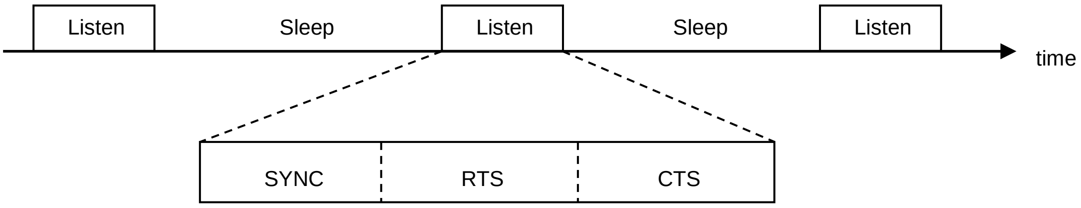


Fig 3.1 Low-duty-cycle operation

3.1.2 Data Management

Sensors gather a lot of data during their operation, which typically is a lot more than the data that can be processed by a single node at the destination. This bulk of data can shorten the life of a sensor to a few weeks which actually is meant for years. Hence the data needs to be processed in the node itself before it is delivered to the next node. This can significantly decrease the load and energy consumption. In some cases, sensors are even required to send only specific data or required data rather than the whole data collected over different time intervals. Thus such in-network processing can save a considerable portion of energy.

3.1.3 Collision and overhearing Avoidance

Collision and overhearing are common problems faced by all wireless networks. These effects can be reduced by collision avoidance and overhearing avoidance procedures. Collision avoidance is used because of the presence of multiple users using the medium at the same time. In collision avoidance an RTS/CTS exchange technique can be used to render the transmission collision free. RTS stands for "request to send" and CTS stands for "clear to send". A sending device has to take the permission of the receiver first, before sending any data packet. Therefore, before every transmission, the transmitting device sends a short RTS frame to the receiver to check if it is ready for the reception. If ready, the receiver can acknowledge this by sending a CTS frame signifying its ability to receive the frame. These frames also carry information about the time for which the present transmission will continue. When other nodes receive this packet they record this time in network allocation vector (NAV). NAV is a variable associated with each node and it can be used as a timer to evaluate when a node can transmit. Once a node records the transmission time of other node in its NAV, it starts decrementing this value until it reaches zero. When a node has to send data, it first checks the NAV and if it is non zero, the node assumes that the medium is busy. Hence a considerable power is saved and collisions can be reduced [Wei02].

Similarly, overhearing avoidance techniques also uses RTS/CTS packets to sleep during transmission of packets intended for other nodes. Since data packets are much longer than the RTS/CTS control packets, this can save a lot of energy.

3.1.4 Message passing

Message passing is yet another technique for energy conservation. In this case a long data packet is fragmented into smaller packets before transmission but only one RTS and one CTS is used. However every time a data fragment is transmitted, the sender waits for an ACK and if it fails to get the ACK, re-transmits the fragment. This save the energy used in sending the whole packet again. Each ACK packet also contains a duration field, so any neighbouring node which listens to the ACK packet can go to sleep for that particular period and increase the period if a data fragment is retransmitted.

3.1.5 Routing

Routing plays a major role in network communication and if used appropriately can save power too. There are two special features of routing schemes that are used in sensor networks. First, they are attribute-based and second they are energy constrained and dynamic. Attribute-based in this case means a routing protocol where data is recognised by fixed attributes such as location, duration, type and others. Therefore devices don’t need to keep a table of addresses of other nodes to track down data. A simple request containing attributes such as location and type can help the nodes to gather information about the event. Also, because they are required to operate in a relatively small area they just need local information instead of big routing tables. Hence less energy is spent on maintaining routing tables and storing them.

3.2 Routing Techniques

Several routing protocols are used in network systems to evaluate the path between the source and the destination. The use of a particular routing algorithm in a network depends upon factors such as density of nodes in the network, distance between the nodes and the energy constraints of the nodes. Depending upon these factors, routing algorithms are divided into three types; *Proactive*, *Reactive* and *Hybrid* routing protocols [Jin04].

3.2.1 Proactive routing

In a proactive routing protocol all the routes are calculated before the actual traffic is sent. Whenever a route is needed to send data from a source to the destination then it is picked from the pre-calculated routing table. In this case the network updates the routes regularly to trace any failure and maintain the route or replace the route with a fresh route. Thus the data can be sent with little network activity. An example of proactive routing is *Destination-Sequenced Distance Vector (DSDV)* [Jan03] [Rab00].

3.2.1.1 Destination-Sequenced Distance Vector (DSDV)

In this routing algorithm each network node maintains a routing table in its memory. This routing table contains the list of all the available destinations with hop count to each destination. To maintain the freshness of each route, all the routing entries are tagged with a sequence number which is originated from the destination. The nodes also keep transmitting periodic updates to advertise their routing table to neighbouring nodes. If a node wants to send the data to a destination then it checks its routing table for a routing entry for that particular destination. If it has an entry then it sends the data through that route. If the source has more than one entry to that destination it looks at the sequence number associated with that entry and picks the route with the highest sequence number. In the case that multiple entries have the same sequence number then it looks at the hop count which counts the number of intermediate nodes between the source and the destination, and picks the one with the lowest hop count. The algorithm is quite efficient for data transfer, but a lot of energy is consumed in maintaining the routing tables.

3.2.2 Reactive routing

In reactive routing the network discovers the route to the destination only when needed. It does not maintain any sort of routing table in advance. As there are no updates, it has a considerable amount of data overhead associated with transmission. It is preferred in case of a network with a small number of nodes communicating less frequently for which maintaining routing tables would be waste of energy. Examples of reactive routing are Ad-hoc On Demand Distance Vector Routing (AODV) and Dynamic Source Routing (DSR).

3.2.2.1 Ad-hoc On-demand Distance Vector Routing (AODV)

In this routing algorithm the nodes do not maintain the routing table in advance but start finding the route on demand. Nodes not in the path of the desired route do not maintain any routing information, however local connectivity is achieved by exchanging *Hello* messages for the purpose of establishing a quick link in case of a request. AODV also follows the concept of sequence numbers to maintain the freshness of the routes by monotonically increasing the sequence numbers of fresh routes. Whenever a source wants to send data to the destination, it sends a route request (RREQ) message containing information about source address, destination address, hop count, broadcast id etc to a neighbour which already has an entry in its routing table. Upon receipt of the request the node sends back a request reply (RREP) to the source and then broadcasts the RREQ to its neighbour with increased hop count. This process goes on until the destination is reached. If a node receives the same request again it checks the broadcast id of the request and if it is the same as that of the previous request, drops the request. As the RREQ progresses from one node to another it creates a reverse path until it reaches the destination. Once a reverse path is set up the source can start sending the data through that route. But meanwhile, if the source comes across a path with a route with a lower hop count then it will start sending data through new route. Fig 3.2 show the reverse and forward path formation used in AODV. Fig b shows forward path formation when the source sends a request to its neighbours who send the request further until it reaches the destination. Before sending the request further each node records the address of the neighbour from which it received the first request. In this way the reverse path is set up as shown in fig a. A route which can not trace the destination or which is not along the path determined by the request is deleted after ACTIVE_ROUTE_TIMEOUT as shown in fig b by “Time out”.

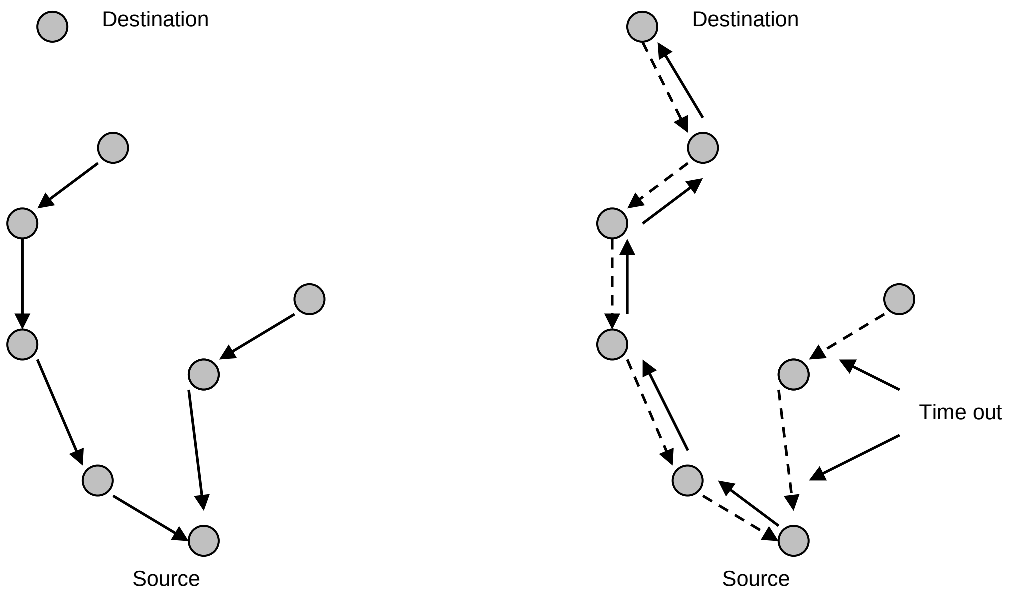


Fig 3.2 (a) Reverse Path Formation (b) Forward Path Formation

3.2.3 Hybrid routing

A hybrid routing protocol is a combination of the reactive and proactive approach where the data is not accessed through some address but through some query or information related to a particular location. This approach bypasses the need for creating and maintaining routing table entries. As sensors are used to gather information relating to a particular area or a particular type so the protocols developed are more data centric. There are a lot of protocols developed under this thinking, for example, *Directed Diffusion*, *Geographical Routing*.

3.2.3.1 Directed Diffusion

In Directed Diffusion, data is named by attribute-value pairs and the data query is referred as “interest” [Int03]. The interest propagates to different neighbouring nodes and nodes with matching data send the data to the sink or the source of interest through a reinforced path. This data can be modified by intermediate nodes and can make it more accurate.

The main elements of Directed Diffusion are:

- Interests
- Data messages
- Gradients
- Reinforcements

Interest: A query for the data needed by the user.

Data message: Collected information in response to the interest. It can be an event or a sensed phenomenon.

Gradient: The backward path created by the nodes which receive the interest. It is created to draw back the data from the destination towards the interest generating sink.

Reinforcements: Reinforcement is the process of strengthening one particular path for further data exchange after getting replies through different paths.

3.2.3.2 Example:

Let us take an example of an interest tracking a vehicle.

```
type = wheeled vehicle                    //vehicle type
interval = 10 ms                          // send event every 10ms
duration = 10seconds                       // send for next 10 seconds
rect = [ 10, 40, 30, 60 ]                  // area of consideration
```

A possible data message could be as shown below:

```
type = wheeled vehicle                     // type of vehicle
instance = car                             // instance type
location = [15, 18]                         // node location
```

intensity = 0.7 // signal amplitude measure

confidence = 0.75 // confidence in the match

timestamp = 02: 34: 45 // event generation time

3.2.3.3 Interest propagation procedure

During interest propagation the sink injects the interest in the network and it is diffused to other network nodes. This interest includes attributes such as type, rect, duration and interval. “rect” in this case stands for rectangle or area of concern. The interval attribute indicates the event data rate. This task state is purged from the sink node after expiry of the period indicated in the duration attribute. Initially, when the interest is broadcast by the sink it has a higher interval attribute than what is really required. This is just to determine if there are any nodes satisfying the broadcast interest. This initial interest is called *exploratory* interest.

The interest sent by the sink is considered a soft state because it is not sure at this stage whether the interest is properly disseminated throughout the network. To ensure this, the sink periodically refreshes the interest with a monotonically increasing *timestamp.* Every node has its own interest cache which contains different interests. Two interests are considered distinct if they differ in their *type* attribute or *rect* attribute. Every interest has several fields. For example, the *timestamp* field indicates the timestamp of the last received matching interest. The gradient attribute specifies the neighbouring nodes with the data rates they are providing which are derived from the *interval* attribute. A duration field is derived from the *timestamp* and *expiresAt* attributes. "expiresAt” attribute gives the time instant at which an event finishes. When a node receives an interest entry, it checks if there is a matching entry in its cache and if not creates a new entry with that interest. Fig 3.3 shows the various steps involved in directed diffusion algorithm. Fig (a) shows the interest propagation through different routes from sink to the source. Fig (b) shows the gradient set up to draw data from source to the sink and fig (c) shows the reinforced path which indicates the route which can transfer the data from source to sink in least time.

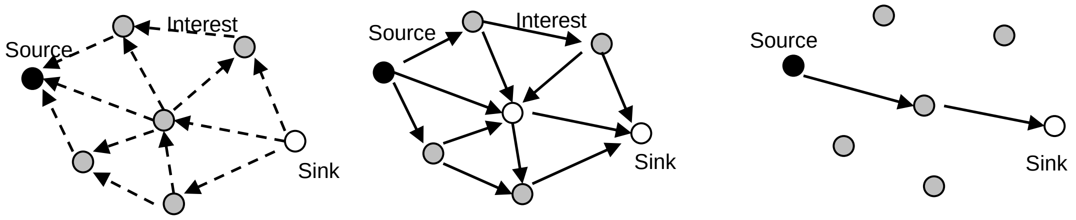


Fig 3.3 (a) Interest propagation (b) Gradient set up (c) Reinforced path

3.2.3.4 Gradient establishment

When an interest is propagated into the network, all the nodes set up a state which can be used to pull down the data from the destination to the sink node. As nodes don't have any information about the sink, a node only creates a reverse path to the node from which it has received the interest. Thus a gradient is set up which has information about the data rate and the direction in which it has to send the data.

3.2.3.5 Data propagation

Data propagation is the process in which the data is sent from the destination to the sink once an event is detected. Whenever an event is detected the node looks at its cache to see whether it has an interest relating to such an event. If an interest has the same "rect" and same "type" as that of the event it is considered a match. Then the node checks the highest requested data rate among all its outgoing gradients, and picks the one with the highest data rate and instructs its processor to generate data at that highest rate. The source node sends an event description every 10ms to all the nodes from which it has a gradient as agreed in the interest. The same process is repeated when the data reaches the next node until the data is delivered to the sink. At every step the same algorithm is run and if no match is found then the data message is dropped without notice to anyone. If a node gets a data message with higher data rate than what it can provide then it can down-convert the data rate.

3.2.3.6 Reinforcement

When a sink sends an interest it sends it with low data rate so as to let the interest propagate throughout the network. This is called exploration. But once the data is received from the event generating node, several paths are used to deliver the data to the sink. The sink then reinforces a particular path with a higher data rate to draw down the data.

3.3 Topology Discovery

Topology discovery is used in sensor networks to construct the topology of the network with respect to a single node [Hac03b]. Three main steps of topology discovery are:

- Initiating the process by sending a topology discovery request
- Spreading the request throughout the network
- Setting up a response action to get the topology information

The initiating node sends a topology discovery request to its neighbours and upon the reception of the request these active nodes send back a response with information about their topology and their respective neighbours. There are different approaches followed with respect to the topology discovery procedure. These have been demonstrated here with the help of a simple example, as shown in fig 3.4.

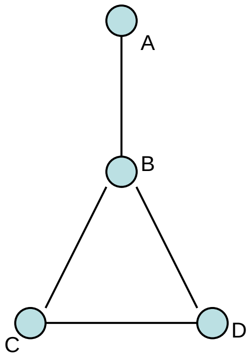


Fig 3.4 Simple topology discovery diagram

3.3.1 Direct response approach

In this type of approach the nodes receiving the request from the initiating node reply back to it directly without gathering topology information about their neighbours. The steps followed in this approach are as follows:

- Node A sends a topology discovery request;
- Node B receives the request and it replies back to the initiating node A;
- Node B forwards the request to node C, which replies back to node B which in turn forwards the reply to node A;
- Node B forwards the request to node D, which replies back to node B which in turn forwards the reply to node A;
- Node A gets the information about the entire topology.

3.3.2 Aggregated response approach

In this type of approach, the active nodes first gather information about their neighbours by sending the request forward and then send the aggregated topology information through a reply to the initiating node. The steps followed in this approach are as follows:

- Node A sends a topology discovery request;

- Node B forwards the request to its neighbours;
- Node C and node D get the request and reply back to node B;
- Node B considers node C and D as its child nodes and aggregates the topology information before sending it to node A;
- Node A gets the information about the whole topology.

3.3.3 Clustered-response approach

In this type of response the network is divided into clusters and the each cluster has one cluster head or coordinator. Each node is associated with at least one coordinator and it’s only the coordinator which can generate the response. With this approach the node aggregates the information before replying back to the initiating node. The steps followed in this approach are as follows:

- Consider node B as a coordinator and node C and D are its child nodes in the cluster;
- Node A sends a topology discovery request;
- As only coordinator can reply so nodes C and node D do not reply;
- Node B sends a reply to node A;
- Node A receives no information about the link between C and D.

3.3.4 Cluster formation in sensor networks

Sensor networks are networks consisting of a large number of energy constrained sensors with the objective to maintain a lifetime of several months to a few years. Hence a low overhead is required in such networks which can be achieved by minimizing the number of clusters. This objective can be achieved by solving the two problems of minimal set of clusters sending replies and a minimal cluster tree with a minimal set of cluster heads.

There are two different node “colouring” schemes used to discover the minimal number of cluster heads. The first scheme is the three-colour scheme and second is four-colour scheme [Hac03c].

3.3.4.1 Three-colour scheme

In the three-colour scheme, the nodes are given three colours according to their status in the network. These colours are black, grey and white. In this topology discovery procedure all the nodes are white initially. A node which starts the topology discovery or the node which first sends the topology discovery request becomes a black node indicating a coordinator. Subsequently any other white node receiving the topology discovery request packet from a black node becomes a grey node and disseminates the discovery packet further. Any white node receiving a discovery packet from a grey node becomes a black node, in other words a coordinator. Hence in this case, a coordinator is always a single hop away from another coordinator.

More details about the three-colour topology discovery scheme can be found in chapter 7.

3.3.4.2 Four-colour Scheme

The Four-colour scheme is potentially more efficient than the three-colour scheme. The four-colour scheme covers more area with fewer coordinators. In the four-colour scheme there are fours colours assigned to different nodes according to their status in the network. In this scheme all the nodes are white initially and a white node which starts the topology discovery by sending a topology discovery request becomes black or a coordinator. When white nodes receive this topology discovery request packet, they turn into grey nodes and further disseminate the request packet. The white nodes receiving the request packet from the grey node turn into dark grey nodes and again disseminate the request further. In the last step, the white nodes receiving request packets from dark grey nodes becomes black or coordinators. Thus in this case, two coordinators are always two hops away from each other, ensuring larger coverage area with fewer coordinators.

More details about the four-colour topology discovery scheme can be found in chapter 7.

Chapter 4

Technologies

With the popularity of wireless communication and mass acceptance of wireless technologies, the development work in these areas has received a push to provide better standards to customers. Various standards have evolved in the past years keeping in mind the needs and different applications. Among all these, the technologies which have received wide acceptance and popularity are;

IEEE 802.11 Standard (Wi-Fi) (Section 4.1)

Bluetooth (Section 4.2)

IEEE 802.15.4 Standard (Section 4.3).

4.1 IEEE 802.11 Standard (Wi-Fi)

4.1.1 Introduction

Wi-Fi is a short for Wireless Fidelity. It is a network technology conforming to the IEEE 802.11 standard [IEEE99] which provides wireless connectivity to various devices. It is a sort of WLAN – a wireless network operating in a short range for example in a small building, an office or in a house. For example a laptop or a device with a wireless interface card can remain connected to a network while moving in an area described as the coverage area of the base station. The base station is an "Access Point" which is connected with a wide area network such as internet which in turn can give wireless internet access to the device in its range. Wi-Fi provides connectivity between different appliances such as television, computer, printer and other electronic devices in the work place and home environments. For a WLAN to be called "Wi-Fi" and carry a WiFi logo, it should confirm to the standards maintained by the Wi-Fi Alliance[*] which formerly was called Wireless Ethernet Compatibility Alliance or WECA [WAlli]. The Wi-Fi logo is shown in fig 4.1.1. This standardisation not only outlines the basic idea of the technology but also ensures greater interoperability of the equipment produced by different manufacturers. It also tracks standards developments and enhances interoperability testing to reflect advances in the field.

[*] The Wi-Fi Alliance is a nonprofit international association formed in 1999 to certify interoperability of wireless Local Area Network products based on IEEE 802.11 specification. Currently the Wi-Fi Alliance has over 200 member companies from around the world

Figure 4.1.1 Wi-Fi logo

Wi-Fi creates a flexible environment where devices can move and can accommodate new devices in the network quickly without much hassle and without installation of new wires. It enables the client to work at remote places with the help of “Hotspots”[†] anywhere ranging from coffee shop to a library or a park without any sort of obstruction. It also allows mobile working using a PDA or other mobile device to constantly remain in contact with the company network and transfer information to the centre.

4.1.2 SETTING UP A Wi-Fi NETWORK

It is very easy to install a Wi-Fi network. A fundamental part in a Wi-Fi network is an Access Point or an AP which can be regarded as the heart of the network. If it stops working then the network would also die. In this respect it resembles a star network although all the other devices can communicate without sending the message via the access point. It can also be called a “base station”. The radio waves from the AP are broadcast and move in all the directions like the light from a bulb. As the distance increases the signal strength decreases so there is a need for suitable placement of the AP such that it covers maximum area and provides connectivity to greater number of devices. This decision is influenced by a number of factors including the number of users, type of the devices, application environment and the area available. It needs to be ensured that if the device moves it remains connected and maximizes the support to mobility. A very basic Access Point is available in the

[†] “Hotspots”- public access points owned by private companies or individuals for providing paid mobile internet access to the people

market for less than 100 pounds and can support up to 256 users depending upon the specification. Any computer or device equipped with a Wi-Fi network interface card (NIC) can establish a link with the base station. The range of the access point is between 150-500 feet and can generally penetrate walls, floors and windows [WFSO4]. Wi-Fi supports data transmission at a rate of 11 megabits per second which is comparable to first generation Ethernet connection. A single AP is sufficient for a home application while two or more may be used for a corporate level application, providing greater coverage.

4.1.3 Wi-Fi MARKET:

The rapid increase in popularity of Wi-Fi has left the market and homes flooded with Wi-Fi products. Every sphere of life is highly influenced by the technology because of its functionality, flexibility and low installation cost. Within a very short time Wi-Fi has made a big market for itself. Be it business houses or corporate offices, homes or public points, libraries or infotainment parks, there is hardly any sphere of life untouched by Wi-Fi technology. Its consumer base is increasing day by day. As Wi-Fi is becoming a standard in laptops it is expected that by the end 2004, 45 million laptops will be equipped with Wi-Fi technology and millions of people worldwide will use Wi-Fi devices regularly [Sen02]. Some major market areas are discussed below.

4.1.3.1 Home Applications

Wi-Fi is finding high growth in the home market for connecting home computers and consumer electronic appliances through wireless networking. It helps consumers to create networks without having to spend time and money to go through the pain of installing new wires. Through this one can use a single internet connection to multiple computers with the help of access point. The access point is hardwired to the internet and further allocates network addresses to the subordinate devices. Various products in the market are introducing innovative new applications. Certain manufacturers are coming up with different networking devices that can help clients to remotely control their home appliances such as security system, lights, car locks and heating system. It can enable one to control an oven while sitting away in the bed room or at work.

4.1.3.2 Public Applications

The most common form of public application of Wi-Fi is in the form of public hotspots to enable people travelling around the world to get internet access outside home and office. One can find such hotspots in airports, public places, shopping centres etc. According to one forecast, by 2007 the WLAN market in US is going to expand to 21 million users which would bring 3 million US dollars in revenues [Pao02]. People are using the idea of Hotspots to boost their other business concerns, for example some fast food outlets have installed hotspots in their premises so that people coming there for internet access can also enjoy their food. Due to increasing market interest this area is becoming more competitive with better services to the client.

4.1.3.3 Corporate Applications

Big corporations are using Wi-Fi to enhance connectivity between employees to ensure better performance and increased productivity. Today the market has some devices which can help workers to check a product in the warehouse and communicate with the ordering systems. It significantly increases the performance of the corporation by increasing productivity because of better connectivity between the workers by constant interaction on the network or through e-mail. Being wireless it is easy and less expensive to deploy in an industry environment without much alteration, which in a hardwired network can be very cumbersome and expensive. It is highly scalable and can coexist with wired network also.

4.1.3.4 Academic Applications

Wi-Fi in the academic level is attractive in order to help students to get help through network points in the campus without staff intervention. The same points can also be used by the students to access university information and library facilities which can raise the level of information awareness among the students. Having seen all these benefits of the technology at the academic level, these students become the future Wi-Fi users, with a full understanding of the capabilities of the technology.

4.1.4 Various Standards

Wi-Fi network standards are governed by WECA which is a non-profit, technical professional association. It is a group which comprises of over 140 vendors including the representatives of the world's leading technology and consumer product companies, e.g. Agere, Cisco, Conexant, Dell, Intel, Intermec Technologies, Microsoft, Nokia, Philips, Sony and Symbol Technologies [WAlli]. The IEEE 802.11 standard is currently in its seventh version after adapting to different changes and making essential modifications to make it more secure and robust [WWPO3]. There are a lot of standards but all are not compatible with each other. Care should be taken while buying some Wi-Fi product to make sure that it is compatible with any Wi-Fi product already installed. Principal technical standards are 802.11a, 802.11b and 802.11g [WFSO4]. Standard ‘a’ and ‘b’ are not compatible but as ‘g’ is an advanced version of ‘b’ it is even compatible with it.

4.1.4.1 IEEE 802.11b

This is the most common Wi-Fi standard which is widely used these days. It operates in the 2.4GHz frequency range and can give data throughput up to 11 Mbps which is quite comparable to standard Ethernet connections. But since a single channel is used by multiple users, they may see lower than 11Mbps. The distance between the device and access point can also affect the reception of the signal and larger distances can affect signal quality. It uses Direct Sequence Spread Spectrum Technology (DSSS). It is compatible with a newer version which is 802.11g. As it is used in same frequency band as Bluetooth, cordless phones and microwave ovens, it can experience some interference from them

4.1.4.2 IEEE 802.11a

802.11a is a standard which operates in the 5GHz frequency band and can give a data rate of up to 54Mbps. It is exceptionally fast and its high bandwidth can be shared by multiple users while maintaining fast access. It uses Orthogonal Frequency Division Multiplexing Modulation Technology (OFDM). It has several advantages and several disadvantages. Some of the advantages are that it can support high throughput. It has a larger number of channels associated with it so the chances of interference decrease. Another factor which is responsible for less interference is that it uses the less crowded 5GHz band rather than the 2.4GHz band. Its disadvantages include low range as it operates at high frequency so any obstruction in line of sight can affect the signal. Moreover because of low coverage area more Access Points need to be installed to cover a particular area for seamless connection.

4.1.4.3 IEEE 802.11g

802.11g is an extended version of 802.11b and it can support a high data rate of up to 54 Mbps operating at 2.4 GHz frequency range. Although it has the same data rate as that of 802.11a, it does not have as many channels so it is more liable to interference between nearby networks. However it is compatible with the 802.11b standard and can be used on existing equipments.

There are certain other wireless LAN technologies besides IEEE 802.11, for example ETSI High Performance Radio LAN Type 2 also called HiperLAN 2 which operates in 5 GHz band at a speed of 54 Mbps and Multimedia Mobile Access Communication Systems (MMAC) which operates in millimetre wave radio band (30-300 GHz) and can transmit up to 156 Mbps of data [Sen02]. These have not achieved the same market penetration as IEEE 802.11a, b and g. Fig 4.1.2 (a) and (b) shows access points for 802.11g and 802.11b respectively manufactured by Netgear.

Fig 4.1.2(a) WG302 Prosafe 802.11g AP (b) ME 103 Prosafe 802.11b AP

4.1.5 Medium Access in WiFi systems

The technique for accessing the medium in WiFi networks is a form of CSMA-CA mechanism, which is about sensing the medium before starting transmission thereby avoiding collisions. The station trying to send the information first senses energy on the medium and if that energy level is above the threshold level then the medium is considered busy and the station waits until the medium is free. There are two types of medium access techniques used in WiFi networks named Distributed Coordination Function (DCF) and Point Coordination Function (PCF) [Gre02] [IEEE99a].

4.1.5.1 Distributed Coordination Function

In DCF protocol the station uses the information contained in the frames sent by other stations to know when the medium would be free. This information is in the control field of each frame in the form of a duration field which indicates how long the transmitting station will keep the medium busy. For a station to transmit it has to make sure that the time indicated in the duration field has expired. It also needs to check for the busy physical channel. Due to the nature of the DCF protocol, the transmission in DCF is asynchronous which means there is no time constraint regarding data delivery, hence the station is not bound to transmit data within a particular time slot. Therefore the delay between successive frames is random. DCF protocol can have some problems, including RF interference. In this case the source of RF interference can create an illusion of the medium being busy by raising the energy level of the medium. Therefore the stations stop transmitting during that time interval considering the medium busy. This causes significant drop of throughput.

4.1.5.2 Point Coordination Function

PCF is an optional access method for the standard. PCF supports a synchronous mode of transmission. In the PCF protocol, there is a network coordination point which allots the time period to each node operating in PCF mode one at a time. First of all the coordinator node makes a list of all the nodes operating in the PCF mode and then starts polling them one by one for transmission. This period is called the contention free period (CFP); therefore every node transmits in its own particular time slot. Due to this the time delay between successive transmissions is regular hence there is greater synchronization. The standard has made it mandatory to use the DCF protocol in all networks while PCF is optional. The standard ensures that stations can switch their modes between DCF and PCF.

4.1.6 Technical Vulnerabilities

There are certain soft points in the Wi-Fi standards which need to be dealt with before approving as safe and secure technology for everyday use in homes and offices. Security and interference are two major problems faced by the manufactures towards making the technology make an impact on the market. Hackers can access weakly secured networks, accessing important data and paralysing the network temporarily or permanently. Interference can also be a great deal of a problem as signals from two or more access points can interfere with each other and can seriously affect their performance. This becomes worse as more users are using the same 2.4 GHz license free band at the same time.

4.1.6.1 Control on Access

Due to the overcrowded market and usage of license free band, Wi-Fi users are coming across the problem of unauthorised access to their network by hackers. Any device or computer with a network interface card (NIC) can have access to the network if it is not properly secured. It is true that the wireless signals travel a very short distance but this depends upon the conditions and the device used. Such weak signals can actually be picked up by some external high gain antennas at a much greater distance [Sen02]. Even a stolen or lost device from a network can compromise the overall security of network.

4.1.6.2 Data encryption

Data sent via Wi-Fi network can easily be intercepted by a third party so there should be some technique which could help the user to avoid such a problem. This can be helped by encrypting the data before sending. Even encrypted data can be accessed by the hackers but still it gives primary level security to the network. Manufacturers of the next generation Wi-Fi network are spending large amount of money on making such systems more secure and reliable. For example, the Netgear Prosafe Wireless WG302 Access Point has the ability to select a suitable channel for reduced interference and also includes “invisible mode” which can make the Wi-Fi network invisible to unauthorised users.

4.1.6.3 Interference

The use of the license free 2.4GHz band by a large number of users is creating a problem of interference. As more and more users are switching over to Wi-Fi the spectrum is becoming crowded leading to interference in the signals. This problem makes the network slow and unfunctional sometimes. There can be a lot of reasons for the interference, such as cordless phones and microwave ovens operating within range of the base station, and Bluetooth devices operating in the vicinity. As these two technologies have a wide acceptance among the users Bluetooth and Wi-Fi signals create interference. There is no solution to the problem of interference since all these devices are working in the license free bands which gives everyone the right to operate them freely with minor restrictions of maximum coverage area. With the help of new frequency bands it can be possible to divert the traffic into different directions which can solve the problem to some extent. The use of more complex spread spectrum techniques can also help.

4.1.7 Security Mechanisms

There are a lot of security mechanisms available which can effectively address the problem of security and interception in Wi-Fi networks. Because Wi-Fi technology is a big and growing market so lots of big companies like Microsoft and Intel are trying hard to achieve more secure systems in the future. Some of these are discussed below [Sen02].

4.1.7.1 Service Set Identifier (SSID)

This mechanism divides the network into different parts with one or more access points and provides a password to each access point. All the devices trying to access the network must be programmed with the same SSID. Sometimes some devices are even programmed with multiple SSIDs to ensure connectivity with different networks. This is not a very effective tool for security.

4.1.7.2 MAC Address Filter

MAC Address filter serves as a sieve which permits devices with permitted MAC addresses only to communicate with the AP. It can very well serve the purpose of access control and security but complicates network management; since all permitted devices must have their MAC address entered in the AP's database.

4.1.7.3 Wired Equivalent Privacy

This is an 802.11b built in mechanism and encrypts data to make the network secure. In this case all the devices and access points use the same key to encrypt and decrypt the data. In an effort to further enhance security the encryption keys are frequently changed to minimize any chances of trespassing.

Besides these standard mechanisms, there are also some proprietary solutions to the problem of security from Cisco Systems, IBM, Proxim and others.

Table 4.1.1 depicts different technologies with their modulation techniques, data rate, range, frequency band and the network size.

Standard and Market name	802.15.4 Zigbee™	802.11b Wi-Fi™	802.15.1 Bluetooth™
Frequency band	2.4 GHz, 868 MHz and 915 MHz	2.4 GHz	2.4 GHz
Modulation	BPSK and O-QPSK	DSSS	FHSS
Data rate	20 – 250 kb/s	11 Mbp/s	~1 Mbp/s
Network size	Unlimited	32	7
Range	40-50 metres	1 – 100 metres	1 - 10 metres

Table 4.1.1 Different standards and modulation techniques

4.1.8 Future scope

Due to the increasing demand and technical awareness, Wi-Fi is facing a bright future. New standards are being developed by the manufacturers with increased security and higher data rates. As Wi-Fi certified equipments are easily interoperable so more and more companies are trying to produce standard products which can be compatible with future technology. Many companies such as Dell, Netgear, Intel, Cisco, Microsoft are working hard towards the future of Wi-Fi.

4.2 Bluetooth

4.2.1 Introduction

With the merger of technologies of computing and communication it is becoming easier to transfer information in personal computing and communication devices to other devices through wireless communication links. Another approach towards meeting such needs is ‘Bluetooth Technology’. Based upon wireless communication, its main objective is to provide the technology to render systems and networks cable free, at the same time enhancing their connectivity through data transfer or voice transfer. It was evolved keeping consumer requirements in mind rather than just designing the technically best possible radio.

Bluetooth is an open standard to enable wireless communication between different communication devices. Anything inside the Bluetooth wireless personal area network can communicate with other devices in the same range, including personal computer, wireless keyboards, wireless mice, headsets and portable devices. Bluetooth realises the vision of a seamless access to a diverse set of devices in an efficient manner. The technology is named after king Herald Blatand of Denmark between 940 and 985 A.D who is known to have unified the Danes and Norwegians. Similarly the Bluetooth technology unifies the communications between different devices. Bluetooth devices start communicating automatically if they are brought in the range of each other and stop communicating otherwise. The normal coverage area of Bluetooth devices is in the range of 10-100 m [DTIB]. Bluetooth is an open industry standard which is royalty free. The main objective is to make the network more cable free; hence connecting new devices is quite easy, unlike wired network where it can be quite cumbersome. Besides being a cable free network it has some other advantages which makes it well suited for its application, including low power consumption and its ability to support both voice and data. With a range from 10-100 m it uses the 2.4 GHz frequency band, which is available worldwide. This is another advantage of Bluetooth in that all the Bluetooth devices can be used worldwide no matter where it is manufactured and bought without any modification.

4.2.2 Bluetooth SIG

Bluetooth has been developed by a group of companies called “special interest group” (SIG) as an open specification. The aim of the special interest group is to develop the standard and to promote the technology worldwide. The group also works with other standards and regulatory bodies and is responsible for testing and certification of Bluetooth products in the market. The founding companies of the SIG are Ericsson, Intel Corporation, International Business Machines Corporation (IBM), Nokia Corporation and Toshiba Corporation. These companies are also known as *promoter* companies. The SIG was publicly announced in May 1998 with a objective to produce a specification for hardware and software which would guarantee a common platform and greater interoperability. Other companies have also joined the group as *adopter* companies [Miller02a]. These adopter companies also have the right to produce the Bluetooth products without paying any royalty. Because of the increasing membership the group has further divided the responsibilities resulting in the formation of sub groups for different tasks such as air interface group which is concerned with radio and baseband layers, the software group which is focussing on protocol stack, the testing group and legal working group which manages the legal affairs. Fig 4.2.1 shows the Bluetooth logo, which is a registered trademark of Bluetooth SIG.

Fig 4.2.1 Bluetooth logo, a trademark owned by the Bluetooth Special Interest Group, Inc., USA

4.2.3 Usage Scenario [Miller02b] [HJ98]:

4.2.3.1 The Cordless Computer

Bluetooth communication can be an easy replacement for the cumbersome wired networks which are difficult to manage and hinder mobility and portability, so its major use could be in the desktop computer and its peripherals where Bluetooth can replace the wired links. So, one can keep keyboard, mice, joystick, speakers, printer and scanner at any place without requiring wires.

4.2.3.2 Ultimate Headset

Bluetooth technology can be used to keep our link between mobile and headset cable free. These headsets are called “hands free”. Using this, a person can place a call through the handset and start communicating through the headset without physically sticking to the mobile handset and can move about freely. Besides this, because the Bluetooth technology supports multiple devices the same headset can be used by different devices shared by the same access point, for example, for communication with PC and fixed line telephone etc.

4.2.3.3 Three-in-one Phone

The three-in-one phone usage model is an interesting concept which can bring down the number of resources needed to make a person connected to the outside world round the clock. In this case the same phone can be used as a standard mobile phone, as a cordless phone in the home environment through the use of an access point and as an intercom for direct contact with other devices in the proximity without use of a standard landline. This allows a single telephone to be used for different purposes unlike using different handsets in the conventional set up.

4.2.3.4 The Interactive Conference

Because of the wireless nature of the Bluetooth system it can be effectively used in conference rooms where different laptops or desktops can be connected through Bluetooth which will make a small wireless network and can help in faster exchange of data among the members. In an interactive conference room set up one can even exchange business cards and files. In this way the data exchange can take place at the same time as the meeting is going on rather than stopping in between to exchange files or doing it in the end.

4.2.3.5 Hidden Computing

Hidden computing is an exiting application where Bluetooth devices can help a user in his work and can perform a task on his behalf. There can be different modes under which a Bluetooth device can do hidden computing, for example:

- A notebook computer “hidden” in a briefcase can be used to send periodic alerts about the e-mails received on the notebook to a mobile phone. The user can even read them on the mobile phone and reply to them.
- In another case a mobile phone lying in the pocket or a briefcase can be used by an appropriately configured computer or laptop to access the internet, where the mobile will act as an internet bridge

Because of the popularity of Bluetooth technology, the market is flooded with Bluetooth products. These products are built in contemporary mobile phones, PDA’s, PC’s and other PC related products. Bluetooth adapters are available in the market at low costs. A Bluetooth adapter costing below 50 pounds can add Bluetooth functionality to a PC or laptop without any major modification [DTIB].

4.2.4 Main Features

Bluetooth basically is a wireless personal area network where devices can be dynamically added and removed without any major change in the network. The Bluetooth network is called a piconet which can support up to eight communication devices. However, among these devices all devices are not the same. There is one “master” device and there are up to seven “slave” devices. The master is the coordinating device and helps in the proper functioning of the slave devices and also synchronizes them. However, any device can take up the task of master device and can become master if there is any need to do so.

Bluetooth networks used the globally available license free 2.4 GHz band. Because of its availability round the globe, it is advantageous to use this frequency band because it will assure greater interoperability of Bluetooth devices manufactured and produced in parts of the world. So any one can use his Bluetooth device anywhere in the world without any change, no matter where he is. Chief features of Bluetooth network system are [Harte04a];

- It operates in 2.4 GHz ISM license free frequency band.
- There are 79 channels in the spectrum.
- Each channel has a bandwidth of 1 MHz.
- Frequency Hop Spread Spectrum (FHSS) type of modulation scheme is used.
- The range of a particular Bluetooth device is ~10m. However, it can be extended up to 100m by using external power amplifier.
- A network can support up to 8 devices.
- The maximum data throughput rate is ~1 Mbps
- Has low power consumption, typically 0.3mA in the standby mode and 30mA in the transmission mode.

4.2.5 Bluetooth Piconet

A Bluetooth piconet consists of a master and up to seven slaves. In a Bluetooth network whenever a network is set up between two or more devices, then one device becomes master and the others becomes slaves. A master node or device is basically a device which is responsible for the overall piconet functioning and helps devices in Frequency Hopping Spread Spectrum (FHSS) synchronization between themselves. FHSS is a modulation technique which is used to divide the spectrum into different channels and packets are sent over these different channels, hence spreading the message over a wider spectrum [Hod03].

4.2.6 FHSS in Bluetooth

FHSS is used in Bluetooth networks to transmit the signal over multiple channels so as to avoid interference. The messages are first divided into number of packets and then they are sent over the divided spectrum. By using the FHSS method, the available spectrum is divided into different frequency channels and the transmitters and receivers then 'hop' between channels following a sequence defined by the master. Then all the packets are transmitted over different channels. Thus the radio selects different channel for every transmitted packet. The process is repeated thereby spreading the entire message across the available frequency spectrum. The master selects the frequency hopping pattern and synchronizes the communication between the devices. Therefore, all slaves hop together in synchronization with the master. FHSS in Bluetooth employs frequency hopping at a rate of 1600 times per second. Fig 4.2.2 shows how different message packets are transmitted over different frequency channels.

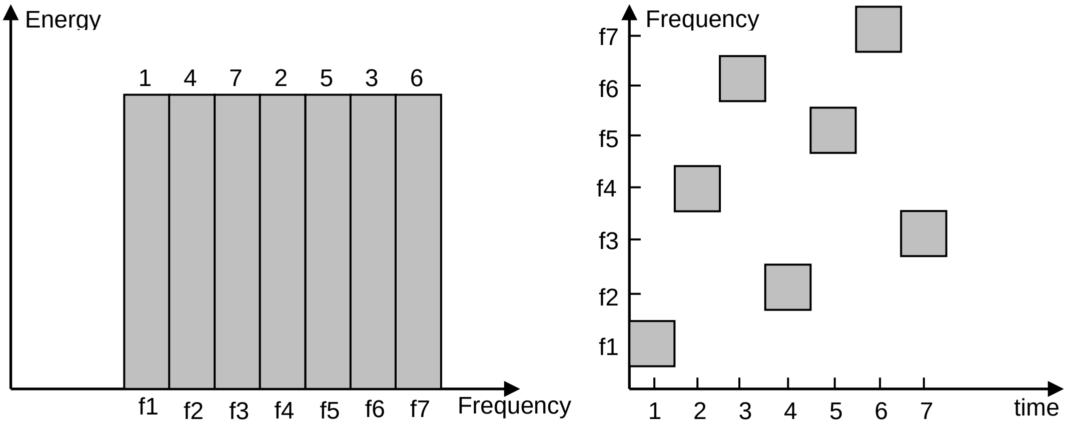


Fig 4.2.2 Packet segmentation and distribution over different frequency channels

The advantages of spreading the message over a wider spectrum are discussed below.

4.2.6.1 Low Interference

As the message is spread over the whole spectrum it picks less interference and in case a particular packet gets corrupted because of interference from other sources, there is no need to retransmit all the packets but the corrupted one.

4.2.6.2 Secure

It is a fairly secure system because the original message can be reassembled only by the receiver knowing the sequence for the frequency hopping spectrum.

4.2.7 Master and Slave

A master device is a device which governs the synchronization of the devices in the piconet and is also responsible for the determination of the pattern of frequency hops. A master device can have no more than 7 active slaves attached to it and can support up to 255 parked slaves. However the roles of the master and slave can be interchanged according to the needs. A device which is master in one piconet can be a slave in another. The communication is usually started by the master once the piconet is formed.

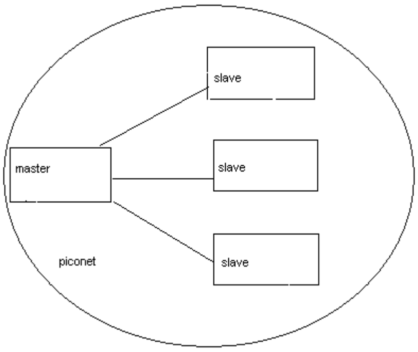


Fig 4.2.3 A Bluetooth piconet with master and slave

In order to save energy, some energy saving modes are introduced in the Bluetooth which also differentiate nodes according to their position in the piconet. These states are discussed below.

4.2.7.1 Active Mode

A device in the active mode is always synchronized with the master and is ready to receive packets from it at any time. It has the fastest response time and also consumes most power because of its continuous listen to the master.

4.2.7.2 Sniff Mode

In this mode the device makes an agreement to receive the packets destined for it at regular intervals. So it sleeps most of the time and wake up periodically to receive packets.

4.2.7.3 Hold Mode

In this mode a device can make an agreement with the master to temporally stop the link between them which can be resumed later. This state allows greater energy saving than sniff state but the response is slow.

4.2.7.4 Parked State

In this mode a device is no longer active in the piconet but still is connected to the master device. Since a master device can support only 7 active slaves, so all the other devices can stay in the parked mode. A master can support 255 parked devices. This mode ensures greater power saving than any other mode.

There may be some devices which are in the proximity of but not connected to the master. They can be said to be in standby mode. Standby nodes are not considered a part of the piconet. Fig 4.2.4 shows a Bluetooth piconet with a master and several slaves. Some slaves are active while other are parked slaves. Fig also depicts the devices which are not actively connected to the master but are in the proximity sphere, which means they can join in if master releases some device already connected in the piconet.

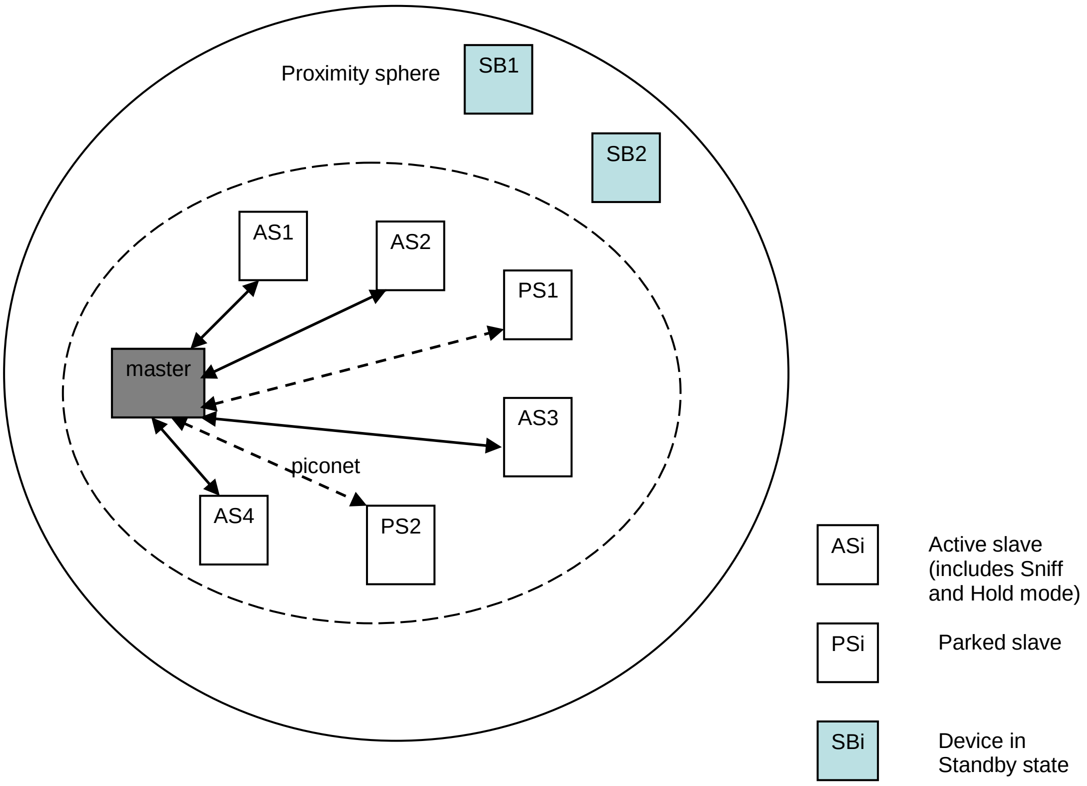


Fig 4.2.4 A Bluetooth piconet with master, active and parked slaves and devices in standby state

4.2.8 Bluetooth Architecture:

The core of the Bluetooth technology lies in its protocol stack, which contains various layers involved in carrying out various functions and ensures accurate transfer of information between two devices. The protocol stack has been divided in to three layers. The lower level protocols or transport protocol group manages the radio links and performs functions like establishing, managing and disconnecting the link. The second protocol stack layer is called the Mid-level protocol group. It facilitates new and existing applications to operate over Bluetooth links. As it is an open standard the middleware protocol includes all the SIG and third party protocols and is aimed to support greater interoperability. It includes internet related protocols, wireless protocols and object exchange protocols and some other protocols to support other applications to run on Bluetooth links. The third level is an application protocol group which consists of all the applications irrespective of whether they are Bluetooth aware, e.g. telephony control protocol or unaware of Bluetooth, e.g. modem dialer application [HJ98] [Harte04b] [Miller02d].

4.2.9 Layers:

4.2.9.1 Radio Layer

The Bluetooth radio operates in the 2.4 GHz license free frequency band. The radio section covers most of the issues towards making a reliable and cost efficient Bluetooth transceiver. The issues include co-channel interference, in-band and out-of-band spurious emissions, intermodulation characteristics frequency accuracy etc.

4.2.9.2 Baseband Layer

The bandband layer is responsible for the link initiation and also assigns the master and slave roles to devices depending upon whether the device which initiates the connection becomes the master and other devices joining the network become the slave devices. As the same channel is used by different devices so the baseband layer also employs them with TDD (time division duplex) scheme. It also defines the methods to support synchronous[‡] and asynchronous[§] traffic and implement

[‡] Synchronous links are connection oriented links, also called (SCO). They are point-to-point full duplex links between the master and a slave. These links are established only by master and kept alive until the master does not release it.

[§] Asynchronous links are also called isochronous or asynchronous connectionless links (ACL). These are momentary links between the master and a slave for a duration of one frame. The master can freely decide which slave to address and in which order.

procedures such as error detection and correction, encryption and packet retransmission to make the system secure and robust

4.2.9.3 Link Manager Layer

Link manager is responsible for the management of the link set up between two Bluetooth devices. It manages the bandwidth allocation to get a desirable grade of output from logical link control and adaptation protocol (L2CAP) which supports multiple protocols and applications to share the air-interface. It manages device pairing (link management between the master and a slave) and carries out the device authentication process for a secure link. In case of failed authentication it can even disrupt the link.

4.2.9.4 Logical Link Control Layer

Logical link control layer is a link between the higher layers and the lower layers. It shields the working of the lower layers from the higher layers. It supports multiplexing allowing different applications to share the Bluetooth air-interface. It also enables the segmentation of large packets and reassembly of those packets at the receiving end. Apart from this it maintains a grade of service by working out an acceptable level of received signal for successful processing.

4.2.9.5 RFCOMM Layer

Bluetooth is basically a cable replacement protocol and RFCOMM is basically a cable replacement for the serial port which is used to transfer data across serial ports. It is designed to help different applications to use cable free serial port utilization without any significant modification to the application itself.

4.2.9.6 TCP/IP Layer

Bluetooth does not support the direct use of TCP/IP but it can be used to connect an IP network via an access point using a point-to-point link between the Bluetooth device and the access point, which in turn is connected to a network providing internet facility.

4.2.9.7 Audio Layer

Audio layer deals with audio traffic in Bluetooth protocol stack. Audio data is isochronous, means it has some time element associated with it, it is routed directly from the base band. Therefore the protocol stack is implemented such that the audio layer creates a direct contact between the higher layers and the lower layers, thus bypassing the intermediate layers like Link Manager and L2CAP layer. Special baseband packet structure called synchronous connection oriented (SCO) packets are defined for audio traffic.

4.2.9.8 IrDA (OBEX) and WAP Interoperability

Object exchange protocol is a higher layer protocol devised by the Infrared Data Association. Its function includes, setting up the logical link and determining the format of the data supposed to be exchanged between two devices. OBEX is a basic necessity of the various file transfer models. It supports applications like electronic business cards, e-mails and calendar entries. Similarly WAP interoperability layer also ensures interoperability between different devices trying to access data wirelessly.

4.2.9.9 Application Layer

Application layer contains the application software implemented by independent manufacturers or software vendors for a desirable purpose. The purpose to keep the application layer separate from other layers is evident from the fact that different applications can be implemented over the Bluetooth protocol stack just by making necessary modifications in the application layer software while keeping the rest of the stack intact.

4.2.10 Packet Structure

A Bluetooth packet consists of three fields: a 72-bit access code, 54-bit header and a variable payload of length between 2-342 bytes [HJ98].

4.2.10.1 Access Code

Access code is used for synchronization and identification. It is 72-bit in length. The channel access code consists of a preamble, a sync word, and a trailer. The preamble is a 4-bit zero-one pattern for DC compensation. It is followed by a sync word of 64 bit length. It is used to ensure larger Hamming distance[**] between two message strings and also improves the synchronization process. The trailer is a fixed 4-bit long and it is used for fine compensation.

4.2.10.2 Header

The header is a 54-bit field which is further divided into 6 subfields containing control information about the link. These 6 subfields are: 3-bit sub address (M_ADDR) which is a temporary address assigned by a master to distinguish an active slave from other devices in the piconet. Other devices do not use M_ADDR but their full 48-bit unique address. A 4-bit packet type (TYPE) specifies the type of packet used. There are 16 different packet types available. A 1-bit flow control bit (FLOW) controls the flow of packets in ACL links. FLOW=0 specifies no buffer space for incoming packet in recipient device, hence a request to stop the transmission. A 1-bit acknowledgement indication (ARQN) is used to indicate to the sender whether the reception of the packet was successful or not. ARQN=0 indicates success. A 1-bit sequence number (SEQN) is used to confirm the freshness of the packet and an 8-bit header error check

[**] The hamming distance can be interpreted as the number of bits which need to be changed (corrupted) to turn one string into another. Sometimes the number of characters can be used instead of the number of bits.

(HEC) to check the header integrity. The packet is discarded if it fails the header check. Fig 4.2.5 shows the Bluetooth packet format with its subfields.

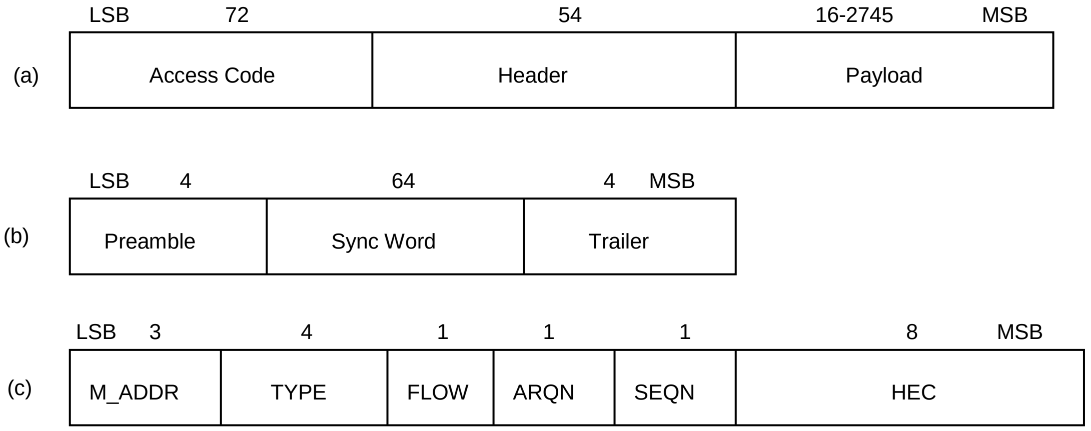


Fig 4.2.5 (a) packet format (b) channel access code (c) Header Format

4.2.11 Medium Access Control in Bluetooth

A Bluetooth network system consists of a piconet which contains a master and maximum of 7 active slaves. The master can be a base station or a fixed access point, while the slave can be any hand held devices such as laptop, cell phone or printer.

Time Division Duplex Multiplexing (TDD) is best suited for medium access control in Bluetooth networks [MMC99]. The master can talk to the slave in the first slot which is 625 μ s and the slave can reply in the very next slot. Thus these two slots in which the master and slave communicate with each other are called a “Bluetooth frame” [Rao01]. A master can only send packets to a slave in even slots and a slave can send packets to a master in an odd slot. The communication is therefore always in the form of a pair of slots. This scheduling structure can be a waste of slots sometimes when there is no data to be sent by either master or the slave. In Bluetooth networks only three types of packet lengths are supported: one slot, three slots and five slots [MMC99]. This is done to make sure that the packets starting from either odd or even slot can end up in odd or even slots respectively without disturbing the frame structure. This protocol guarantees an absence from collisions during data transfer. Fig 4.2.6 shows the communication model of a Bluetooth network. Fig shows alternate slot assignment to master and slave for communication.

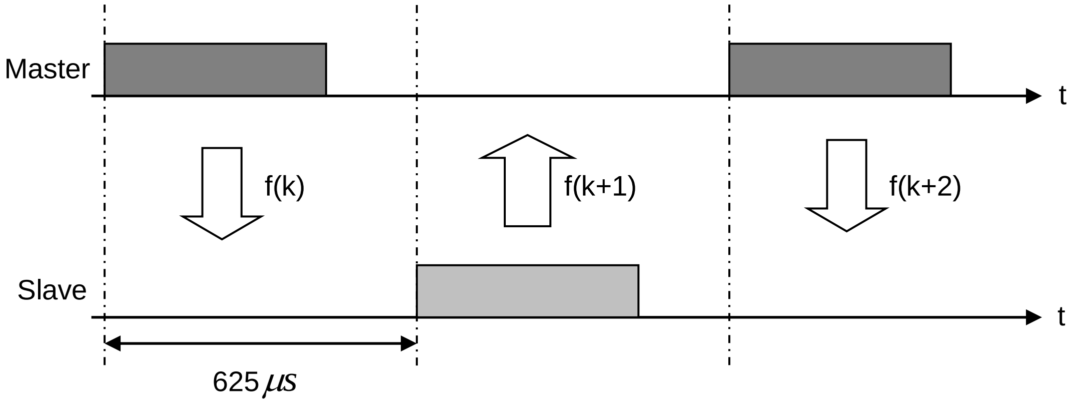


Fig 4.2.6: Single master-slave communication

Bluetooth networks support two types of links: Synchronous Connection Oriented (SCO) and Asynchronous Connectionless (ACL) for voice communication and data communication respectively [Rao01]. Asynchronous links are initiated by the master device and the slave responds to the master in the very next slot while in synchronous link the master and the slave talk to each other at regular intervals and fix the timing interval before the communication starts. Therefore it can be seen that the voice link or SCO links occupy fixed slots and ACL links adjust themselves between two SCO links. A voice link with two voice slots can be followed by four slots for data communication. Fig 4.2.7 shows the communication model between the master and multiple slaves. Communication includes both synchronous and asynchronous data.

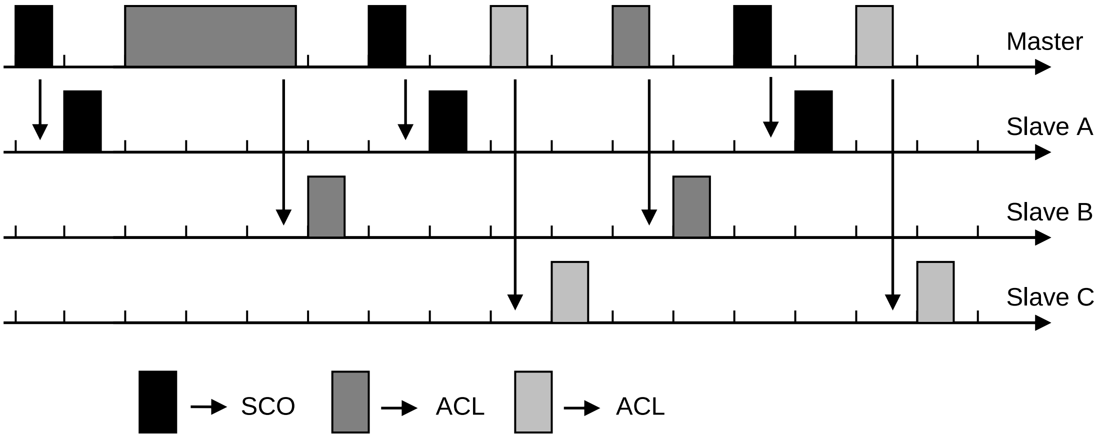


Fig 4.2.7 Master and multiple slave communication with variable size data link (ACL) and reserved voice link (SCO).

The process of connection between two devices starts with the master device, which sends identification (ID) packets on different hop frequencies and gives the indication to the devices about its desire to connect to a device. When the specific device with matching ID receives that packet it can respond to the master without any delay. When the master gets the response from the connecting device, it sends a frequency hopping synchronization packet to the device. Both the devices change their frequency hopping sequence to the master's frequency hopping sequence and if the master receives the response from the device, the connection between the master and the slave is considered successful.

4.3 IEEE 802.15.4

4.3.1 Introduction

Growth of technology and emerging wireless standards have made it possible to have high data rates in communication systems but still there are certain application sectors where the priority is not high data rate but low energy consumption. Most of the network systems today have high data rates, low latency and high Quality of Service (QoS) but these are not the requirements of every application. Some applications do require connectivity with low power consumption and low data rates. So it would not be appropriate to use the same systems for both applications. To meet this requirement the IEEE has come up with a standard which is particularly developed keeping in mind the requirements of what we call “Low Rate – Wireless Personal Area Networks” or LR-WPAN. It emphasises network systems with extremely low power consumption. These networks are applicable to the areas where mains power supplies cannot be used and frequent replacement of batteries is also difficult. The trade off for lower power consumption is usually lower data rate, higher latency and shorter range.

IEEE 802.15.4 is a standard which describes the technical aspect of the MAC and PHY layers of these LR-WPANs. The purpose of the document is to formulate the techniques for low complexity, low power consumption, low cost and low data rate wireless connectivity among inexpensive devices. It is designed to operate in the following license free bands [Gut03a] [IEEE03a]:

- 868.0 - 868.6 MHz: There is only one channel in this band which can give a data rate of up to 20 kb/s. This band is available in most European countries.

- 902.0 – 928.0 MHz: There are 10 channels in this band and each can support a data rate of up to 40 kb/s. The band is available in North America.
- 2.4 – 2.48 GHz: There are 16 channels in this band and can support a data rate of up to 250 kb/s. This band is available all over the world.

Because of the low data rates the first two bands are called “low bands” and the third is called “high band”.

All these bands are also called ISM bands (Industrial, Scientific and Medical). These are license free bands and anyone can use them. Devices operating in these frequency bands need to comply with the local regulations. In US, the Federal Communication Commission (FCC) is the regulatory body and in Europe it is European Telecommunications Standard Institute (ETSI) which is the regulatory body. Other countries around the world have their own regulations and authorities governing them.

4.3.2 IEEE and the Zigbee Alliance

The IEEE 802.15.4 standard was formulated by the IEEE 802.15 TG4 (“Task Group 4”) which had the main job of creating the solutions for the network in which the batteries can run from months to a few years and have low data rate requirement. However the main job of the group was restricted to the formulation of rules for the Physical layer and some portions of the data link layer (DLL) including the MAC layer [Rad04]. The higher layer specification is being done by another group of more than 100 industries under the name of the Zigbee Alliance. Zigbee Alliance is a non-profit group of companies working together to enable low cost, low power, reliable and secure wireless networks based on an open global standard. It has some major companies which have gained the status of the “promoter” companies. These promoter companies are Honeywell, Ember, Mitsubishi, Motorola, Ivensys, Philips and Samsung. It is believed that with applications like control and monitoring to home and industrial automation, the Zigbee market would grow to hundreds of millions of dollars by the end of 2006, with approximately 800 million Zigbee devices deployed around the world in different sectors [Rep04]. Zigbee Alliance is growing fast and in the last year it has almost doubled its members with the figure exceeding 100 now. Some of the recently joined members are Cisco Systems, Exegin Technologies, Orange Logic, Smarthome Inc, Yamatake Corporation etc. The tremendous interest shown by these world leading companies very well highlights the need of products based on an open global standard. Zigbee is built on the robust physical and MAC layer defined by IEEE 802.15 TG4 and ensures greater interoperability of products from different manufacturers. It addresses all the needs of the standard including low data rate, easy deployment, reliable link, low cost, long battery life etc. Unfortunately, unlike the IEEE standard, Zigbee standards are not available free of charge to non-alliance members. Fig 4.3.1 shows layer structure according to IEEE 802 Model and ISO-OSI Model. The IEEE 802.15.4 standard only deals with MAC and PHY layer of the IEEE 802 Model.

ISO-OSI Model	IEEE 802 Model	
7. Application	Higher Layers	
6. Presentation		
5. Session		
4. Transport		
3. Network		
2. Data Link	IEEE 802 LLC	LLC
	SSCS	
	IEEE 802.15.4 (MAC)	
1. Physical	IEEE 802.15.4 868/915 MHz (PHY)	IEEE 802.15.4 2.4 GHz (PHY)

Fig 4.3.1 Layer structure according to the IEEE 802 Model and ISO-OSI Model [Par04]

4.3.3 Technology Issues

The basics of the standard are influenced by the applications such as monitoring remote places, where low data rate is required, where battery powered sensors can be deployed with ease and where the network is cost efficient. The most important issue in the physical layer is the use of the standard in the unlicensed bands where there are so many other technologies using the same frequency band. This necessitates techniques to deal with interference from other sources. This is done with the help of a suitable modulation technique. The duty cycle of the transceiver is another concern which is accomplished by introducing sleep cycles by switching off the transceiver at regular intervals in order to increase battery life.

4.3.3.1 Physical Layer Issues:

As the standard operates in three different license free bands, 2.4GHz worldwide, 918 MHz for North America and 868 MHz in Europe, it can cover the whole world, although the specifications for different frequency bands are slightly different from each other. For example, the data rates supported by all these frequency are different and so is the modulation scheme and interference level. The 2.4 GHz band has lesser chances of interference than the other two bands because of the spread spectrum modulation technique used. When the spectrum is spread over a wider frequency range it also gives higher security because the chance for data interception decreases.

4.3.3.1.1 Modulation

Two different types of modulation techniques are used in low power radio networks which depend upon the frequency band. 868/915 MHz frequency bands use Binary Phase Shift Keying (BPSK) while the 2.4 GHz frequency band uses Offset Quadrature Phase Shift Keying (O-QPSK) [Gut03d]. But to protect data from interference and noise, more security measures need to be taken. For this, all frequency bands employ a spread spectrum technique on top their respective modulation schemes. This spread spectrum technique is called Direct Sequence Spread Spectrum (DSSS).

DSSS is the technique which is used to increase the bandwidth of a transmitted signal. Therefore the resultant signal after the DSSS is a signal with much greater bandwidth and lower spectral density [ECE4321]. Fig 4.3.2 shows how the spreading and the modulation operations in DSSS change the energy level of the signal and spread it over a wider frequency range.

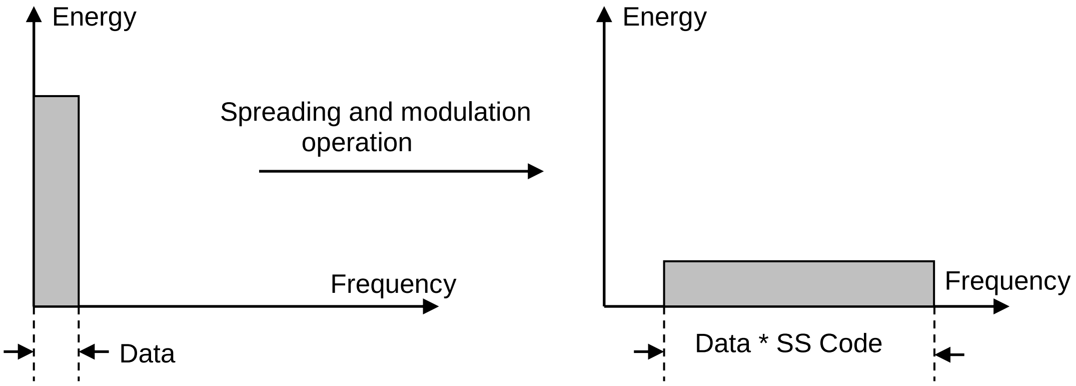


Fig 4.3.2 Energy level and frequency range of a data signal before and after the spreading and modulation operation

In DSSS a pseudo-random number (PRN) is multiplied directly by the data entering the carrier modulator. This increases the bit rate of the entering data because of the higher "chip" rate of the PRN sequence. Therefore the result of the modulator is a spread spectrum with wider bandwidth in ratio with the PRN bit rate. When the signal is received at the receiver it is demodulated with the help of the same PRN sequence. Hence the original signal can be recovered. The spread of signal in a wider bandwidth can be seen as a wastage of bandwidth but this loss is compensated with the use of same bandwidth by multiple users as the transmission is interference free because of unique PRN sequences used by different users allowing them security and privacy. The other advantage of DSSS is that when the signal is spread then the spectral power density of the signal decreases hence it seems like a noise on the receiver. This property is particularly very useful for military applications where information is highly secret and needs adequate measures for preventing it being discovered by possible enemies. DSSS gives better noise immunity since noise power is usually concentrated at certain frequencies whilst DSSS is spread. Fig 4.2.3 shows the block diagram of the DSSS transmitter and receiver. In this case, binary data is first modulated and then multiplied by a PRN sequence to make a spread spectrum. This signal is then transmitted over the channel and at the receiving end, the PRN sequence is extracted by DS De-spreader and data demodulated to get the original signal.

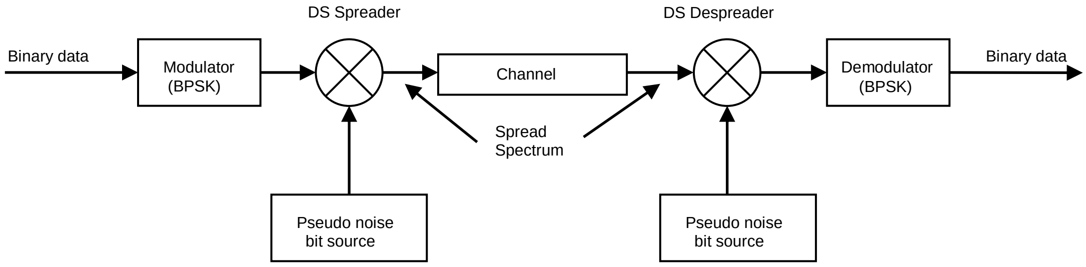


Fig 4.3.3Block diagram of DSSS transmitter and receiver

Given below is the table for different frequency bands and channels, data rate and modulation techniques with them.

Band	Number of Channels	Data Rate	Modulation
2.4 GHz	16	250kb/s	BPSK
915 MHz	10	40kb/s	BPSK
868 MHz	1	20kb/s	O-QPSK

Table 4.3.1 Different frequency bands and associated modulation and bit rate

4.3.3.1.2 Transmit Power

Because of the simultaneous use of the same frequency band by different users there has been tight regulations regarding the maximum transmit power associated with them to minimise interference. Increased transmit power can make the signal travel beyond the regulatory limits and hence violate law. Power limits are regulated by the local authorities and can vary from place to place. For example, in the United States, some services using DSSS in the 2.4 GHz band are allowed up to 1 Watt of transmitter power, however, in Europe the limit is 100 milliwatts in the same band [Gut03g].

4.3.3.2 Mac Layer Issues

The standard supports three different types of network topologies: star, mesh and cluster tree topologies. It is the function of the MAC to assign the master node in the network which controls the functioning of the network and synchronises the whole network. MAC also keeps track of the new joining nodes and gives addresses to them according to their state. MAC uses two types of addressing schemes; 64 bit and 16 bit [TG404]. The 64 bit address uniquely identifies a node in the network while it joins the network in the same way as Ethernet MAC addressing, but once the network is set up, the master node gives node another, short 16 bit address to the device to work within the network. In this way a network can support over 64000 nodes. MAC is also responsible for assigning time slots to different nodes in the channel for communications. It may also grant guaranteed time slots (GTS) if requested by the node and if it is feasible to accommodate those guaranteed time slots without hindering proper functioning of other nodes [Gor02] [IEEE03d]. Besides this MAC functioning includes providing a frame structure with unique id's (sequence numbers) and also includes an algorithm to check the correctness of the frame (Frame Check Sequence). Power consumption is an important feature which needs to be addressed skilfully. MAC does this by introducing sleep cycles between frames so as to keep the active node time as short as possible, whilst ensuring the synchronization between the sleeping and active nodes [CE04]. Nodes are categorised according to their states in the network, such as active state, sleeping state, idle state etc. MAC provides a timer which is used by the master node or the coordinator node for synchronization between it and other nodes in the network. The "coordinator" controls synchronization by sending periodic beacons to the member nodes which includes information about the frame structure, sleep cycles, PAN coordinater id etc. Last but not the least, the 802.15.4 MAC provides robustness to the network by making the nodes contend for the channel through a CSMA-CA algorithm. This diminishes the chances of collisions and saves energy used in retransmitting the data.

A detailed description about the IEEE 802.15.4 standard can be found in chapter 5.

Chapter 5

IEEE 802.15.4 STANDARD

5.1 Introduction

The IEEE 802.15.4 Standard was drafted for the connectivity of low power consuming devices which can operate in a small coverage area and require low data throughput with battery life lasting from months to a few years. The IEEE 802.15.4 Standard is the blueprint which describes the MAC and Physical layer specifications in Low Power Radio Networks. The standard defines the physical features such as transmitting power, frequency ranges, receiver sensitivities, modulation scheme, channel selection etc and MAC features such as superframe structure, MAC frame format, collision avoidance mechanism, acknowledgement frames and security, which can depend upon the type of application. The standard encompasses only layers up to and including portions of data link layer and the higher layers are left for the companies involved in the manufacturing of application specific products. The standard is a reliable solution for the wireless connectivity of the network devices involving low data rate and long battery life where other wireless connectivity technologies like IEEE 802.11 and Bluetooth may not fulfil the needs because of high complexity and low battery life. The existence of a standard makes it likely that products from different vendors will interoperate.

5.2 Overview

The IEEE 802.15.4 Standard applies to Low Rate- Wireless Personal Area Networks (LR-WPAN) and is intended to operate over a short distance known as Personal Operating Space (POS), which is of the order of 10-15 metres. A higher range can be achieved depending upon the application, at the expense of lower data rate. The aim basically is to implement a network system which is inexpensive, power efficient, not complex, tolerant towards interference from other networks and has the likelihood to be implemented over a wide range of devices. The data rates range from 250 kb/s to 20 kb/s depending upon the application and the frequency band used. For example the 2.4 GHz frequency band can support a maximum of 250 kb/s while the other two frequency bands 868MHz and 915 MHz can support maximum data rate of 20 kb/s and 40 kb/s respectively. PC enhanced toys and interactive games require maximum data rates of 112.5 kb/s so 2.4 GHz frequency range can be used. Other applications such as home automation and consumer electronics require a low data rate of the range of 10 kb/s, so 868 MHz and 915 MHz frequency ranges can be used.

5.3 General Description

Following are some of the characteristics of the LR-WPAN [IEEE03b];

- Supports star and peer-to-peer topologies
- Maximum data rate of 250 kb/s scalable down to 20 kb/s
- 16 bit short or 64 bit extended addressing scheme
- high reliability ensured by fully acknowledged data transfer
- collision avoidance scheme CSMA-CA used
- low power consumption
- link quality indication (LQI)
- energy detection schemes (ED)
- guaranteed time slots for urgent data (GTS)
- 27 channels; 16 in 2.4 GHz frequency band, 10 in 915 MHz frequency band and 1 in 868 MHz frequency band.

Low Rate WPANs support two different types of devices which are classified on the basis of their functionality in the network. These two type of devices are; full-function device and reduced-function device.

Full Function Device (FFD): A device which can operate as a coordinator, a PAN coordinator, or a device. It can talk to all other FFDs and RFDs.

Reduced Function Device (RFD): A device which can act as a device only and can talk only to other FFDs.

5.4 Network Topologies

LR-WPANs support both star and peer-to-peer network topologies [IEEE03f]. It depends upon the application scenario which topology is most appropriate. Both the topologies use a 64 bit extended addressing scheme for easy communication between the PAN coordinator and the device, however once the network is established, the PAN coordinator can assign a 16 bit short address to neighbouring devices. Usually, it is expected that the PAN coordinator is mains powered and devices are battery powered.

5.4.1 Star Topology

In the star topology the PAN coordinator is the main controller of all devices in the network. All communication is between the PAN coordinator and the device and there is no mutual communication between the devices. Once an FFD is activated, it starts looking for neighbouring devices and may become a PAN coordinator. This is accomplished by assigning itself a PAN identifier which is unique in the network and not being used by any other device in the POS. Once the id is assigned, it starts allowing other devices to join the network by allotting them addresses. A PAN coordinator is an FFD, and it can allow both FFDs and RFDs to join the network. This simple network structure is usually used in applications like home automation, computer peripherals, toys and games and personal health. Fig 5.1 shows star topology consisting of both full function and reduced function devices.

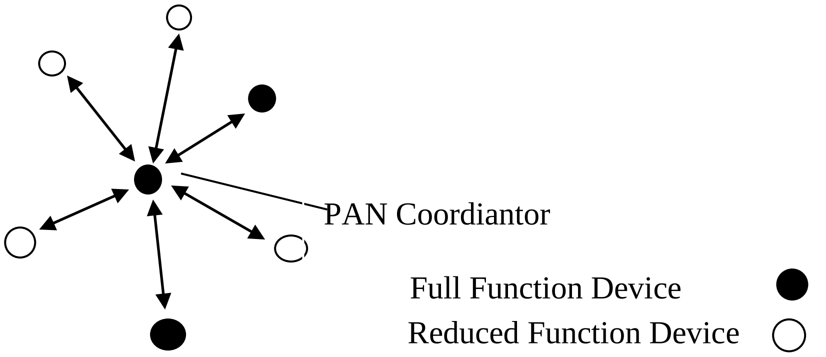


Fig 5.1 Star network topology

5.4.2 Peer-to-peer Network Topology

Unlike the star network, in peer-to-peer networks all the devices can directly communicate with each other within the POS. Any device which initiates the network or acquires the channel first can become the PAN coordinator with a cluster identifier (CID) of zero. The first job after becoming the PAN coordinator is to assign itself a PAN identifier. Once it has assigned itself an unused PAN identifier it can start establishing its own cluster. To make a cluster, the PAN coordinator starts broadcasting periodic beacons to the neighbouring devices which can join the cluster by requesting the coordinator. If the PAN coordinator accepts the request then the device can join the cluster as the child node in the neighbour list of the cluster head. The child device will also add the cluster head as its parent node in the neighbour list and start sending beacon frames to its neighbours. Other candidate devices can join the cluster at these nodes also instead of directly joining with the cluster head. As the distance of the PAN coordinator increases from the candidate devices the chances of the candidate device joining the same cluster decreases so the candidate device can request the PAN coordinator, via an intermediate node, to make another cluster and become the cluster head of that cluster. Upon receiving permission to create a new cluster from the PAN coordinator, the candidate device declares itself as a cluster head with cluster identifier (CID) value of 1. Only FFDs can become cluster heads while the RFDs join the network as leaf nodes at the end of the branch of at most one FFD. In this way a large cluster tree is formed with a single PAN coordinator coordinating and synchronizing the proper functioning of the devices associated in the network.

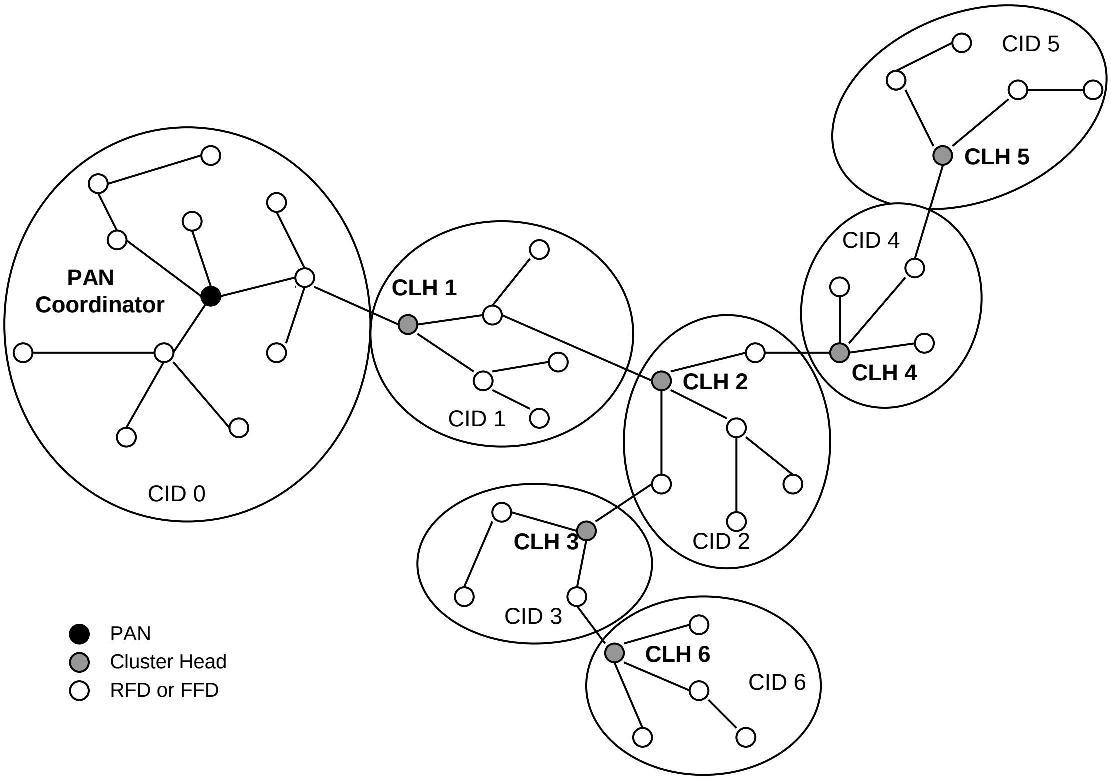


Fig 5.2 Cluster tree network topology

5.5 Functional Overview

5.5.1 Superframe structure

The superframe structure is the structure which describes the working of the PAN coordinator and its communications with other devices. It shows the allotment of time by the PAN coordinator to the child device to contend for the channel acquisition. The superframe structure is defined by the coordinator however its use is optional. Every superframe is bounded by network beacons and the superframe itself is divided into 16 equal slots. The network beacon resides in the first slot of the superframe. Devices trying to access the channel use a slotted or unslotted CSMA-CA mechanism to compete with other network devices to access the channel during the interval between two beacons, which is also called the Contention Access Period (CAP). Fig 5.3 shows a superframe structure with contention access period and inactive period between two beacon frames. Every superframe starts with a beacon frame.

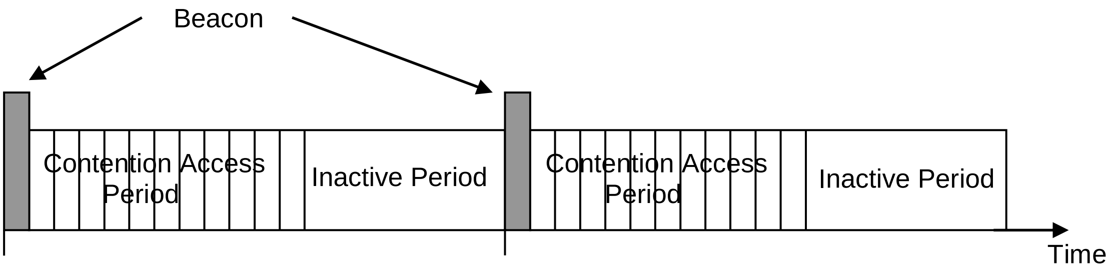


Fig 5.3 Superframe structure without GTSs

Due to the variable nature of the application area, some applications require low latency where it becomes impractical to hold the data for a long time in a queue while the device contends for the channel. The standard has introduced some portions of the superframe as contention free dedicated slots which are also called Guaranteed Time Slots, to ensure quick transfer of data. The GTSs appear at the end of the contention access period and there can be at most 7 GTSs allotted during a superframe. Each GTS can occupy one or more slots however a sufficient portion of the contention access period need to be ensured for the undisrupted functioning of other network devices. As there is no contention during this period it is also called contention free period (CFP). All the contention ends before this portion. All the devices enjoying the facility of GTS shall also ensure that their transactions finish before the start of the next GTS or otherwise before the end of contention free period. Fig 5.4 shows a superframe structure with contention access period along with contention free period.

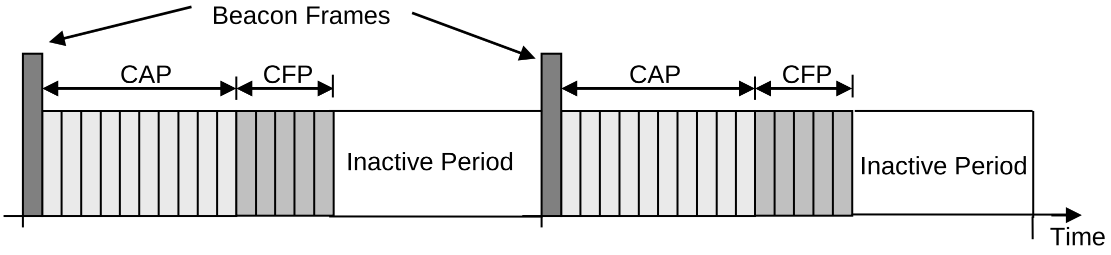


Fig 5.4 Superframe structure with GTSs

5.5.2 Data Transfer Models

There are three possible type of data transfer which can take place in the network. The first type of data transfer is when the coordinator sends the data towards the device, the second type is when the device sends the data towards the coordinator and the third type is data transfer between two peer devices. A peer-to-peer network topology supports all three types of transfers while the star topology involves only those between the devices and the coordinator. The models are further divided for all the three types depending upon whether the system is beacon-enabled or not. Shown below are communication models for the three types of data transfers both for beacon-enabled and beaconless transmissions.

5.5.2.1 Data transfer from the coordinator

In this case the data travels from the coordinator to the device. Whenever the coordinator wishes to send data to a device in a beacon-enabled network it indicates in the beacon frame that a message is pending for that particular device. The device periodically listens to the beacon frame and if a message is pending for it, it sends a MAC command requesting the data, using the slotted CSMA-CA mechanism described below. The coordinator sends an optional acknowledgement of the reception of the data request and then sends the pending data using the slotted CSMA-CA mechanism. The device sends an optional acknowledgement of the reception of data. The transaction is then complete. Fig 5.5 shows the message chart for communication from a coordinator in a beacon-enabled network.

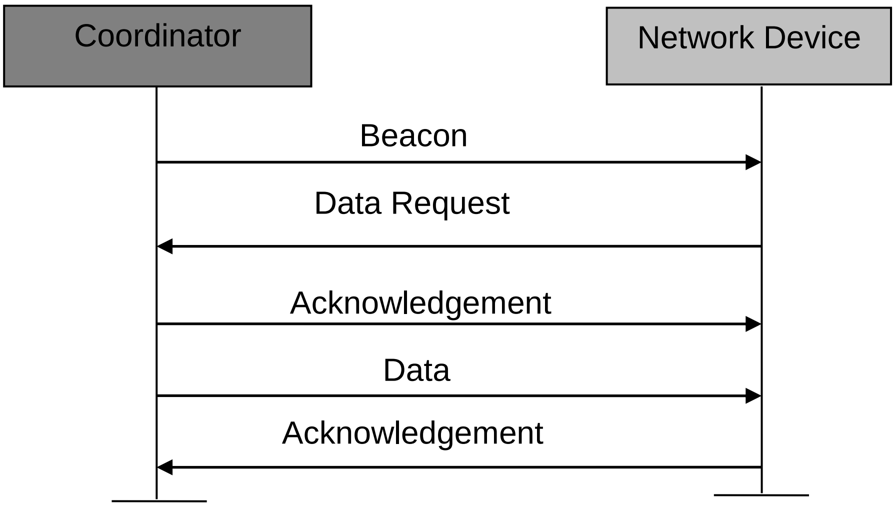


Fig 5.5 Communication from a Coordinator in beacon-enabled network

In a beaconless network the coordinator has no way to tell the device about the pending data so it stores the data and waits for the device to make contact and request the data. When the device contacts the coordinator requesting pending data using unslotted CSMA-CA then the coordinator sends an optional acknowledgement frame for the reception of the request and then sends the pending data using unslotted CSMA-CA. If there is no pending data then the coordinator sends a data frame with zero pay-load to the device. The device acknowledges the reception of the data frame in either case. The transaction is then complete. Fig 5.6 shows the message chart for communication from a coordinator in a beaconless network.

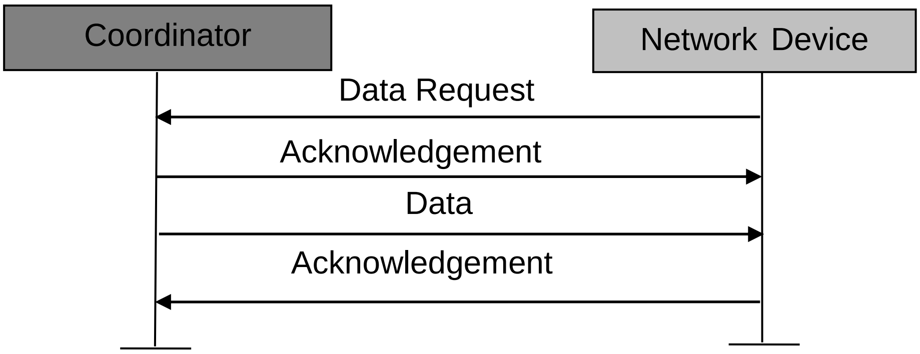


Fig 5.6 Communication from a coordinator in a beaconless network

5.5.2.2 Data transfer to the coordinator

In this case the data transfer is from the device to the coordinator.

In a beacon-enabled network, whenever the network device wishes to send some data towards the coordinator, it waits for the network beacon and then synchronizes itself with the superframe. It then sends the data to the coordinator during the CAP by using slotted CSMA-CA. Upon reception, the coordinator sends an optional acknowledgement to the device. The transaction is complete. Fig 5.7 shows the message chart for communication to a coordinator in a beacon-enabled network.

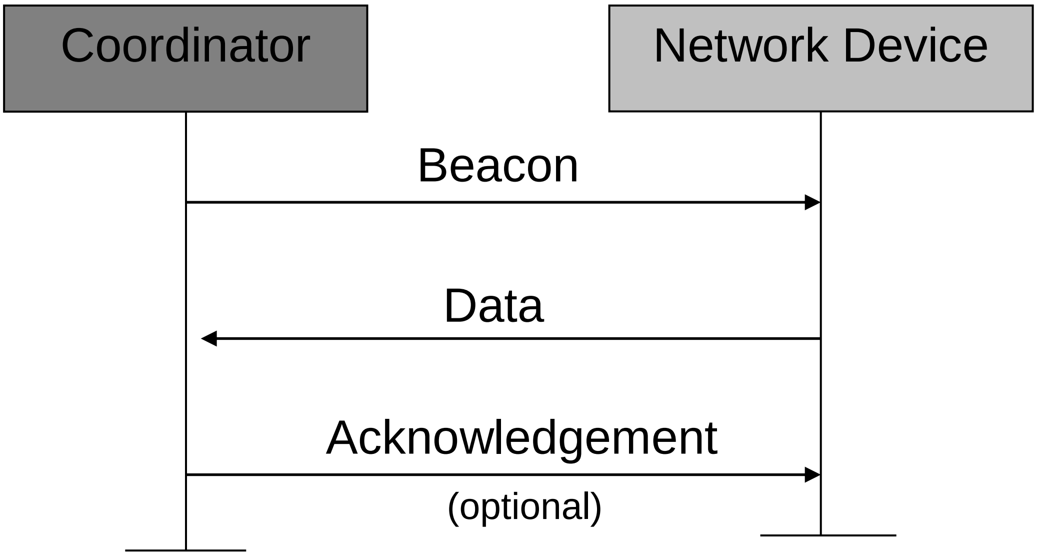


Fig 5.7 Communication to a coordinator in a beacon-enabled network

In a beaconless network, the device sends the data towards the coordinator without any delay using unslotted CSMA-CA and the coordinator sends an optional acknowledgement of the successful reception of the data. The transaction is complete. Fig 5.8 shows the message chart for communication to a coordinator in a beaconless network

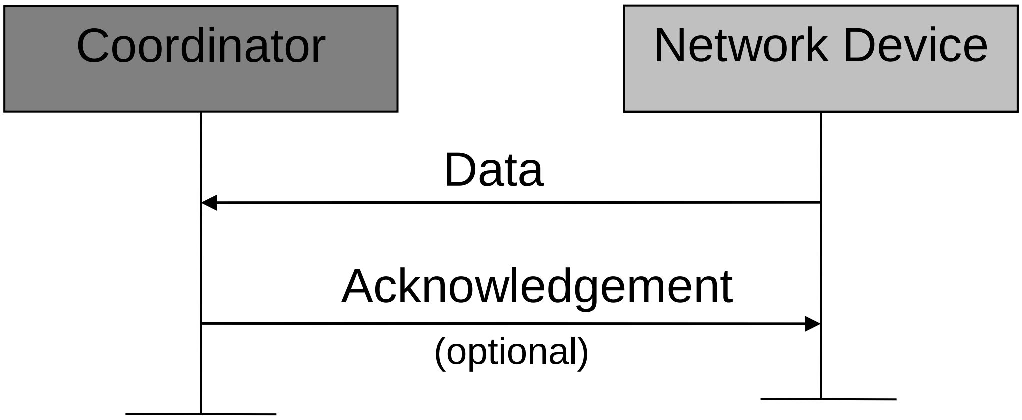


Fig 5.8 Communication to a coordinator in a beaconless network

5.6 Frame structures

The frame structure is very important because it explains all the aspects of the communication between two devices. It needs to be very clear, not complex and robust. There are four types of frames in the standard [IEEE03g]. These are:

- Beacon frames
- MAC command frames
- Data frames
- Acknowledgement frames.

5.6.1 Beacon frame format

Beacon frames are sent by the coordinator in a beacon-enabled network to other network devices. Basically a beacon frame originates from the MAC sublayer and as it passes through the PHY, a physical header is prefixed to the beacon frame. When it originates from the MAC it is called MAC service data unit (MSDU) and it contains the superframe specification, pending address specification, address list and beacon payload along with MAC header (MHR) and a MAC footer (MFR). A MAC header contains frame control, sequence number and addressing field information while the MAC footer contains a 16 bit frame check sequence (FCS). MHR, MSDU and MFR together form MPDU and are forwarded to the PHY as PSDU. There it is prefixed with a synchronization header (SHR) which is used for receiver synchronization and a PHY header which contains the information about the length of PSDU in octets. These together form the PHY beacon packet or MPDU.

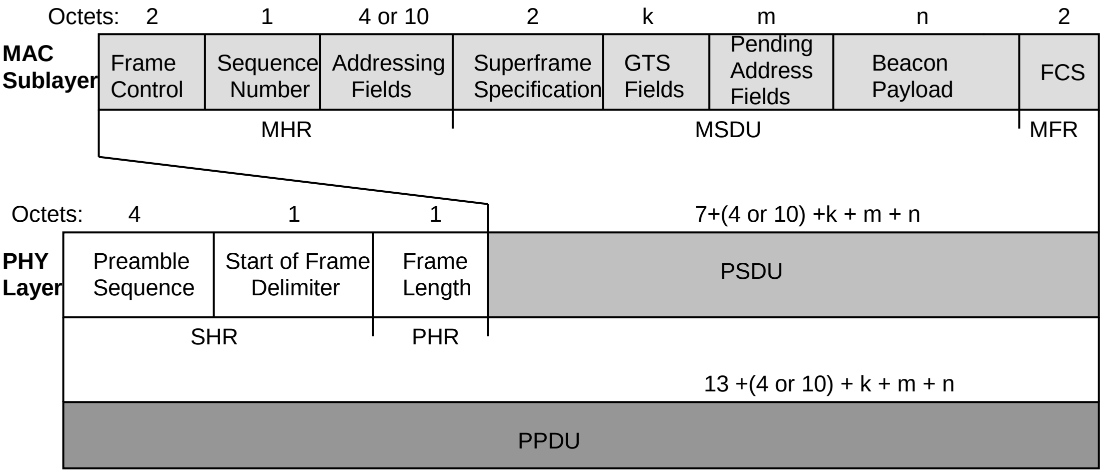


Fig 5.9 Beacon frame format

5.6.2 Data frame format

The data frame originates from the upper layers and as it passes through the MAC it is referred as MSDU and prefixed with the MAC header (MHR) containing frame control, sequence number and addressing field information and a MAC footer containing 16 bit FCS. Together they form the MPDU. This MPDU is then passed to the PHY where a PHY header is prefixed to this MPDU. This PHY header contains a SHR containing preamble sequence and start of frame delimiter (SFD) and PHR containing the length of PSDU in octets. The SHR, PSDU and PHR together form the PHY data packet.

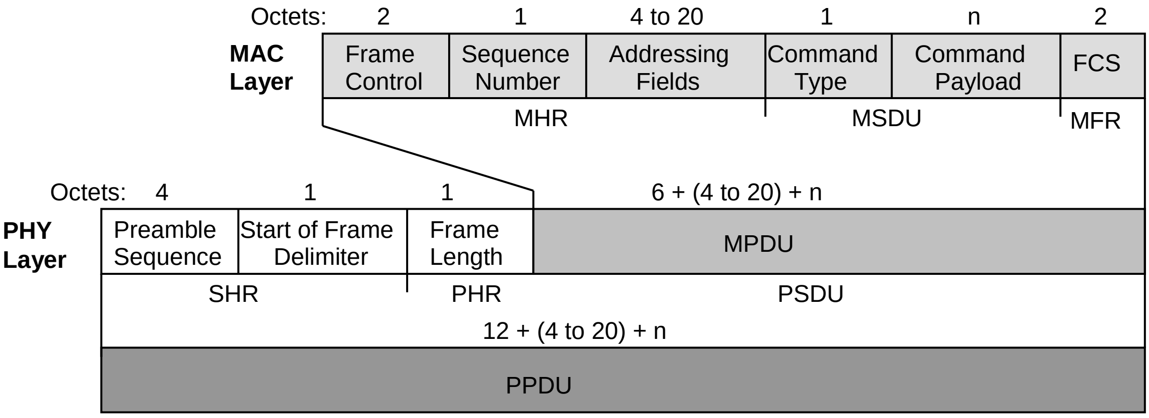


Fig 5.10 Data frame format

5.6.3 Acknowledgement frame format

An acknowledgement frame originates from the MAC sublayer. It is made up of MHR and MFR. MHR contains frame control and   sequence number while the MFR contains 16 bit FCS. MHR and MFR together make MAC acknowledgement frame (MPDU). When this MPDU is passed to the PHY, it is called PSDU and a SHR and PHR are prefixed with it. SHR contains preamble sequence and SFD and PHR contains the frame length which indicates the length of PSDU in octets. SHR, PHR and PSDU together form the PHY acknowledgement frame (PPDU).

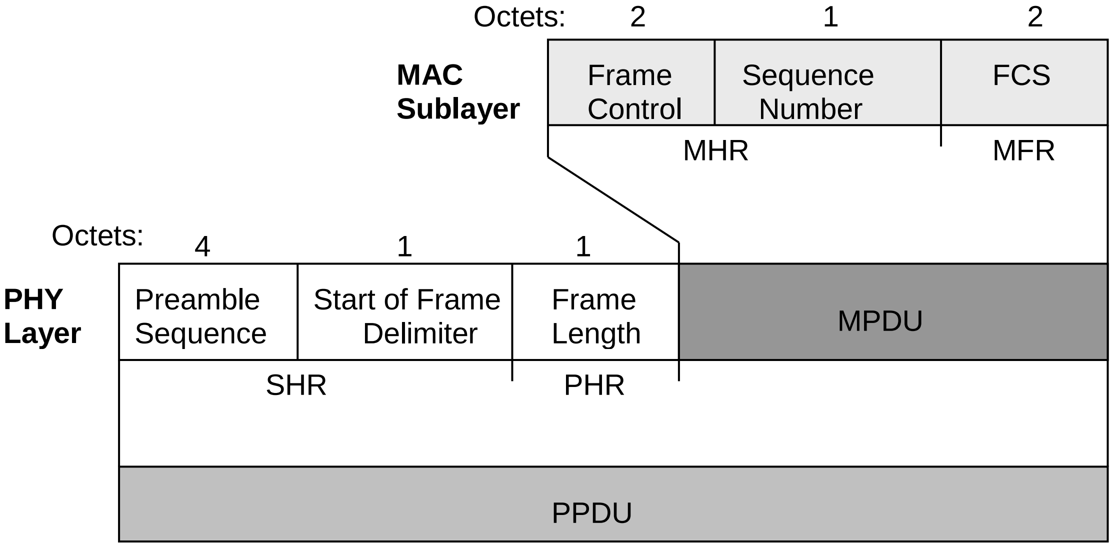


Fig 5.11 Acknowledgement frame format

5.6.4 MAC command frame

MAC command frame originates from the MAC sublayer. It contains MAC header containing frame control, sequence number and addressing field information, MSDU containing command type field and command payload and MFR containing16 bit FCS. Together MHR, MSDU and MFR form MPDU. Then it is passed to the PHY as PHY command payload (PSDU) and prefixed with SHR contains preamble sequence and SFD and PHR contains the frame length which indicates the length of PSDU in octets. SHR, PHR and PSDU together form the PHY command packet.

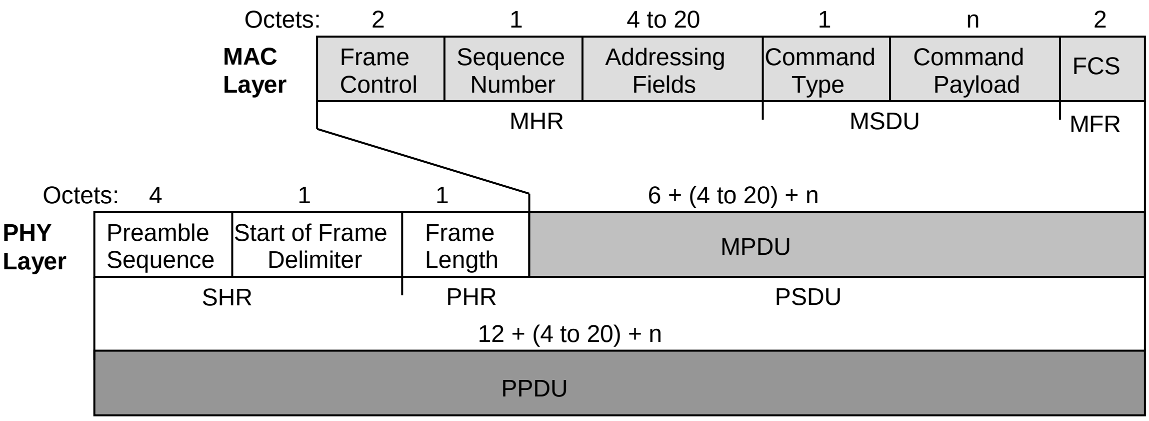


Fig 5.12 MAC command frame format

5.7 Robustness

Another important aspect of the IEEE 802.15.4 standard is its robustness in data transmission. The standard defines mechanisms such as CSMA-CA, frame acknowledgement and data verification scheme to ensure the reliable transfer of the information between the devices [IEEE03c].

5.7.1 CSMA-CA mechanism

The standard uses a CSMA-CA technique to access the channel. Different devices trying to communicate with the coordinator use this mechanism to access the channel. As the standard defines both beacon-enabled and beaconless networks, there are two types of CSMA-CA techniques used. These are called slotted CSMA-CA and unslotted CSMA-CA mechanism [IEEE03e].

Beaconless networks use an unslotted CSMA-CA technique to access the channel as shown in fig 5.13. Every time a device wishes to transmit data or MAC command, it has to wait for a random period of time and perform the CCA. If the channel is free, after a random backoff, the device can start transmitting; otherwise it undergoes another random backoff in the range 0-5 before trying to access the channel and increments the values of *NB* and *BE* by 1. It goes on until it acquires the channel.

However, there is a limit on the maximum number of backoffs defined by *macMaxCSMABackoffs* normally 5 after which the device gives up trying. Acknowledgement frames are transmitted without CSMA-CA. Fig 5.13 shows the flow chart for CSMA-CA mechanism as mentioned in IEEE 802.15.4 standard. The diagram depicts both slotted and unslotted versions of CSMA-CA. Slotted version of CSMA-CA is used in case of beacon-enabled networks while the unslotted version is used in case of beaconless networks.

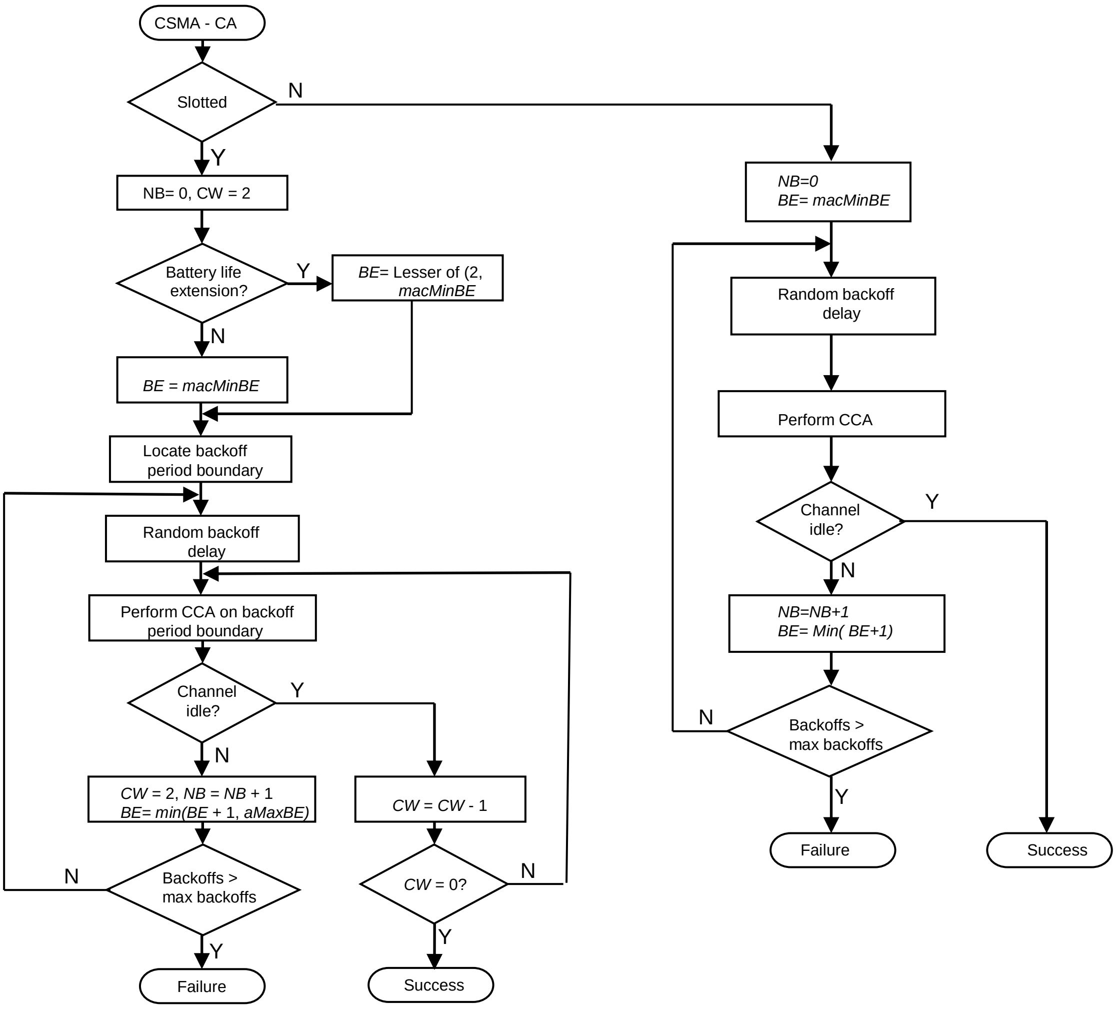


Fig 5.13 Flow chart showing CSMA/CA algorithms in beaconless and beacon-enabled networks

Beacon-enabled networks use slotted CSMA-CA. In this case the backoff slots are aligned with the start of the beacon transmission and every device has to locate the boundary of the next backoff slot before undergoing a random backoff. Each device maintains three variables during each attempt for transmission. These are *NB*, *CW* and *BE*. *NB* denotes the number of times the backoff algorithm was used while attempting transmission. This value should be initialized to 0 before each new transmission attempt and can go up to *macMaxCSMAbackoffs*. *CW* denotes the contention window length which should be free of any channel activity before the transmission can commence. This value is initialized to 2 before each transmission and reset to 2 every time the channel is sensed busy. *BE* denotes the backoff exponent from which the number of backoff periods can be calculated for a device to wait before attempting to transmit a data frame. In unslotted CSMA-CA with *macBattLifeExt* set to FALSE, *BE* shall be initialized to the value of *macMinBE* . However in case of slotted CSMA-CA with *macBattLifeExt* set to TRUE, the value lesser of 2 or value of *macMinBE* is used to initialize the value of *BE*.

In case of slotted CSMA-CA with battery life extension subfield set to 0, the MAC needs to check whether the entire transaction can be completed before the end of CAP after the backoff. If the number of backoff periods to wait is greater than the remaining number of backoff periods in the CAP, then it must pause the backoff count down and start it in the CAP of the next superframe. If the number of backoff periods is less than the number of backoff periods in the CAP then it should proceed only if the remaining transaction including two CCA's, frame transmission and the acknowledgement frame, if any can be transmitted in the same superframe. If it can proceed then it requests the PHY to perform the CCA in the same superframe. If not it waits for the next superframe and repeats the evaluation.

In the case of slotted CSMA-CA with battery life extension subfield set to 1, the MAC ensures that the whole transaction including backoff, CCA’s, frame transmission and the acknowledgement frame, if any, be completed before the end of the CAP. The backoffs can only take place in the first six backoff slots and the data frame starts in one of the first six backoff slots. If there is time then the MAC requests the PHY to perform Clear Channel Assessment (CCA) in the same superframe. If it is not possible then the MAC should wait for the next superframe and repeat the evaluation.

If the channel is found busy after that, then the device has to wait for another random number of backoff slots before trying to access the channel and increment the value of *NB* and *BE* by 1. On the other hand, if the channel is found idle, the device can start transmitting at the start of next backoff slot boundary.

5.7.2 Frame acknowledgement

Acknowledgement frames are used to confirm the successful reception of the data frame or the request both by the coordinator and the device. If the originator does not receive the acknowledgement from the receiver then the transmission is considered unsuccessful and the originator tries again to send the same data. On the other side, if the acknowledgement is received then the transmission is considered successful. Acknowledgement frames can be sent without using the CSMA-CA mechanism.

5.7.3 Data verification

Every frame sent either by the coordinator or the device has some information with which the errors occurred during the transmission can be detected. The standard uses

16 bit International Telecommunication Union – Telecommunication Standardization sector cyclic redundancy check to detect such errors.

5.7.4 Other important features

The standard provides a whole range of other operations to make the network system more secure, safe and long lasting.

As the standard is aiming to form a network which can last long without any battery replacement so some measures have been incorporated in the standard to realise such aims. The standard introduces the concept of duty- cycles to minimise the energy consumption. In duty–cycling the devices need to go to sleep and wake up periodically to determine whether there is pending data. However certain devices like the coordinator can be mains powered and can continuously listen to the RF.

Along with this the standard ensures security by implementing some security procedure like maintaining an access list (ACL) through which a device can select which device it can communicate with, and maintains a list of devices in the ACL with which it is expecting communication. Symmetric cryptography is used for protecting transmitted frames, message integrity code to protect data from being modified by a third party without cryptographic key and a sequential freshness scheme is implemented through which a device can know the freshness of its data. If the freshness value is more than the value of the previous data, it is processed, otherwise it is discarded.

5.8 PHY Layer Specification

The PHY (physical) section describes all the physical layer specifications related with the standard [IEEE03h] [Gut03f]. The main functions of the PHY layer are as listed below

- Activation and deactivation of the radio transceiver

- Energy Detection (ED) within the current channel
- Clear Channel Assessment (CCA) for CSMA-CA mechanism
- Link Quality Indication (LQI) for received packets
- Channel frequency selection
- Data transmission and reception.

All the constants and attributes are written in italics and constants have a general prefix ‘a’, e.g. *aMaxPHYPacket Size.* Attributes have been given a general prefix of ‘phy’, e.g *phyCurrentChannel*

5.8.1 General Requirements

The network can operate in three different frequency ranges depending upon the location of use and the requirement of the data throughput of the application. There are various other issues which need consideration for the successful working of the network such as power management, signal transmission, interference and sensitivity.

5.8.1.1 Channel Assignment:

There are 27 channels under the standard belonging to different frequency ranges. Depending upon the range of the frequency bands, channels are allotted while making sure that the two consecutive channels don’t overlap and they have proper spacing. Hence there is one channel in 868 MHz band, ten channels in 915 MHz band and there are 16 channels in 2450 MHz band. Shown below is the scheme showing the allotment of channel number in respective frequency bands.

F_C = 868.3 MHz, for k = 0

F_C = 906 + 2 (k – 1) MHz, for k = 1, 2, 3 …,10

F_C = 2405 + 5 (k -11) MHz, for k = 11, 12 ….,26

where, k stands for the channel number.

However different channels are used according to the region where they are supported and comply with the local regulations.

5.8.1.2 Transmit Power

As the same RF spectrum is used by different users so there is an urgent need to keep a check on the amount of power used during transmission so that they don’t interfere with other devices of other network using the same frequency. The local regulatory bodies have declared some regulatory limits to check such interference and all the users should make sure that their devices conform to those limits. The PHY parameter, *phyTransmitPower* indicates the nominal transmit power level.

5.8.1.3 Receiver sensitivity

Receiver sensitivity is measured by two factors named Packet Error Rate (PER) and Receiver sensitivity.

Packet error rate is defined as the average fraction of the transmitted packets that are not detected correctly. It is measured over random PSDU data.

Receiver sensitivity is defined as the threshold input signal power that yields a specified PER. It is measured over a PSDU length of 20 octets, with a PER less than 1%, power measured at antenna terminals over an interference free channel.

5.8.2 PHY service specifications

The PHY provides a link between the MAC sublayer and the physical radio channel. It includes a management entity called PLME. This entity is responsible for providing layer management interface through which all the management functions can be invoked. It can also maintain a database of all the managed objects in a database called PHY PAN Information Base (PIB). The physical layer provides two services named, the PHY data service and PHY management service through two Service

Access Points (SAP) i.e. PHY data SAP (PD-SAP) for PHY data service and PLME-SAP for PHY management service. Fig 5.14 shows the PHY reference model.

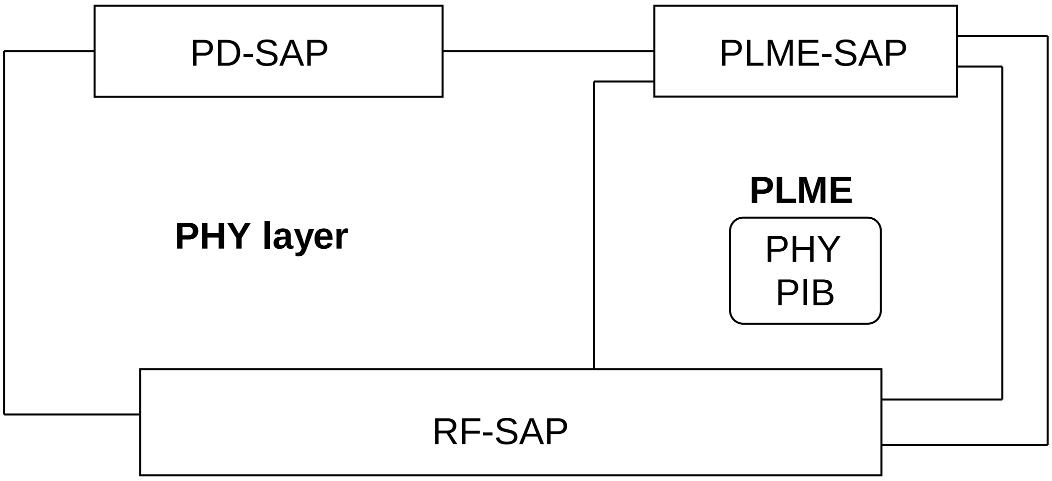


Fig 5.14 PHY reference model

5.8.2.1 PHY data service

The PD-SAP transfers the MDPU's between peer MAC sublayer entities and supports three primitives of PHY data service i.e. PD-DATA.request, PD-DATA.confirm and PD-DATA.indication. Table 5.1 shows primitives associated with PHY data service.

Primitive	Purpose
PD-DATA.request	Request for the transfer of MPDU from MAC sublayer to PHY
PD-DATA.confirm	Confirms the end of transmission of an MPDU
PD-DATA.indication	To indicate the transfer of a MPDU from PHY to MAC sublayer

Table 5.1 PHY data service primitives

5.8.2.2 PHY management service

The PLME-SAP allows the transport of management commands between the PLME and MLME. The PLME-SAP supports several primitives which are described below.

Primitiv	Purpos
PLME-CCA.request	Requests the PLME to perform a CCA
PLME-CCA.confirm	Reports the results of CCA
PLME-ED.request	Requests the PLME to perform an ED measurement
PLME-ED.confirm	Reports the results of ED measurement
PLME-GET.request	Requests the information about a particular PHY PIB attribute
PLME-GET.confirm	Reports the results of an information request
PLME-SET-TRX-STATE.request	Requests the PLME to change the internal operating state of the transceiver
PLME-SET-TRX-STATE.confirm	Reports the result of the request to change the internal state
PLME-SET.request	Attempts to set the PHY PIB attribute to a given value
PLME-SET.confirm	Reports the result of the attempt to set a PIB attribute

Table 5.2 PHY management service primitives

5.8.2.3 PPDU Format

A PPDU packet structure consists of three parts named a Synchronization Header (SHR) which is responsible for synchronization of the receiving device with the bit stream, a PHY Header (PHR), which contains information about the frame length and a variable length payload which carries the MAC sublayer frame. The PPDU frame structure is presented such that the field written on the leftmost side will be transmitted or received first. All octets are transmitted or received with least significant octet first and each octet is transmitted with least significant bit first. Table 5.3 shows general packet format.

Octets: 4	1	1		variable
Preamble	SFD	Frame Length (7 bits)	Reserved (7 bits)	PSDU
SHR		PHR		PHY payload

Table 5.3 PPDU frame format

5.8.2.3.1 Preamble field

Preamble is used by the transceiver to obtain chip and symbol synchronization which is lying in the incoming message. It is made up of a 32 bit binary number.

5.8.2.3.2 SFD field

SFD field indicates the end of synchronization field and the start of packet data. It is 8 bit in length.

5.8.2.3.3 Frame length field

As the name suggests it gives information about the number of octets in the PSDU or PHY payload. It has a value between 0 and *aMaxPHYPacketSize.*

5.8.2.3.4 PSDU field

PSDU field carries the data of the PHY packet. It has a variable length up to 127 octets. **.** All the packets with frame length 5 or greater than 7 contain a MAC sublayer frame also called MPDU.

5.8.3 2450 MHz band specifications

The 2450 MHz PHY employs DSSS technique**.** Each four bits of data selects one of 16 nearly orthogonal 32-bit pseudo-random noise (PN) sequences and the aggregate chip sequence is then modulated onto the carrier using offset quadrature phase-shift keying (O-QPSK). Fig 5.15 shows a block diagram of modulation and spreading process.

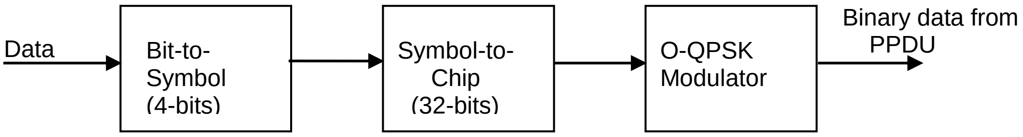


Fig 5.15 Modulation and spreading process

The 2450 MHz band has a symbol rate of 62.5 ksymbols/s ± 40 ppm with four bits in each symbol. Hence the final data rate comes out to be 250kb/s. the 32-chip pseudo-random sequence is divided into two orthogonal I and Q channels of the modulator. All the even chip sequences are placed on the I-channel and all the odd chip sequences are placed on the Q-channel and a one half chip delay is added to the Q-channel. This creates the offset for the O-QPSK. As 32 chips are transmitted in one symbol time of 16 μ s the overall chip rate comes out to be 2 Mc/s.

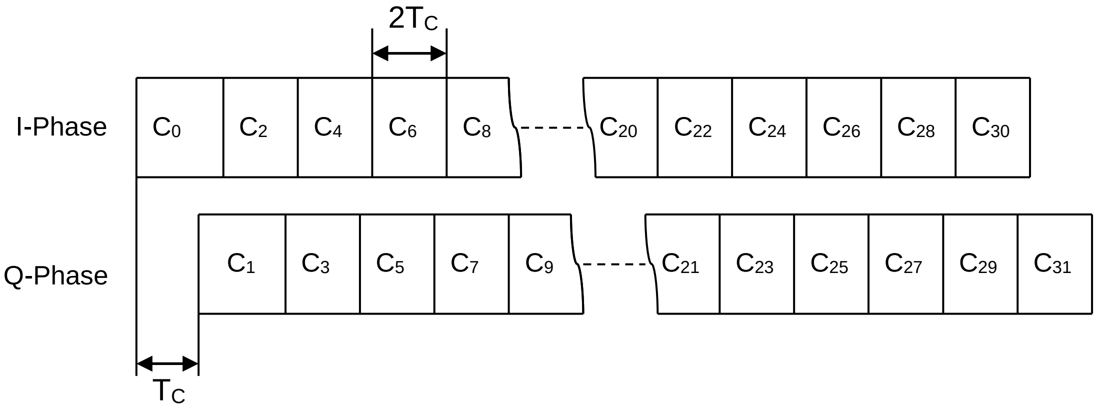


Fig 15.16 O-QPSK chip offsets

5.8.4 868/915 MHz Band specification

The 868/915 MHz PHY employs BPSK modulation and has a data rate of 20 kb/s in 868 MHz band and 40kb/s in 915 MHz band. Both bands employ differential encoding of the transmitted data bits. If the raw data encounters “0” bit, then the BPSK data bit is transmitted in the same phase as the previous one, but if the raw data encounters “1” bit then the BPSK data bit is transmitted in phase opposite to the previous BPSK bit. The 868/915 MHz PHY employs a direct sequence spread spectrum (DSSS) in which a single 15-chip pseudo-random sequence is transmitted in a symbol period representing “1” and the inverse of the sequence is used for the “0”. The chip rate is specified as 300 kchips/s in case of 868 MHz band, giving a data rate of 20 kb/s, while the chip rate of 915MHz is specified as 600 kchips/s which enables a data rate of 40kb/s. The modulation and spreading process is depicted in diagram below.

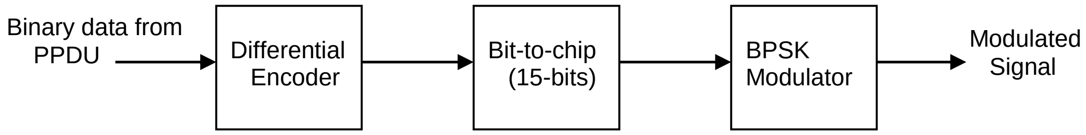


Fig 15.17 Modulation and spreading process

5.8.5 IEEE 802.15.4 Radio characteristics

The IEEE 802.15.4 standard is quite similar to other radio technologies but it has some of its own characteristics which makes it unique and helps it to exist distinctively among other radio systems in the vicinity. Some of its characteristics are discussed below.

5.8.5.1 Power output

The standard employs a wide output range. The device must be capable of transmitting -3dbm while the upper limit is designated by the local regulatory agencies to ensure an interference free transmission. For example, in US the government allows a transmitter power limit of 1 Watt in 2.4GHz band while the European government allows 100 milliwatts in the same frequency band.

5.8.5.2 Sensitivity

The standard has fixed that the receiver should be capable of decoding the incoming signal with an input power of -85 dbm or less in the 2.4 GHz band while in the lower

868 MHz and 915 MHz bands the receiver should be capable of decoding a signal with -92 dbm of input power.

5.8.5.3 Range

In ideal cases the devices operating in the lower frequency bands are capable covering a distance of 1 km and the higher frequency band has a range of about 220 meters. However in real-life models these figures change a lot because of atmospheric conditions, type of antenna used and path loss.

5.8.5.4 Receiver jamming resistance

IEEE 802.15.4 standard specifies an adjacent channel rejection for 2.4 GHz which is having 16 channels and for 915 MHz having 10 channels. The standard specifies an adjacent channels rejection of 0db which means that the channel must reject an interfering signal from the adjacent channel which is at same level i.e. 0db difference.

It also specifies an alternate channel rejection of 30db which means that the channel must reject the interfering signal from a channel which is two channels away from the channel of operation having a difference of 30 db from the simultaneous on channel signal.

5.8.5.5 Link Quality Indication (LQI)

The LQI is the measure of the strength or quality of the signal. This is measured by using an Energy Detection (ED) scan, a signal-to-noise ratio or a combination of two. The result of the scan is reported to the MAC through the PD-DATA.indication primitive with maximum and minimum values of 0x00 and 0xff referring to lowest and highest quality of the signal.

5.8.5.6 Clear Channel Assessment (CCA)

CCA stands for clear channel assessment. There are three modes of operation to perform a CCA.

- CCA Mode 1: In this case, the PHY will issue a report of medium being busy if it detects an energy level above the ED threshold.
- CCA Mode 2: In this case the CCA will report a medium busy only if it detects a signal with modulation and spreading characteristics of IEEE 802.15.4.
- CCA Mode 3: In this case the CCA will report a medium busy only if it detects a signal with modulation and spreading characteristics of IEEE 802.15.5 and energy above threshold.

5.9 MAC sublayer specification

The MAC sublayer handles all the data transfer between the MAC layer and the physical radio channel. Following are some of the functions of the MAC sublayer [IEEE03i] [Gut03e]:

- Generating network beacons in case of coordinator device.
- Association and disassociation of devices in the PAN.
- Synchronization
- Device security.
- Employing CSMA-CA for channel access.
- Handling the GTS mechanism.
- Providing a reliable link between peer MAC entities.

The MAC sublayer is responsible for providing two types of services to the higher layers. These are MAC data service and MAC management service. The MAC data service is accessed by MAC common part sublayer service access point (MCPS-SAP) while the MAC management service is accessed by MAC sublayer management entity service access point (MLME-SAP). Fig 5.18 shows the MAC sublayer reference model.

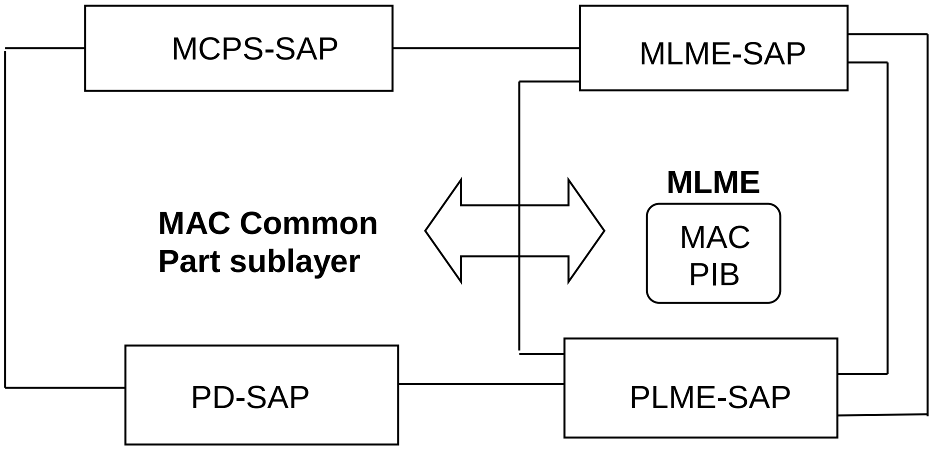


Fig 15.18 The MAC sublayer reference model

5.9.1 MAC data service

The MAC data service supports three types of primitives for the successful transfer of data between two devices. These three primitives are MCPS-DATA.request, MCPS-DATA.confrim and MCPS-DATA.indication. Table 5.4 shows primitives associated with MAC data service and fig 15.19 shows the message sequence chart describing MAC data services.

Primitive	Purpose
MCPS-DATA.request	Requests the transfer of a data SPDU from local entity to peer entity
MCPS-DATA.confirm	Reports the result of the request to transfer data
MCPS-DATA.indication	Indicates the transfer of data from MAC to a peer SSCS entity

Table 5.4 MAC data service primitives

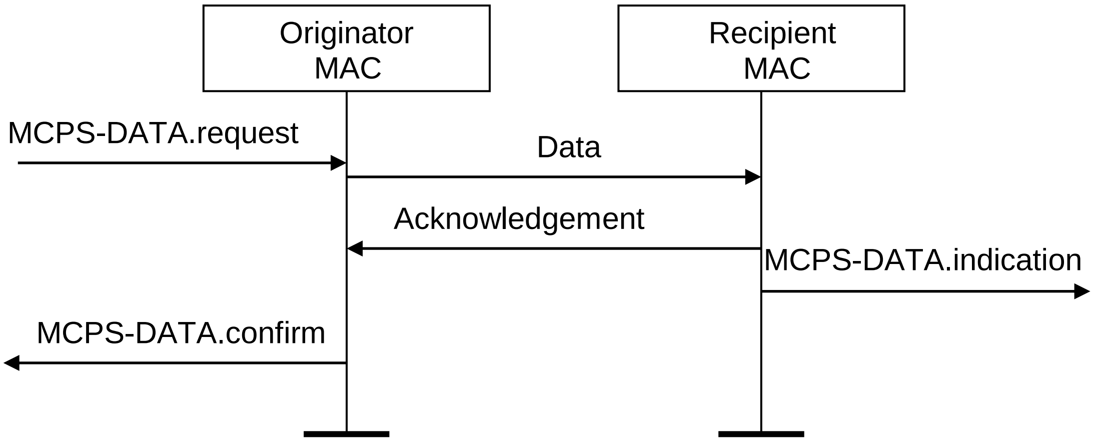


Fig 15.19 Message sequence chart describing MAC data service

5.9.2 MAC Management Services

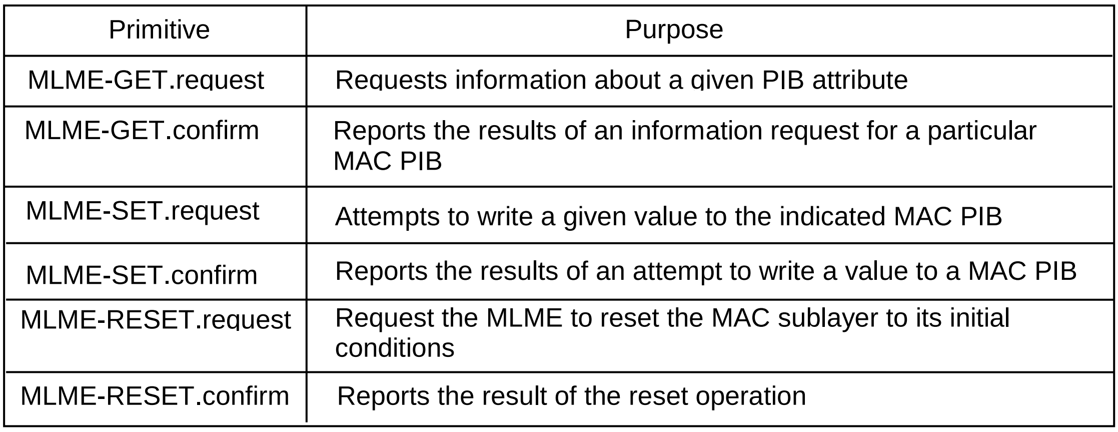
MAC management services provide support for commands related to communication settings, radio control and networking. Table 5.5 shows primitives associated with MAC management services.

Primitive	Category	Description	Request	Confirm	Response	Indication
GET	Communication Settings	MAC PAN Information base management	X	X		
SET			X	X		
RESET			X	X		
RX-ENABLE	Radio Control	Enables/Disables radio	X	X		
SCAN		Scan radio channels	X	X		
ASSOCIATE	Networking	Association control by coordinator	X	X	X	X
DISASSOCIATE			X	X		X
GTS		GTS Management	X	X		X
ORPHAN		Orphan device management			X	X
SYNC		Device synchronization	X			
SYNC-LOSS						X
START		Beacon Management	X	X		
BEACON-NOTIFY						X
POLL		Beaconless Synchronization	X	X		
COMM-STATUS		Communication Status				X

Table 5.5 MAC management service primitives

5.9.3 MAC PAN Information base management primitives

The PAN information base contains configurable attributes to control the MAC sublayer. These attributes can be read with the help of MLME-GET primitive and can be written with the help of MLME-SET primitive. In case the user wants to set the default values, he can do so with the help of MLME-RESET primitive. Table 5.6 shows primitives associated with MAC PAN Information base management service.

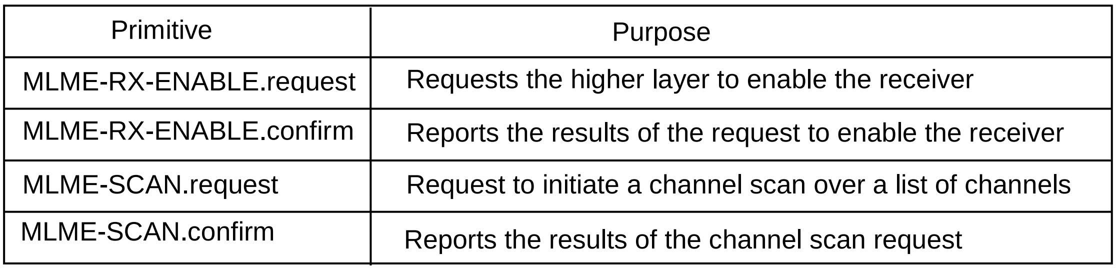


Primitive	Purpose
MLME-GET.request	Requests information about a given PIB attribute
MLME-GET.confirm	Reports the results of an information request for a particular MAC PIB
MLME-SET.request	Attempts to write a given value to the indicated MAC PIB
MLME-SET.confirm	Reports the results of an attempt to write a value to a MAC PIB
MLME-RESET.request	Request the MLME to reset the MAC sublayer to its initial conditions
MLME-RESET.confirm	Reports the result of the reset operation

Table 5.6 MAC PAN information base management primitives

5.9.4 MAC Radio Control primitives

MAC radio control primitives deal with enabling or disabling the radio system and performing the scan operation over a list of channels. The scanning operation can be performed to scan a channel to measure energy on the channel, search for a coordinator or search all the coordinators transmitting beacons. Table 5.7 shows primitives associated with MAC radio control service.



Primitive	Purpose
MLME-RX-ENABLE.request	Requests the higher layer to enable the receiver
MLME-RX-ENABLE.confirm	Reports the results of the request to enable the receiver
MLME-SCAN.request	Request to initiate a channel scan over a list of channels
MLME-SCAN.confirm	Reports the results of the channel scan request

Table 5.7 MAC radio control primitives

5.9.5 MAC network management primitives

MAC network management primitives deal with all the issues involved in creating and maintaining a network. It covers all the aspects such as association, disassociation, synchronization, orphan device detection, allotting GTS and beacon management, thus making the network efficient, reliable and flexible. Table 5.8 shows primitives associated with MAC network management service.

Primitive	Purpose
MLME-ASSOCIATE.request	Requests an association with the coordinator
MLME-ASSOCIATE.confirm	Reports the result of an association request
MLME-ASSOCIATE.indication	Indicate the reception of an association request
MLME-ASSOCIATE.response	Response to the MLME-ASSOCIATE.indication primitive
MLME-DISASSOCIATE.request	Used by the device to notify the coordinator about its intent to leave the PAN
MLME-DISASSOCIATE.confirm	Reports the results of the MLME-DISASSOCIATE.request command
MLME-DISASSOCIATE.indication	Indicate the reception of the disassociation notification command
MLME-GTS.request	Request to the coordinator to allocate a new GTS
MLME-GTS.confirm	Reports the results of a request to allocate a new GTS
MLME-GTS.indication	Indicates that the GTS has been allocated
MLME-ORPHAN.indication	Notifies the next higher layer of the presence of an orphan device
MLME-ORPHAN.response	Response to the orphan indication primitive by the coordinator
MLME-SYNC.request	Request to synchronize with the coordinator
MLME-SYNC-LOSS.indication	Indicates the loss of synchronization with the coordinator
MLME-START..request	Request for the device to start using new superframe configuration
MLME-START.confirm	Reports the results of the attempt to start using new configuration
MLME-BEACON-NOTIFY.indication	Used to send parameters contained within a beacon
MLME-POLL.request	Prompts the device to request data from the coordinator
MLME-POLL.confirm	Reports the results of the request to poll coordinator for data
MLME-COMM-STATUS.indication	Allows the MLME to indicate the communication status

Table 5.8 MAC network management primitives

5.9.5.1 Association Procedure

MLME-Association primitives define how a device can associate with a PAN. The association process starts with a MLME-ASSOCIATE.request primitive which allows a device to request association with coordinator. It is generated by the next higher layer of an unassociated device and issued to the MLME of the coordinator. On receipt of the request primitive, the MLME of the unassociated device first updates the *phyCurrentChannel* with the LogicalChannel by using PLME-SET.request primitive and then updates the *macPANId* with the value of CoordPANId parameter.

The MLME then generates an association request and sends it to the coordinator. On receipt of the request the coordinator sends the MLME-ASSOCIATE.indication primitive is to indicate the reception of the request to associate. It is generated by the MLME of the coordinator and issued to its next higher layer. When the next higher layer receives the indication primitive it issues a response primitive. It is generated by the next higher layer of a coordinator and issued to its MLME. Upon receipt of the response primitive the coordinator tries to add the information in its list of pending transaction. If there is no capacity to store the information, then an indication primitive with a status of TRANSACTION_OVERFLOW is issued. If the transaction is not handled within *macTransactionPersistenceTime*, then an indication primitive with a status of TRANSACTION_EXPIRED is issued. If the CSMA algorithm fails, then a status of CHANNEL_ACCESS-FAILURE is issued. In case of no acknowledgement, a status of NO_ACK is issued. If transaction is successful, then a status of SUCCESS is issued and if the parameter is not supported or out of range, then a status of INVALID-PARAMETER is issued. The MLME-ASSOCIATE.confirm primitive is used to inform the initiating device about the result of its request. It is generated by the MLME of the initiating device and issued to the next higher layer in response to the associate request primitive. Fig 5.20 shows a full message sequence chart for association process.

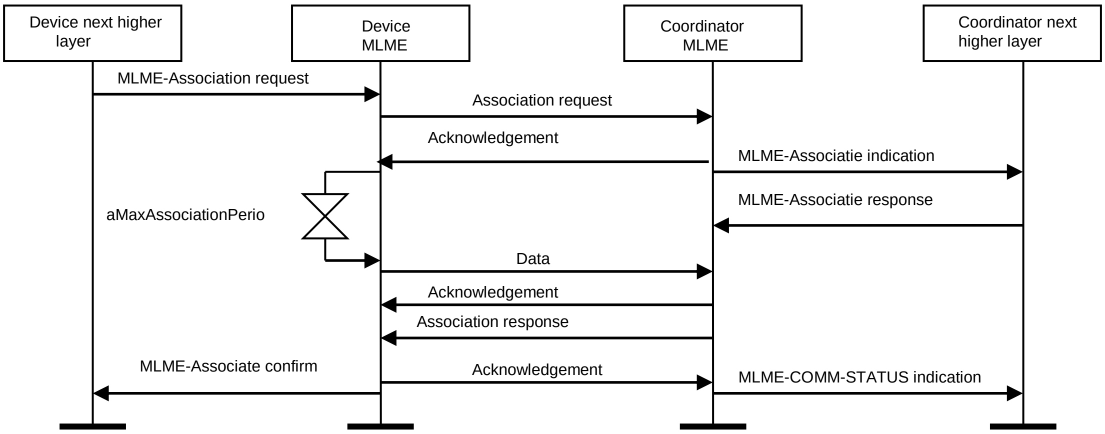


Fig 5.20 Message sequence chart for association

5.9.5.2 Orphan Notification Procedure

Orphan notification procedure is used by the coordinator to notify the presence of an orphan device and if possible, realign the device with the coordinator. Whenever the MLME of the coordinator comes across a orphan, it sends a MLME-ORPHAN.indication to its next higher layer to tell about the orphan device. The next higher layer then determines if the device was previously associated and issues a response primitive within *aResponseWaitTime*. It will send an AssociatedMember parameter of “TRUE” if the device was associated and “FALSE” otherwise. In response to the indication, the next higher layer sends a MLME-ORPHAN.response primitive to the MLME after it has made a decision whether the device was associated or not. If the AssociatedMember parameter is set to “TRUE” then the coordinator sends a coordinator realignment command with a short address to the orphan device. If it is set to “FALSE” the device is not associated. If it does not receive a response within *aResponseWaitTime* it will assume that it is not associated with any coordinator. Fig 5.21 shows a message sequence chart depicting message exchange during orphan notification procedure.

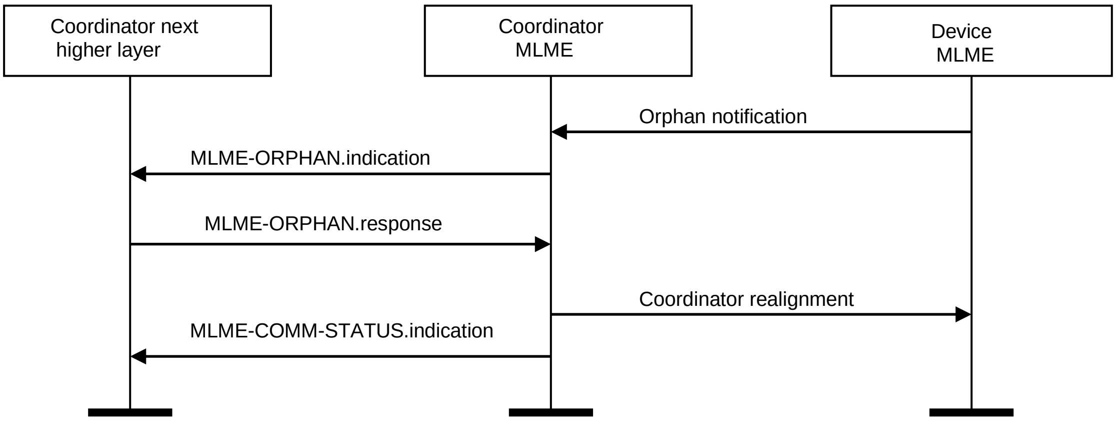


Fig 5.21 Message sequence chart for orphan notification

Chapter 6

CSMA-CA Algorithm

6.1 Introduction

A simulation study of the IEEE 802.15.4 CSMA-CA algorithm was carried out. First the operation of the algorithm is explained. In a medium where there are more than two users sharing the same medium, there is always a possibility of collision of data at the receiver. This collision not only garbles or corrupts the message but also add on to the wastage of time and energy involved in the transmission of that message. This overall brings down the efficiency and reliability of the network. There needs to be some mechanism that can help us avoid this wastage.

CSMA/CA which is short for Carrier Sense Multiple Access/Collision Avoidance is one such mechanism that can be used for this problem. Unlike CSMA/CD which is a collision detecting mechanism, retransmitting the data after the detection of collision, CSMA/CA mechanism works by listening to the medium for any transmission before transmitting its own data to avoid any collision. This not only minimizes the latency involved in transmitting the signal but also saves a lot of energy involved in the process. In radio channels where the data is sent through radio signals it becomes difficult to detect the collision using CSMA/CD because of the attenuation of radio signals as they move away from transmitters. These attenuated signals are difficult to pick by the transceiver which is already in the transmitting mode. This problem makes CSMA/CD ineligible for radio networks while CSMA/CA solves the problem by first sensing the medium and if busy, backing off the transmission for a random time period which is also called random exponential backoff.

In CSMA/CA whenever a client wants to send data or transmit data, it undergoes a random backoff and then senses the medium again before transmitting the data. If the medium is busy after that backoff, then the node freezes its timer until the medium is free again but if it is free, it starts transmitting after its timer value reaches zero. This brings down the chances of collision to a large extent. Every time the node encounters a collision it increases the range of the random delay exponentially until the transmission is successful and it is reset to the minimum value after that. Different standards fix different limits on the maximum and minimum value of the backoff exponentials or the contention window.

6.2 Operation of CSMA-CA

Initially all the stations are in idle mode. When a station has some data to send, it first undergoes a random backoff and starts sensing the medium. If the channel is sensed free then it starts decrementing its backoff timer until it reaches zero. It can only start transmitting after the timer reaches zero. This backoff timer is also called Network Allocation Vector (NAV). It is worth mentioning here that a station can start decrementing its NAV timer only if the channel is free but if the channel becomes busy, it freezes it until the channel is free again. If the NAV expires and the channel is still busy then the station undergoes another backoff with a random number selected from a wider range, increasing the length of the contention window. This procedure goes on until the data is successfully transmitted to the destination station. However there is a limit on the maximum number of retries that a node can make before finally giving up the transmission until the next time period which again depends upon the standard in which the CSMA/CA is used. If the transmission has been successful after attaining the channel then the state goes to the idle state again.

6.3 CSMA/CA in 802.15.4

There are two types of CSMA/CA algorithms in 802.15.4 which depend upon the type of application in which they are used [IEEE03e]. They are:

- Slotted CSMA/CA
- Unslotted CSMA/CA

6.3.1 Unslotted CSMA/CA

Unslotted CSMA/CA algorithm is used in a system in which no beacons are used. Unlike slotted CSMA/CA algorithm, the backoff time of different devices in unslotted CSMA/CA are not synchronised to each other. There are certain variables that the device need to keep track of, including *NB* and *BE. NB* is the number of times the CSMA/CA algorithm undergoes backoff while trying to transmit the data and *BE* is the value of backoff exponent. It is used to determine how many backoff periods it has to wait before trying to transmit again.

When the algorithm starts these values of *NB* and *BE* are set to their minimum values which is ‘0’ for *NB* and *macMinBE* (usually 0-3) for *BE.* Then the device waits for a random delay in the range 0 to (2^{BE}- 1) unit backoff periods and then performs a clear channel assessment (CCA). If the channel is found idle after that then the data can be transmitted otherwise it will again undergo a random backoff with incremented value of *NB* and *BE* by one. This increases the possible length of the backoff period which further brings down the possibility of collision. This process is repeated until the data is successfully transmitted or until the value of *NB* exceeds the maximum permitted value which is *macMaxCSMABackoff* (usually 0-5) after which the devices gives up trying and the transmission is assumed to have failed.

6.3.2 Slotted CSMA/CA

The slotted version of the CSMA/CA algorithm is used in networks in which beacons are enabled. However CSMA/CA algorithm is not used while transmitting beacon or acknowledgement frames. In the slotted version of the CSMA/CA algorithm, the MAC first initialize *NB, CW,* and *BE* and then it tries to locate the boundary of the next backoff period. The MAC delays for a random number of backoff periods in the range of 0-2^{BE}-1 and then the physical layer performs the CCA. This CCA starts at the next boundary of a backoff period. The MAC makes sure that after this random backoff, all the remaining CSMA/CA operations and data transmission is complete before the end of the CAP. If the number of backoff periods required is greater than the number of backoff periods remaining in the CAP, the MAC should stop decrementing its backoff timer and resume it at the start of the CAP of the next superframe. If the number of backoff periods is less than the number of backoff periods in the CAP, then the MAC should first evaluate whether the whole transaction including two CCA’s, frame transmission and acknowledgements if any, can be completed before the end of the CAP.

After performing the CCA operation on a backoff period boundary, if the channel is accessed to be busy then the variables *BE* and *NB* are incremented by one and the device again goes for random backoff delay. This procedure is repeated until the data is successfully transmitted or till the variable *BE* reaches its maximum permitted value which is *aMaxBE*. If the transmission is successful after that then the algorithm terminates with channel access failure status.

If the channel is sensed idle after CCA then the MAC waits until the contention window expires (*CW* backoff slots) before starting transmission. The MAC can not start transmission until the contention window is equal to zero, so every time when the MAC wants to transmit is should first decrement the value of CW and then check whether it is 0 or not. If it is not equal to zero then the MAC will undergo CCA mechanism again, but if the value is equal to 0, then the MAC can begin transmission at the boundary of the next backoff period.

6.4 Simulation Model

A simple simulation model of the CSMA-CA algorithm was constructed. Every node is assumed to attempt transmission at a random time in every transmission slot. A slot is the time period during which a node can transmit and represents the CAP part of a superframe. All packets are assumed equal in length, and take ‘T’ time to transmit. All nodes are assumed within range of each other. The unslotted version of CSMA/CA algorithm defined in IEEE 802.15.4 is used by each node.

For each slot period, a transmission start time is calculated for each node. This is a random starting time within the slot, plus the initial CSMA random backoff generated from the backoff algorithm. This time represents the start of Clear Channel Assessment (CCA) for the node wishing to transmit.

At the beginning of the transmission slot, the node with the earliest transmission start time is found. Since the channel is free at this time, that node will start transmission. Collisions occur if another node completes CCA before the first node begins transmission. After completing CCA which takes *CCAdetectionTime* seconds equal to

8 symbol times according to IEEE 802.15.4 standard, the first node will switch its transceiver to transmit mode. This takes 12 symbols times as defined in IEEE 802.15.4 and is called *aTurnaroundTime*. If another node's transmission start time is less than or equal to *CCAdetectionTime* before the expiry of *aTurnaroundTime*, then the CCA process will indicate a clear channel, and transmission will proceed, resulting in a collision. In other words, any other node whose transmission time is scheduled for *aTurnaroundTime* or less after the first node will result in a collision.

The transmission starting time is compared with the starting time for every other node to identify every node which would collide in this way with the next transmitted packet. Each node thus colliding results in the collision count and the transmitted count both being incremented. The new time at which the channel will be free (last colliding node completes transmission) is calculated. Collision events can thus involve 2 or more nodes, and the collision count indicates the number of packets lost due to collsions. The time for which the channel is busy during collision is also calculated. This collision time is lost transmission time and reduces the utilisation of the channel.

Any other node wishing to transmit at some later time, but during the time that transmission is taking place would have performed a clear channel assessment (CCA), found that the channel is busy and performed a backoff. After collisions have been dealt with, these nodes are identified and for each of them a new transmission start time is calculated using the CSMA backoff algorithm. If the new transmission start time calculated is before the end of the current transmission, another backoff is performed and a new transmission start time is calculated. This is repeated until either a calculated transmission time would occur after the end of the current transmission time but still within the current slot (superframe) time, or the maximum number of backoffs is reached, or the calculated transmission time would result in the packet not being able to be transmitted within the current slot.

The whole process is repeated until the end of the slot time and then repeated for any further slots required for each pass. At the end of the simulation pass, the random number generator seed is incremented so that the next simulation pass uses a different random number set.

When each node transmits, the time elapsed from the initially scheduled transmission time to the actual transmission start (latency) is recorded, in addition to the number of times the backoff algorithm was invoked before transmission. Also recorded is the number of transmission attempts requiring 1, 2, 3, 4, 5 or more than 5 backoffs.

6.5 Simulation results

The study of the CSMA-CA algorithm shows various factors which are responsible for successful transmission of the data packets. These factors also influence the latency involved and the collision rate. These factors are superframe size “S”, time taken to transmit packet “T”, the maximum number of backoffs the CSMA-CA algorithm will attempt before declaring a channel access failure “*macMaxCSMABackoffs*”, the minimum value of backoff exponent “*macMinBE*”, the maximum value of backoff exponent in the CSMA-CA algorithm “*aMaxBE*” and number of nodes “N” within range. The goal of the simulation is to find out the characteristics of the CSMA-CA algorithm as described by the IEEE 802.15.4 standard and to see how it works if the variables are altered. The primary goal is to see the number of collisions involved in transmission of packets with different lengths and with varying number of nodes. There is a matrix of scenarios where we can see the effect of all these factors on the number of collisions and latency. Some of these factors are discussed here with the results shown in graphical form. All the results are represented as averages of a number of passes.

To investigate the behaviour of the CSMA/CA algorithm, a C program was implemented with all the features supported by IEEE 802.15.4 including "*aMaxBE*", *NB*, packet length, number of nodes and superframe size. The program was run for different number of nodes with different values of "*MaxBE*", *NB* and packet length to see the number of collisions under different node densities. For each set of readings the program was made to run for 1000 times and the average was found to bring out the final results.

The first simulation result represents the variation of number of collisions verses the number of nodes operating in the network. The graph in figure 6.1 depicts three different packet length used for comparative study of collisions in three different cases. It can be seen through the graph that for obvious reasons when the number of nodes increases the number of collisions also increases due to more traffic or more nodes trying to transmit at the same time. So with smaller networks the packet length makes only a small difference in the number of collisions. However the number of collisions does follow an increasing trend as the packets become smaller. But as the networks become bigger the results diverge. Bigger networks experience more collisions in case of smaller packets but as the packets become larger the number of collisions starts decreasing. This can be because of the fact that bigger packets keep the superframe busier than smaller packets. Moreover the number of inter-packet intervals for a particular superframe is higher in the case of smaller packets than larger packets, which provides the opportunity for other nodes to contend for the medium resulting in more collisions. This can be seen in the curve depicting the collision curve for 133 bytes packet. The collisions have almost doubled in case of smaller packets of 64 bytes compared to 133 byte packet.

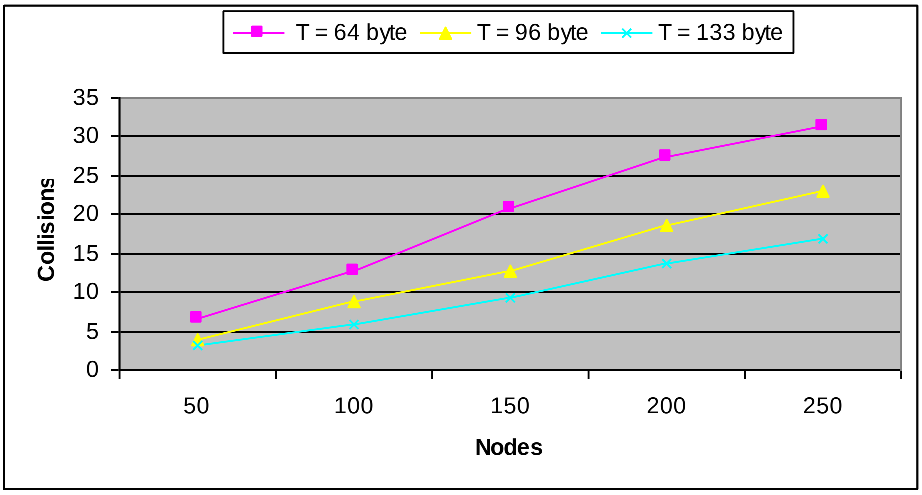


Fig 6.1 Collisions against number of nodes with different packet size

The time taken by the packet to reach its destination may not be a serious problem in low power radio networks but still in some event detecting networks it can prove critical. Through our simulations we have tried to see the extent of delay faced by different networks with varying number of nodes and packet size. It has been noticed that the average latency increases with increasing number of nodes. In all three cases the average latency increases as the network gets bigger. This can be explained with the argument that as the network becomes bigger, the number of nodes trying to transmit also increases resulting in more time taken by the node to access the medium. Hence before transmission a node may have to undergo multiple backoffs. Hence latency increases with larger number of nodes. Networks with smaller packet size has lesser latency than networks with bigger packets because bigger packets take longer time to be transmitted, making other nodes to wait longer before transmitting, thereby increasing latency in that case. However with smaller packets this period is small resulting in lower latency.

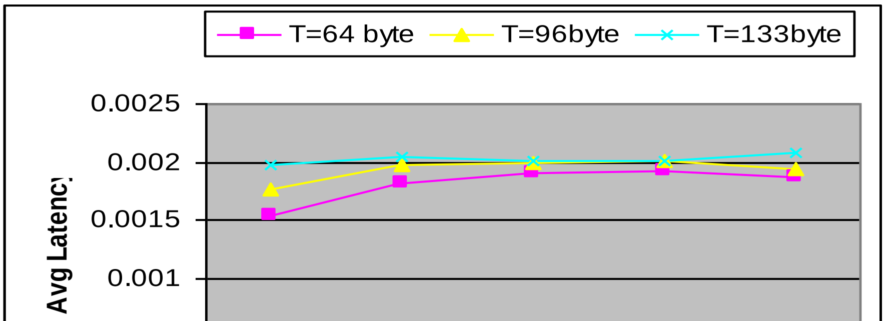
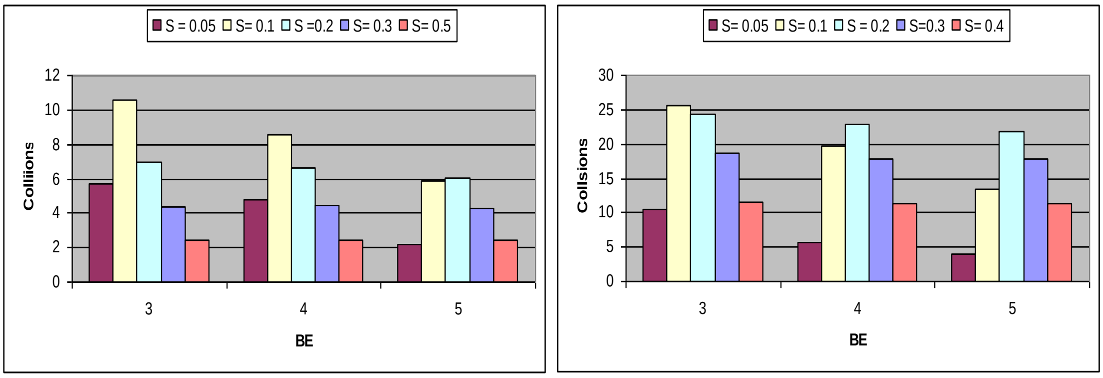


Fig 6.2 Average latency associated with different networks using different packet sizes

Another factor affecting the collision rate and average latency is the value of backoff exponential in *aMaxBE* which is the maximum value of the backoff exponent in the CSMA-CA algorithm used in IEEE 802.15.4 standard. The maximum value assigned to this is ‘5’. This exponent declares the number of backoff periods a device needs to wait before trying to access the channel for the transmission. This exponent value increases each time a device undergoes a backoff. Thus this increased backoff exponent can be a waste of energy but it also ensures lesser probability of collision. Hence by varying the value of this exponent we can try to find the affect on the number of collisions and hence find the optimum conditions for successful transmission. In this case the results are shown differently with each curve, depicting behaviour of a network with different number of nodes whilst changing values of backoff exponent. The graph depicts the number of collisions with smaller packets and changing value of *BE*. It can be observed that the number of collisions decreases for values of “S” as we increase the value of backoff exponent. This is due to the wider range in the number of backoff periods that a particular device has to wait before trying to transmit, which brings down the number of collisions. Also, for a particular value of *BE*, it can seen that collisions are higher with smaller value of “S” except that of S = 0.05. This is because this value of “S” is too small for 50 nodes to transmit a packet of 64 byte each. Therefore the periods finishes even before all the nodes have tried to send a packet, hence lesser collisions can be seen in case of S= 0.05. For all other values of “S” the number of collisions decreases because of the bigger transmission period. The same is true for the next graph, which depicts the collision count for a bigger network with 100 nodes.



(a) (b)

Fig 6.3 Collisions with varying value of BE (a) network of 50 nodes (b) 100 nodes

Although the number of collisions decreases as we increase the value of backoff exponent, the utilization of the superframe decreases. Utilization is the ratio of the time spent on sending successful data and total superframe length. It can be seen from the graph that the utilization decreases in all the cases because of the increase in time spent on backoffs as we increase the value of *BE*.

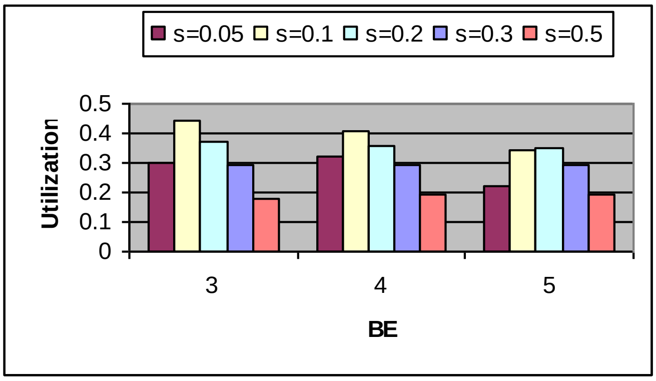
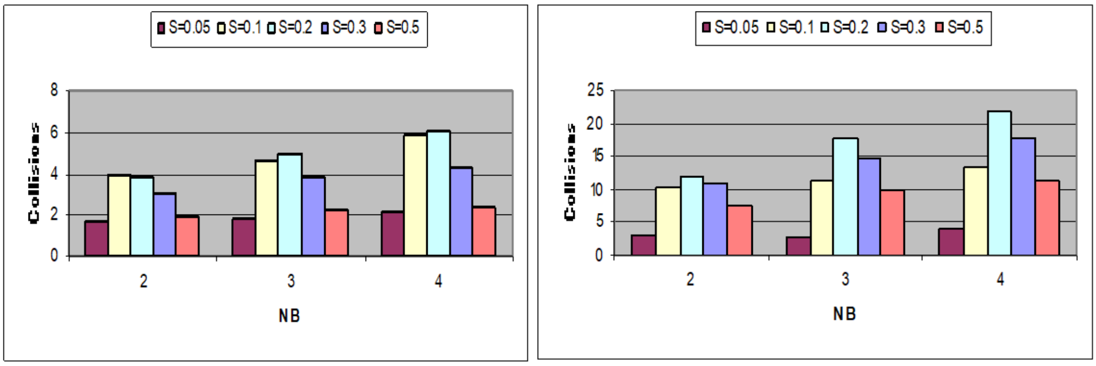


Fig 6.4 Utilization in case of network with 50 nodes and of changing value BE

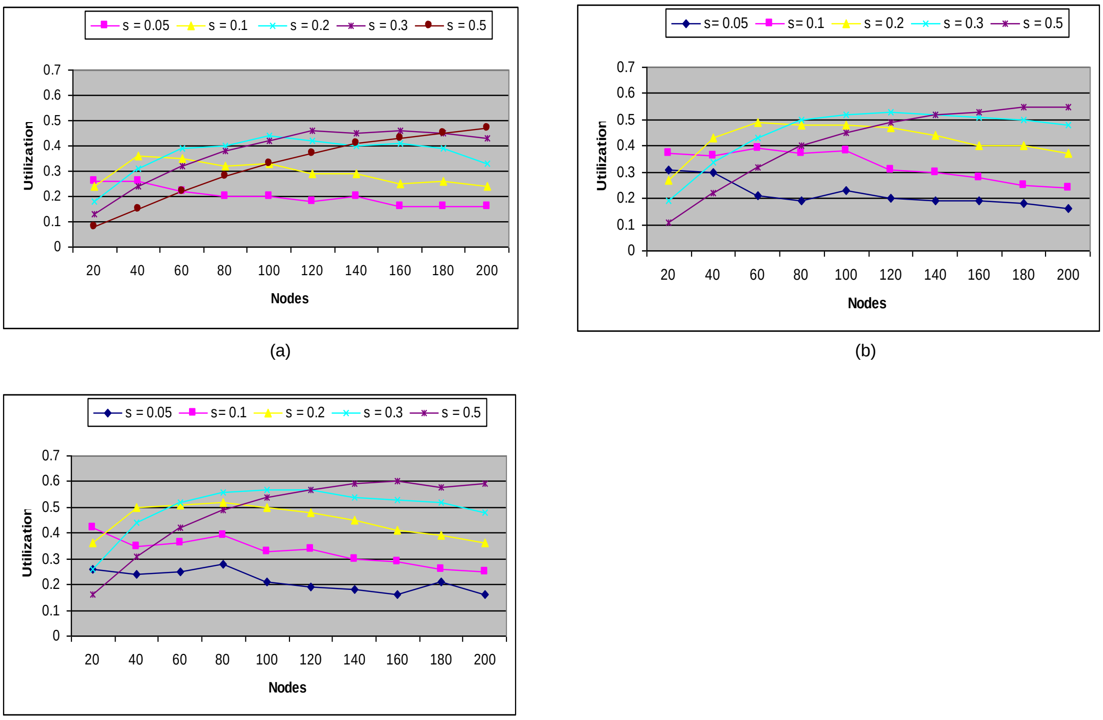
The effect of increasing the value of *macMaxCSMABackoffs* i.e. “*NB*” is discussed in the next graph. *macMaxCSMABackoffs* declares the maximum number of times a device can try to send a packet. The device exceeding this limit stops trying and declares channel access failure. It can be seen in the graph that the number of collisions is increasing as the value of "*NB*" increases because of an excessive try each time leading to increase in the number of collisions. The same reason supports the graphical results in case of adjoining graph for 100 nodes.



(a) (b)

Fig 6.5 Collisions with varying values of NB (a) 50 nodes (b) 100 nodes

The use of a proper superframe for the transmission of packets is a major concern while using different number of nodes. A smaller superframe size can result in insufficient space for data transfer of all the nodes and a large superframe can result in energy loss. To overcome these problems we have tried to depict the utilization of the superframe with different number of nodes and the same is repeated for different packet sizes. It can noted from these graphs that utilization graphs for each value of "S" reaches a peak point before starting to decline. So this peak point can be related to number of nodes for which that particular curve or value of "S" can give the maximum utilization. It can seen here that a smaller superframe suits best for a small network where it gives highest utilization as in case of network with 20 nodes but as we move towards bigger networks other curves start taking the lead. We can see that at 200 mark the curve supporting S = 0.5 gives maximum utilization making it a best option for networks with 200+ nodes.



(a)

(b)

(c)

Fig 6.6 Utilization of a network with different number of nodes with (a) 64 byte packet (b) 96 byte packet (c) 133 byte packet

6.6 Conclusion

The IEEE 802.15.4 CSMA/CA algorithm brings down the possibility of collisions due to its features such as CCA and random backoffs. The collisions are extremely low and latency associated with the transmission of packet is also very low. By varying parameters like '*NB*' and '*BE*' one can evaluate the optimum conditions for highest utilization and utilisation graphs can be used to evaluate the best value of superframe size that can be used for a give set on nodes.

Chapter 7

Topology Discovery

7.1 Introduction

Topology discovery is used in sensor networks to construct the topology of the network with respect to a single node. Two topology discovery algorithms are described here in detail. Simulation studies of these algorithms were carried out and the results are presented here.

7.2 Three-colour scheme

In this scheme nodes are considered to be of three colours: black, grey and white. White is an undiscovered node, meaning a node which has not received any topology discovery request [Hac03c]. A black node is a coordinator and a grey node is a node within range of a black node.

7.2.1 Procedure

In the three-colour scheme all the nodes are white at the initial state. A single node initiates the topology discovery procedure by sending a topology discovery request and turning black after broadcasting this request to all its neighbours. All the white nodes which receive a request packet turn grey. Hence a white node turns grey if it receives a request packet from a black node. Each grey node then broadcasts the request packet to all its neighbours with a random delay which is inversely proportional to the distance from which it has received the request packet. Thus, this inverse relation provides the opportunity for the farther nodes to disseminate the request packet first so as to increase the chances of greater coverage. When a white node receives a request packet from a grey node it turns black after a random delay which is inversely proportional to its distance from the grey node from which it has received its request packet. However if in the meantime a white node receives a request packet from a black node it turns grey. Once a node becomes grey or black it ignores all other discovery packets. The process can be described with the help of an example as shown below where node A becomes a black node by sending a request packet. The request reaches node B and node C, but its node B which forwards the packet before the node C because of its larger distance from the black node. Fig 7.1 shows a simple illustrative model for three-colour scheme.

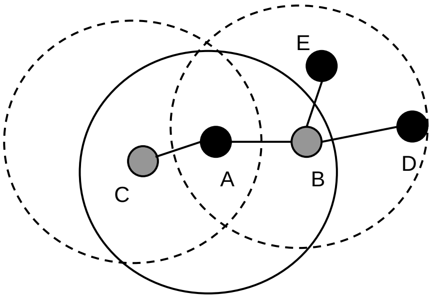


Fig 7.1 Simple illustrative diagram for three-colour scheme

7.3 Four-colour scheme

The four-colour scheme goes a step further in realising the objective to find a minimal cardinal set of cluster heads (coordinators). In the four-colour scheme nodes have four colours depending upon their roles and from which node it receives the request packet [Hac03c]. White is an undiscovered node which has not received any discovery request packet, black is a cluster head, grey is a node which has received a discovery request packet from at least one black node and dark grey is a node which receives a discovery packet from a grey node. Therefore a white node which is two hops away from a black node and receives a request packet from a grey node becomes a dark grey node. Therefore the clusters made by a four-colour scheme are spread further than the cluster made by the three-colour scheme. Fig 7.2 shows a simple illustrative model for four-colour scheme.

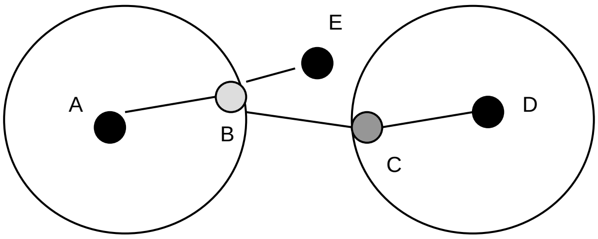


Fig 7.2 Simple illustrative diagram for four-colour scheme

7.3.1 Procedure

All the nodes are white at the beginning of the process. When a white node initiates the topology discovery by sending a discovery packet it turns black. The discovery packet is broadcasted to all its neighbours. All the neighbours who receive the discovery packet turn grey. All these grey nodes further broadcast the request to their neighbours after a random delay inversely proportional to their distance from the node from which they have received the request packet. When a white node receives a packet from a grey node it turns dark grey and broadcasts the request further to its neighbours. The dark grey node also starts a timer to become a black node. When a white node receives a request packet from a dark grey node it turns black with some random delay, but if in the meantime it receives a request packet from a black node it becomes grey again because it has been covered by a black and would be a waste to make one more cluster head. A dark grey node only waits for some limited time before it becomes a black node. Once the timer expires it becomes black as it is not covered by any black node. Once a node is grey or black, it ignores the requests from other nodes. This can be explained with the help of a simple example as shown in fig 7.2. The node A broadcasts a topology discovery request and turns black. The node B turns grey when it receives the request packet and further broadcasts the request packet to its neighbours, node E and node C. Once node E and node C receives the packet they turn dark grey and start a timer to turn black if they don’t receive any request from a black node before the timer expires. As can be seen in the diagram the node C is farthest from node B so it broadcasts the packet towards node D which turns black and rebroadcasts the request. The node C receives the request and stays dark grey but on the other side the node E which is closer to the node B forwards the request later and turns black after failing to hear from any other black node. Hence in this case the black nodes are two hops away from each other which mean that they are covering a greater area and a greater number of nodes.

7.4 Description of Simulation Model

The public domain ns simulator was used to create the simulation model for the experiments. The simulator was significantly extended to include a model of the IEEE 802.15.4 MAC and PHY layers and three-colour and four-colour models. The model did not use the physical distance of nodes in calculating the timeout value, but instead the received signal strength of the discovery packet, since in practice, sensor nodes would be unlikely to know their position, but real transceivers readily provide the received signal strength. Two scenarios were investigated regarding the placement of nodes. One, with a placement of nodes on a regular matrix and other was, random positioning of nodes within an area. The simulator was run, until no more discovery packets are transmitted. The observations in each row show the results from 100 runs, each using a different random number seed.

7.5 Simulation Results of Three-colour scheme.

The physical layer model used in the simulator gives a range of 180 m. In this first case we are dealing with a network of nodes arranged in a regular matrix and the distance between the nodes always remain constant. In this case the distance between the rows and columns is 20 m. Also the initiating coordinator in this case is placed at the corner of the network. Fig 7.3 shows the distribution of nodes and position of coordinator at the corner.

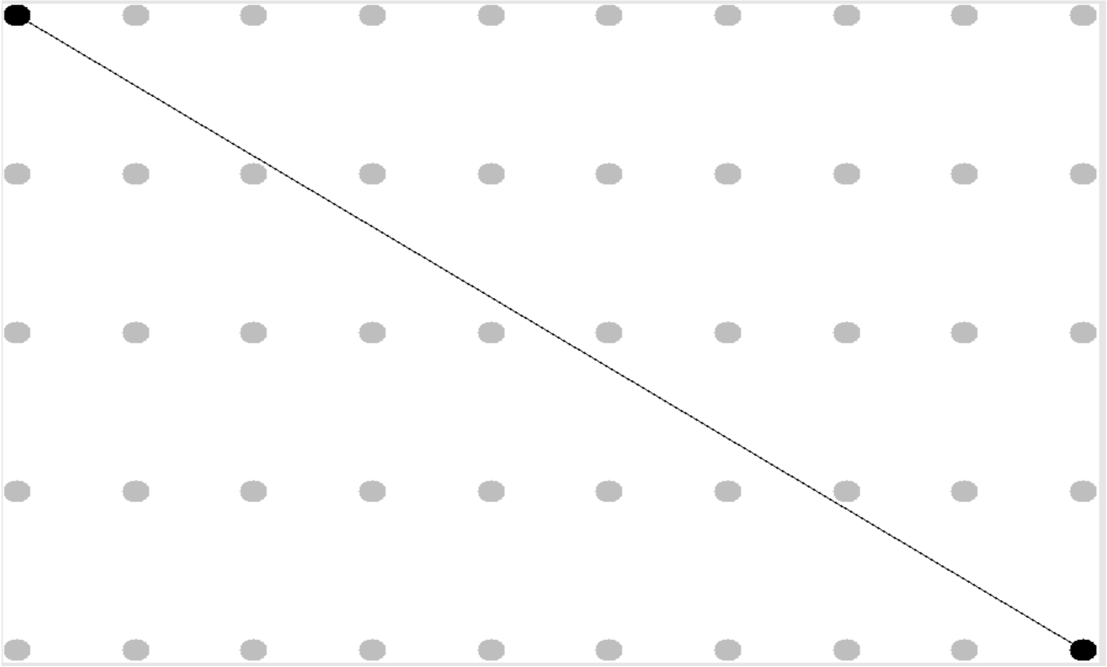


Fig 7.3 Node distribution with coordinator at the corner

The coordinator at the top corner is the initiating coordinator while the coordinator at the bottom corner was a white node which turned into coordinator as topology discovery progressed.

The table below shows the number of coordinators resulting from the simulations of the algorithm in case of networks with different number of nodes. Coordinators in these simulations depict the black nodes.

Nodes	Minimum Coordinators	Maximum Coordinators	Average Coordinators	Standard Deviation
50	2	4	3.31	0.706
100	5	25	14.79	6.045
150	12	74	35.81	15.603
200	17	78	48.74	17.993
250	21	98	53.04	16.150

Table 7.1 Coordinators with different network sizes and initiating coordinator at the corner

It can be seen that the average number of coordinators keeps increasing as we increase the number of nodes. It is worth mentioning in this case that we are just increasing the number of nodes, not the density of nodes in a given area. The standard deviation values in the above table shows some irregular increase which can be attributed to some important factors included in the simulation script. First of all, each of these observations is a result of 100 runs, fed with different randomising seeds to depict a true picture of the network. These random seeds affect the random delays involved in the three-colour scheme while assigning a white node as coordinator. The large standard deviation, evidenced again in the large difference between the minimum and maximum values for the set of simulations, shows the significant effect of the random process used in computing delays before retransmission of discovery packets. Hence in some cases a smaller delay can result in a large number of coordinators relative to the network size.

Keeping a coordinator at the corner of the network, means that a coordinator would cover only a quarter of the area of that covered by a coordinator at the centre. So the next simulation result depicts the case of networks with the initial coordinator at the centre, keeping all other features same. Fig 7.4 shows the distribution of nodes and coordinator at the centre.

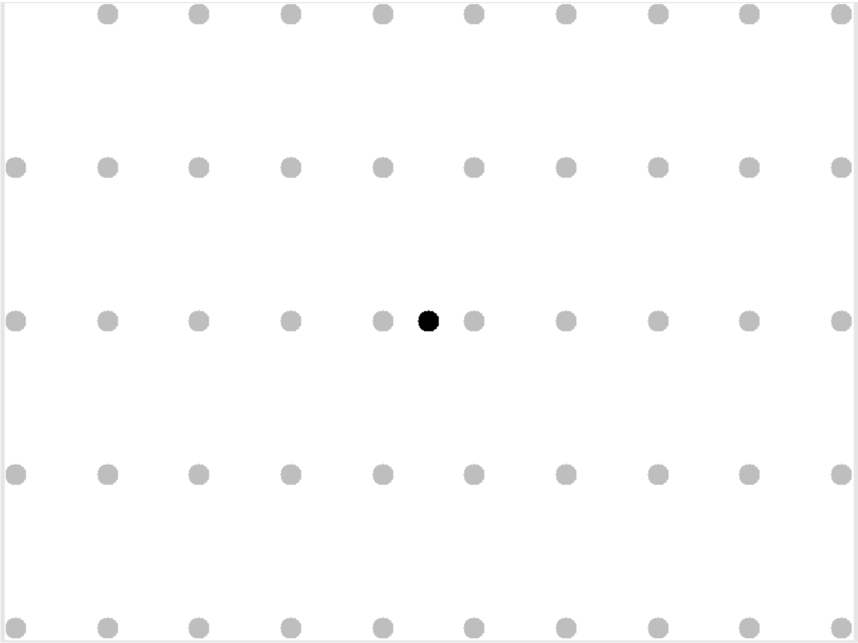


Fig 7.4 Node distribution with coordinator at the centre

Table 7.2 shows the number of coordinators when the initial coordinator is placed at the centre.

Nodes	Minimum Coordinators	Maximum Coordinators	Average	Standard Deviation
50	1	1	1	0
100	1	1	1	0
150	1	1	1	0
200	10	23	17.1	4.35
250	14	61	39.11	13.481

Table 7.2 Coordinators with different network sizes and initiating coordinator at the centre

It can be seen from the above results that after placing the coordinator at the centre the average number of coordinators goes down. The effect is huge in smaller networks where the coordinator can reach all the other nodes in a single hop. However in larger networks the effect is a moderate reduction in coordinators as nodes are arranged in regular fashion in this case and nodes equidistant from the grey node can turn into coordinators at the same time being equidistant from the grey node from which they are receiving the discovery packet, resulting in more coordinators.

The efficiency of the algorithm can be measured by the number of coordinators required to cover the whole network:

$$\text{Efficiency} = \text{nodes} \div \text{coordinators}$$

The graph in fig 7.5 shows the efficiency of the algorithm for different numbers of nodes. In all these cases the coordinator is at the corner. It can be seen that for large networks above 100 nodes, the algorithm results in about 5 nodes per coordinator.

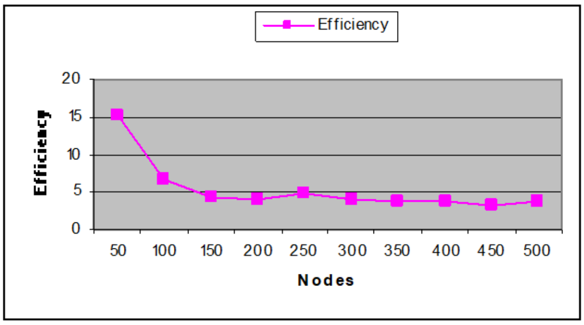


Fig 7.5 Graph depicting efficiency of networks with different sizes

It is not always possible to set up a real network in a regular array because of the hostile conditions of the terrain in which it is deployed. Therefore, in many cases the sensors would be deployed randomly without prior knowledge of their connectivity range. We have simulated some examples of random distribution of nodes to observe the behaviour of a set of nodes in different areas. The table below shows the average number of coordinators in case of a network with 50 nodes and initial coordinator at the corner.

Area Size	Minimum Coordinators	Maximum Coordinators	Average Coordinators	Standard Deviation	Unconnected
200x200	3	18	7.55	3.444	0
300x300	5	16	9.42	3.188	0
400x400	8	14	10.95	1.799	0
500x500	9	18	12.82	2.532	0
600x600	1	1	1	0	49

Table 7.3 Coordinators in case of network with 50 nodes and different distribution areas

In this case as we can see that increasing the area of the network brings up the number of coordinators which is because of greater distance apart of nodes making it difficult for the coordinator to cover more nodes in single hop. This trend follows till a certain limit and after that the nodes becomes so far apart that they are out of communicating range of each other resulting in all nodes being unconnected. This can also happen because of lack of bridging between two or more different set of nodes in the topology which can leave the entire network unconnected.

7.6 Simulation Results of Four-colour Scheme

The four-colour scheme is just an extension of the three-colour scheme. It is more efficient in the way that the coordinators in this case are at least two hops away from each other ensuring wider coverage of area with fewer coordinators.

In the first simulation result table we study the behaviour of different networks in a random array from 50 to 250 nodes. In the four-colour scheme there two types of connected states. One is depicted by "grey node" connected to the coordinator and the second is "dark grey node" which is connected to the grey node. These dark grey nodes are in a state where they can either become black nodes or grey nodes, because dark grey is a temporary state which continues until all nodes are discovered. Therefore, at the end all nodes in the network are either black or neighbours of black node i.e. grey nodes. It was noticed sometimes that dark grey nodes at the boundary of the network area fail to turn either black or grey till the end of the simulation. This happens because there are no white nodes after that boundary. Thus those dark grey nodes do broadcast the discovery packets but don't get any reply. If they neither get any discovery packet from any other black node, they remain in dark grey state, which is a temporary state. This can happen to all the nodes located near the boundary. One solution to this problem is to assign a timer after which they can become black, but this can increase the number of coordinators hugely, turning all dark grey nodes into black. This can bring up misleading results. Therefore in the following results, only grey nodes were taken into consideration. Fig 7.6 shows the distribution of nodes in four colour scheme and corresponding coordinators, grey nodes and dark grey nodes.

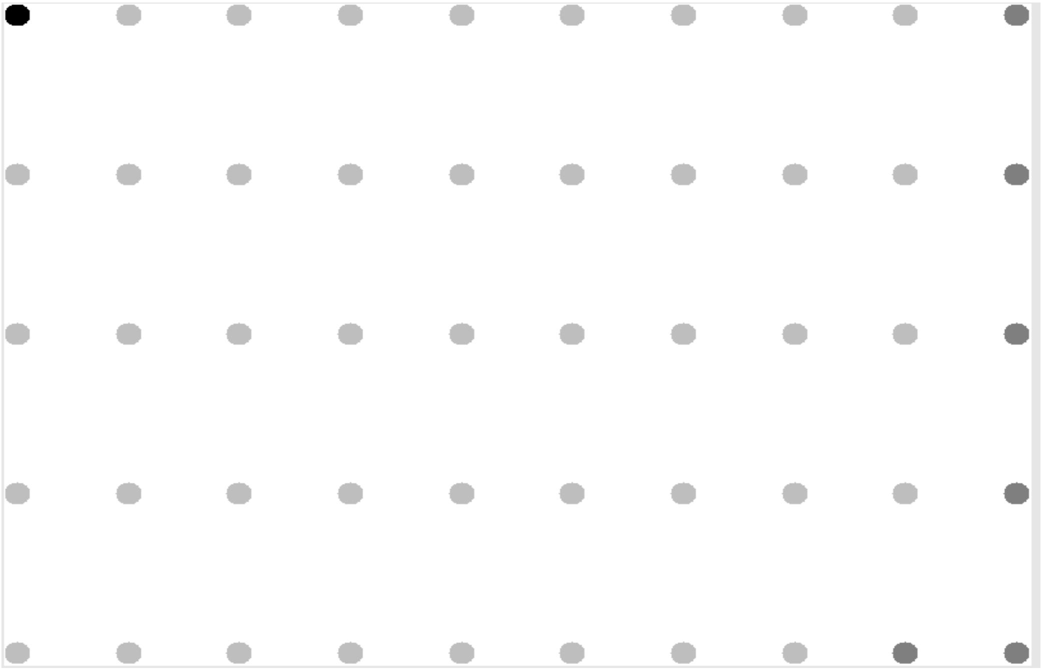


Fig 7.6 Node distribution and position of initial coordinator, grey nodes and dark grey nodes.

It can be noticed in fig 7.6 that all the dark grey nodes are at the boundary of the distribution area and hence unable to forward the discovery request packet. Therefore they remain in dark grey state.

In this case we are just taking into consideration, the grey nodes which are directly connected to the coordinator which is placed at the corner.

Nodes	Minimum Coordinators	Maximum Coordinators	Average Coordinators	Standard Deviation	Minimum Grey	Maximum Grey	Average Grey	Standard Deviation
50	1	1	1	0	43	43	43	0
100	1	1	1	0	68	68	68	0
150	1	23	3.22	3.44	68	134	89.18	30.43
200	5	65	25.59	12.58	93	181	155.95	16.74
250	4	78	32.78	16.67	112	225	193.38	22.53

Table 7.4 Coordinators with different network sizes and initiating coordinator at the corner

It is worth mentioning here that the distance between the regular arrays in this case is kept the same as it was in case of three-colour scheme. This helps to compare the results of both the schemes. The observations reveal that for smaller networks in this case there is only one coordinator which is quite similar to three-colour scheme. However the difference lies in that fact that there are some nodes in the dark grey state in four-colour scheme which are ready for further connectivity. As we move towards bigger networks the difference widens because the four-colour scheme is covering more nodes with fewer coordinators. The difference can be seen in case of networks with 150, 200 and 250 nodes and the results can be compared with readings of table 7.1. This decrease in the number of coordinators is because of the characteristics of the four-colour scheme which ensure two hops between the coordinators as compared to one hop in case of three-colour scheme. Thus each coordinator covers more nodes and brings down the number of coordinators to almost half.

The same experiment is repeated while placing the initiating coordinator at the centre. The table below shows the results.

Nodes	Minimum Coordinators	Maximum Coordinators	Average Coordinators	Standard Deviation	Minimum Grey	Maximum Grey	Average Grey	Standard Deviation
50	1	1	1	0	49	49	49	0
100	1	1	1	0	99	99	99	0
150	1	3	1.04	0.24	134	147	134.12	1.76
200	1	8	1.4	1.13	168	192	170.06	4.89
250	1	13	1.98	2.13	166	207	173.89	14.99

Table 7.5 Coordinators with different network sizes and initiating coordinator at the corner

The results show a huge difference from the 3 colour scheme. The number of coordinators has gone down greatly while maintaining a fair level of nodes in the grey state. The variation in the minimum and the maximum coordinators is because of the random number seed which affects the delay in each case. The huge difference can also be attributed to the fact that a large number of potential coordinators are in dark grey state which would raise the number of coordinators to some extent, but still keeping it considerably lower than three-colour scheme. For comparison between the two cases refer to table 7.2.

The maximum coverage area of a four-colour network with a given set of nodes has been found with the help of a randomly distributed network. By increasing the area we can discover how far a given set of nodes can be spread so that they remain connected. A further change has been made in this simulation by changing the position of the initial coordinator, which in this case is in the middle of one of the axes rather than at the centre or the corner. The table below shows the simulation results of a network comprising 50 nodes which are randomly distributed in square areas ranging from 200x200 metres to 1000x1000 metres. The coordinator in this case is in the middle of y-axis.

Area Size	Minimum Coordinators	Maximum Coordinators	Average Coordinators	Standard Deviation	Minimum Grey	Maximum Grey	Average Grey	Standard Deviation
200x200	1	1	1	0	38	38	38	0
400x400	5	16	8.54	2.40	35	44	39.34	3.30
600x600	5	14	9.42	1.82	24	40	31.51	3.21
800x800	2	13	7.5	2.23	9	37	26.29	9.61
1000x1000	4	5	4.07	0.25	9	10	9.93	0.25

Table 7.6 Coordinators in case of network with 50 nodes and different distribution areas

In the first case the area is small so the nodes are distributed in an area which can be covered almost completely by a single coordinator. As mentioned in the earlier part of this chapter the maximum range of a single node has been calculated to be around 180 metres which means that most of the nodes are in range of the coordinator. The number of grey nodes are however less because of the presence of nodes which are farther than 180 m range. These nodes are actually within range of some other nodes rather than directly in range of a coordinator and are left in the dark grey state. As we increase the area, the number of coordinators increases because of the inability of a single coordinator to cover other nodes beyond its coverage area and communication beyond second hops materialises, giving birth to new coordinators. However in the case of 1000 sq metre area the average has decreased because the nodes have gone so far apart that only a few are in communication range. This can be seen from the statistics showing the average connected nodes which in this case, is below 10.

In the last simulation we have investigated the effect of initial coordinator placement at different positions in a randomly distributed network. The number of nodes in this case is 50 and the area of distribution is 400 sq metres. Three different positions for the initial coordinators are: corner, centre and at the middle of the y-axis.

Coordinator Position	Minimum Coordinators	Maximum Coordinators	Average Coordinators	Standard Deviation	Minimum Grey	Maximum Grey	Average Grey	Standard Deviation
(0,0)	6	19	10.19	3.12	27	44	36.49	4.64
(0,200)	5	16	8.54	2.40	35	44	39.34	3.3
(200,200)	1	4	1.54	0.74	28	40	30.07	3.08

Table 7.7 Coordinators in case of network with 50 nodes and different coordinator positions

The table 7.7 shows the increased efficiency of the algorithm with initial coordinator at the centre. In this case, the network with initial coordinator at the corner has an average coordinator value of 10. This is because being at the corner the coordinator is in direct contract with only a few nodes and has to cover longer distance to cover widely distributed nodes. The coverage area in this case is 90 degrees. Bringing the coordinator at the centre of the axis helps the coordinator to influence nodes on both of its sides and spreads the coverage area to 180 degrees and hence the average number of coordinators comes down. The system reaches its peak efficiency when the coordinator is placed in the centre and coverage area becomes 360 degrees and most of the nodes come in direct contact with the coordinator. Therefore the average in this case is close to 2.

A comparison between the two topology discovery algorithms is shown with the help of graphs depicting the average number of coordinators with different initial coordinator positions. The distribution area in the cases is 400 sq metres.

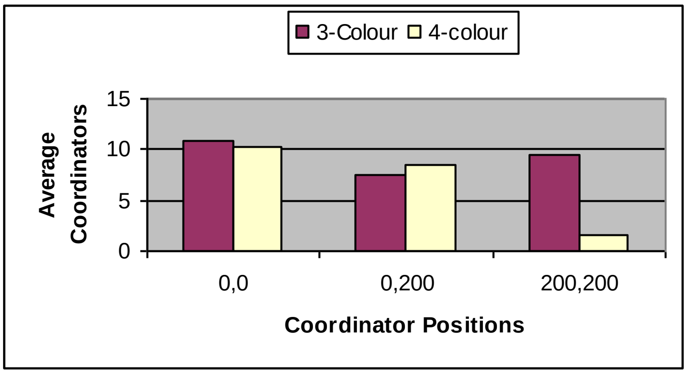


Fig 7.7 Average number of coordinators with 3-colour and 4-colour scheme

The graph shows the decrease of average number of coordinators as we move the coordinator towards the centre. The abnormal rise in average coordinators in case of 3-colour can be due to the selection of delay which is affected by the random seed used. The 4-colur scheme shows a considerable drop in average number of coordinators.

7.7 Conclusion

Both topology discovery techniques proved successful in discovering nodes and setting up cluster trees. The three-colour algorithm has only two states; coordinator and connected nodes while four-colour scheme has three states; coordinator, connected nodes and connecting nodes. Therefore three-colour scheme leaves the network fully connected although the number of coordinators is more than four-colour scheme. Four- colour scheme is efficient but may cause concern sometimes over the number of nodes which are in connecting state which can turn into coordinators at some stage and decrease the efficiency.

Chapter 8

Conclusion and future work

8.1 Conclusion

A background study of low power radio networks has been completed and discussed, along with applications and features of low power radio networks. Energy concerns of low power radio networks were studied and various factors affecting the energy consumption are explored. Techniques for energy conservation are also discussed. Various wireless technologies are mentioned along with their applications and their medium access methods. The IEEE 802.15.4 standard for low power radio networks is studied in detail and it’s Physical and MAC layer features are discussed along with different frame structures, topologies and commands supported by the standard. Its CSMA/CA algorithm is studied in detail and simulations carried out to shed light on its operation. Therefore through simulation results it was confirmed that by right selection of variables such as packet size, number of backoffs and backoff exponent, the number of collisions can be decreased. Topology discovery is a major concern in low power radio networks to render the network self configuring. Two different schemes for topology discovery are discussed and simulations performed to find the optimum results. The two algorithms are compared for a given set of parameters. Both topology discovery algorithms proved efficient for discovery procedures. Good selection of node and coordinator placement can further enhance the connectivity and efficiency of the algorithms.

8.2 Future Work

Future work needs to be done on the CSMA/CA algorithm to synchronize the sleep cycles of different nodes. In most low power radio networks, nodes are inactive for long times (sleep) and wake up only for a short time. For such cases there is an urgent need for the network nodes to have an advanced knowledge about the sleep cycles of neighbouring nodes so that data can be forwarded without delay.

The topology discovery algorithms also need some more work, particularly the four-colour algorithm which has a problem regarding nodes left in the dark grey state. The delay factor in these topology discovery schemes plays an important part, so a more detailed study of this is called for.

References

[Call04a] **Callaway, Edgar H**. Wireless Sensor Networks: Architectures and Protocols. CRC Press, USA 2004. pp 1-16

[Call04b] pp 88-89

[CE04] **Callaway, E**. Secure Low Power of Wireless Sensor Networks. Sensors Magazine. Jan 2004

http://www.sensorsmag.com/articles/0104/22/

[Dell01] **Dell.** Deploying 802.11b (Wi-Fi) in the Enterprise Network. White Paper by DELL^{TM}. April 2001.

http://www.dell.com/downloads/global/vectors/wireless_deployment.pdf

[DM03] **Dallas Maxim**, Wireless, RF and Cable. An Introduction to Direct-Sequence Spread-Spectrum Communications. Application note 1890: Feb, 2003.

[DTIB] **DTI Bluetooth** Factsheet.

http://www.dti.gov.uk/bestpractice/assets/bluetooth.pdf

[ECE4321] **ECE4321 Computer Networks**. SpreadSpectrum

http://www.d.umn.edu/~tkwon/course/4321/PPP/Lec22-SpreadSpectrum.ppt

[Eph02] **Ephremides, A**. Energy Concerns in Wireless Networks. IEEE Wireless Communications, Vol 9, No 4, August 2002, pp 48-59.

[Estrin02] **Estrin, D**. Sensor Network Protocols.

http://nesl.ee.ucla.edu/tutorials/mobicom02/slides/Mobicom-Tutorials-4-DE.pdf

[FCC] **FCC Code of Federal Regulations** 47, Part 15.

http://wireless.fcc.gov/rules.html

[Gan03] **Ganeshan, D**; Cerpa, A; Ye, W; Zhao, J; Estrin, D. Networking Issues in Wireless Sensor Networks. June 2003.

[Gero05] **Gerold, J**. Networks create low-cost sensor links. Automation World, Jan 2005.

http://www.automationworld.com/articles/Features/102.html

[Gor02] **Gorday, P**; Callaway, E; Hester, L; Gutierrez, Jose A; Naeve, M; Heile, B; Bahl, V. Home Networking with IEEE 802.15.4: A Developing Standard for Low-Rate Wireless Personal Area Networks. IEEE Communications Magazine, Aug 2002, Vol 40, No. 8, pp 70-77.

[Gre02] **Greier, J**. 802.11 Medium Access Control Methods. WiFi Planet, 26 Nov, 2002.

http://www.wi-fiplanet.com/tutorials/article.php/1548381

[Gut03a] **Gutierrez, J**; Callaway Jr, Edgar H; Barret Jr, Raymond L. Low-Rate Wireless Personal Area Networks- Enabling Wireless Sensors with IEEE 802.15.4TM. IEEE Standards Wireless Networks Series. Standard Information Network IEEE Press, USA, Nov 2003, pp. 10

[Gut03b] pp. 4

[Gut03c] pp. 13-21

[Gut03d] pp. 26-27

[Gut03e] pp. 25-58

[Gut03f] pp. 59-93

[Gut03g] pp. 48

[Hac03a] **Hac, A**. Wireless Sensor Network Designs. John Wiley & Sons Ltd, England 2003. pp-63

[Hac03b] pp-169-177

[Hac03c] pp-171-180

[Harte04a] **Harte, L**. Introduction to Bluetooth: Technology, Market, Operation, Profiles, & Services. Althos Publishing, 2004, USA, pp 8-13.

[Harte04b] pp 28-47

[Heid01] **Heidemann, J** and Ye, W. Medium Access Control in Wireless Sensor Networks. Oct 2001.

http://www.isi.edu/~weiye/pub/isi-tr-580.pdf

[HJ 98] **Haartsen, J;** Allen, W; Inouye, J; Joeressen, Olaf J; Naghshineh, M. Bluetooth: Vision, Goals, and Architecture. Mobile Computing and Communications Review, Oct 1998, Vol 2(4), pp 38-45.

http://citeseer.ist.psu.edu/haartsen98bluetooth.html

[Hod03] **Hodgdon, C**. Adaptive Frequency Hopping for Reduced Interference between Bluetooth and Wireless LAN. D & R Industry Articles. May 2003.

http://www.us.design-reuse.com/articles/article5715.html

[IEEE99] **ANSI/IEEE Standard 802.11.** Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Edition 1999 (R2003).

[IEEE99a] pp. 70-73

[IEEE03a] IEEE Standard for Information Technology. Part 15.4-MAC and PHY Specifications for Low-Rate Personal Area Networks (LR-WPANs) 2003. pp 17

[IEEE03b] pp 13

[IEEE03c] pp 23-24.

[IEEE03d] pp 160-164

[IEEE03e] pp 142-144

[IEEE03f] pp 14-15

[IEEE03g] pp 21-23

[IEEE03h] pp 29-54

[IEEE03i] pp 55-109

[Int03] **Intanagonwiwat, C**; Govindan, R; Estrin, D; Heidemann, J; Silva, F. Directed Diffusion for Wireless Sensor Networking. ACM/IEEE Transactions on Networking, Feb 2003, vol. 11, Issue 1, pp 2-16.

http://www.isi.edu/~johnh/PAPERS/Intanagonwiwat03a.html

[Jan03] **Jansen, R**; Hanemann, S; Freisleben, B. Proactive Distance-Vector Multipath Routing for Wireless Ad Hoc Networks. Proceedings of the IASTED International Conference on Communication Systems and Networks 2003 (ICCSN2003), Benalmadena, Spain, ACTA Press, pp. 1-6,

[Jin04] **Jinag, Q**; Manivannan, D. Routing Protocols for Sensor Networks. Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC 2004). Jan 2004, pp 93-98.

[Miller02a] **Miller, Brent A; Bisdikian, C**. Bluetooth Revealed, The Insider Guide to an Open Specification for Global Wireless Communication. Edition 2, Printice Hall PTR, 2002, pp 6

[Miller02b] pp 29-47

[Miller02c] pp 25-28

[Miller02d] pp 63-78

[MMC99] **Kalia, M**; Bansal, D; Shorey, R. MAC Scheduling and SAR Policies for Bluetooth: A Master Driven TDD Pico-Cellular Wireless System. In Proceedings of IEEE International Workshop on Mobile Multimedia Communications (MoMuC), Nov 1999, San Diego, USA.

[Pao02] **Paolini, M**; Farrar, T; Pow, R. The WLAN Opportunity for Wireless Service Providers. Analysys for Transat Technologies, 2002.

http://www.analysys.com/pdfs/wlan.pdf

[Par04] **Parker, Andrew D**. A Guide for the Clueless: IEEE 802.15.4 Standard for Low-Rate Wireless Personal Area Networks (LR-WPAN). July 14, 2004. http://lecs.cs.ucla.edu/~adparker/EE202A/hw2/

[PCS02] **PalChaudhari, S**; Johnson, David B. Power Scheduling for Ad-Hoc Networking. Proceedings of the 10th IEEE International Conference on Network Protocols (ICNP'02)

[Rab00] **Rabaey, J**; Ammer, M.Josie; Silve, J; Patel, D; Roundy, S. PicoRadio Supports Ad Hoc Ultra-Low Power Wireless Networking. Computer Magazine, July 2000, vol. 33, No. 7, pp 42-48.

[Rad04] **Radio-Electronics.Com** :: Zigbee and IEEE – for remote data sensor and control applications. Oct 11, 2004. http://www.radio-electronics.com/info/wireless/zigbee/zigbee.php

[Reid04] **Riedel, T**. Self-organising, Wireless Sensor Networks. Remote Site & Equipment Management Magazine, April/May 2004.

[Rep04] **Report: Zigbee-** the next revolution in wireless technology. Sep 28, 2004. http://www.windowsfordevices.com/news/NS4235010315.html

[Rao01] **Rao, R**; Baux, O; Kesidis, G. Demand-based Bluetooth Scheduling. Presented at the 3rd IEEE Wireless LAN (WLAN) Conference. Sept. 2001,Boston, MA. http://labs.ee.psu.edu/faculty/kesidis/public_html/Bluetooth1.doc

[Sen02] **Senkowski, R.** Michael; DeSilva, E and Dombrowsky, T. WiFi – 802.11, The Shape of Things to Come. Wiley Rein & Fielding, July 2002. http://www.wrf.com/wifi.pdf

[Stoj01] **Stojmenovic, I** and Lin, X. Power-Aware Localized Routing in Wireless Networks. IEEE Transaction on Parallel and Distributed Systems 2001

[TG404] **IEEE 802.15.4 WPAN^{TM} Group 4 (TG4).** Sep 29, 2004. http://www.ieee802.org/15/pub/TG4b.html

[Til02] **Tilak, S**; Abu-Ghazaleh, Nael B; Heinzelman, W. A Taxonomy of Wireless Micro-Sensor Network Models. Mobile Computing and Communications Review, Vol 6, No.2, April 2002.

[WAlli] **WiFi Alliance**. Why is Wi-Fi Important? http://www.wi-fi.org/opensection/wi-fi_important.asp

[Wei02] **Wei, Y;** Heidemann, J; Estrin, D. An Energy-Efficient MAC protocol for Wireless Sensor Networks. In Proceedings of the IEEE Infocom, pp. 1567-1576. New York, NY, USA, USC/Information Sciences Institute, IEEE. June, 2002. http://www.isi.edu/~johnh/PAPERS/Ye02a.html

[WFS04] **WiFi Fact Sheet**. Department of Trade and Industry (dti), April 2004. http://www.dti.gov.uk

[WWP03] **WiFi White Paper**. mwr InfoSecurity Limited, 2003. http://searchnetworking.techtarget.com/searchNetworking/Downloads/mwr.wifi.pdf

[You04] **Young, E**. Stop the rot. The Guardian. 1st July, 2004. pp 8-9