A NOTES ON WIRELESS TECHNOLOGY AND MOBILE COMPUTING

Edited By :

Md. Azaz

(Admin way2mca)

VESIT

Shared by :

VESIT MCA-2009-12 batch

WIRELESS TECHNOLOGY NOTES

Chapter 1

1.1.1 Introduction

Radio or wireless–there must be a big difference; they are spelled a little different. I hate to disappoint and disillusion some of you who have counted so much on a 'big difference'. However, just brace yourselves and prepare for the shock: *there is absolutely no difference between radio and wireless except the spelling*. Wireless does not mean sparks, noise, or a lot of switches. Wireless means communication without the use of wires other than the antenna, the ether, and ground taking the place of wires. Radio means exactly the same thing: it is the same process. Communications by wireless waves may consist of an SOS or other messages from a ship at sea or the communication may be simply the reception of today's top 10 music artists, or connecting to the Internet to check your email. It does not become something different in either spelling or meaning.

1.1.2 Where it all began – Marconi

In February 1896, Guglielmo Marconi journeyed from Italy to England in order to show the British telegraph authorities what he had developed in the way of an *operational* wireless telegraph apparatus. His first British patent application was filed on June 2 of that year. Through the cooperation of Mr. W.H. Preece, who was at that time the chief electrical engineer of the British Post-office Telegraphs, signals were sent in July 1896 over a distance of one-and-three-fourths miles on Salisbury Plain.

1.1.3 Packet Data

Packet Data technology was developed in the mid-1960s and was put into practical application in the ARPANET, which was established in 1969. Initiated in 1970, the ALOHANET, based at the University of Hawaii, was the first large-scale packet radio project. Amateur packet radio began in Montreal, Canada, in 1978 with the first transmission occurring on May 31. This was followed by the Vancouver Amateur
Digital Communication Group (VADCG) development of a Terminal Node Controller (TNC) in 1980. The current TNC standard grew from a discussion in October of 1981 at a meeting of the Tucson Chapter of the IEEE Computer Society. A week later, six of the attendees gathered together and discussed the feasibility of developing a TNC that would be available to amateurs at a modest cost. The Tucson Amateur Packet Radio Corporation (TAPR) was formed from this project. On June 26 1982, Lyle Johnson and Den Connors initiated a packet contact with the first TAPR unit. The project progressed from these first prototype units to the TNC-1 and then finally to the TNC-2 which is now the basis for most packet operations worldwide. Packet has three great advantages over other digital modes: transparency, error

correction, and automatic control. The operation of a packet station is transparent to the end user. Connect to the other station, type in your message, and it is sent automatically. The Terminal Node Controller (TNC) automatically divides the message into packets, keys the transmitter, and then sends the packets. While receiving packets, the TNC automatically decodes, checks for errors, and displays the received messages. Packet radio provides error-free communications due to the built-in error detection schemes. If a packet is received, it is checked for errors and will be displayed only if it is correct. In addition, any packet TNC can be used as a packet relay station, sometimes called a digipeater. This allows for greater range by stringing several packet stations together. Users can connect to their friends' TNCs at any time they wish, to see if they are at home. Another advantage of packet over other modes is the ability for many users to be able to use the same frequency channel simultaneously. Since packet radio is most commonly used at the higher radio frequencies (VHF), the range of the transmission is somewhat limited. Generally, transmission range is limited to 'unobstructed line-of-sight' plus approximately 10–15% additional distance.

The transmission range is influenced by the transmitter power and the type and location of the antenna, as well as the actual frequency used and the length of the antenna feed line (the cable connecting the radio to the antenna). Another factor influencing the transmission range is the existence of obstructions

(hills, groups of buildings, etc.). Connections made in the 144–148 Mhz range could be 10 to 100 miles, depending on the specific combination of the variables mentioned above.

1.1.4 Voice Technologies

In the November 7, 1920 issue of the Boston *Sunday Post* there was an article authored by John T. Brady covering the topic of 'Talking by Wireless as You Travel by Train or Motor,' which noted 'It is now possible for a business man to talk with his office from a moving vehicle.' This was a review of two-way radio conversation tested by Mr. Brady and with Harold J. Power who was then the head of the American Radio and Research Corporation, while Power was in a moving automobile. It would not be until the 1980s that the technology needed for such things as pagers and wireless telephones would be perfected to the point that they became widely available consumer products. Although the telephone's use for individual communication largely overshadowed applications for distributing entertainment and news, the reverse would be true for radio, with broadcasting dominating for decades, before radio transmissions would be significantly developed for personal, mobile communication.

1.1.5 Cellular Technologies

In cellular networks there are radio ports with antennas that connect to base stations (BSs) that serve the user equipment known as mobile stations (MSs). The communication that takes place from the MS to the BS is known as the uplink while the communication from the BS to the MS is known as the downlink. The downlink is contention-less, however several MSs access the uplink simultaneously. This uplink uses a very important characteristic, which is the multiple-access technique. Frequency-division multiple access (FDMA), time-division multiple access (TDMA) and code-division multiple access (CDMA) are the most widely used physical-layer multiple access techniques in use today. The infrastructures of cellular networks include mobile switching centers (MSCs). These control one or more BSs and provide the interface for them to the wired public switched telephone network (PSTN), a central home location register (HLR)

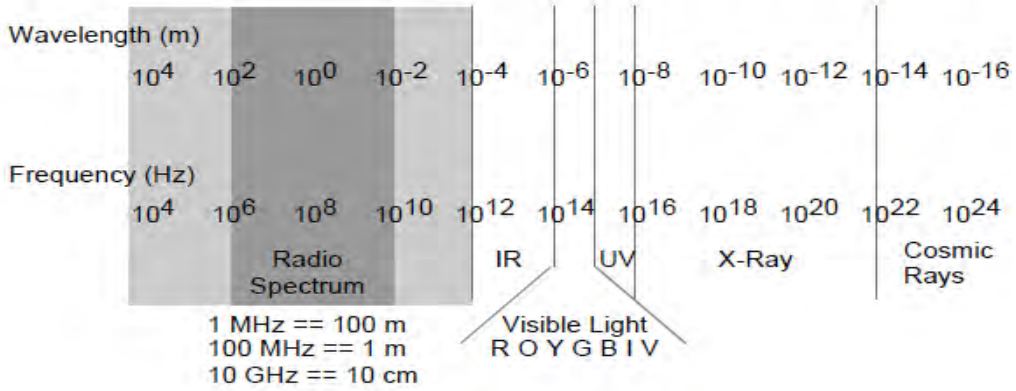
and the visiting location register (VLR) for each MSC. The VLR and HLR are databases that keep the registered and current locations of MSs to be used in the handoffs. Handoff is the process of handing a call from one cell to a new cell as the MS moves around.

1.2 Wireless Communication Technology

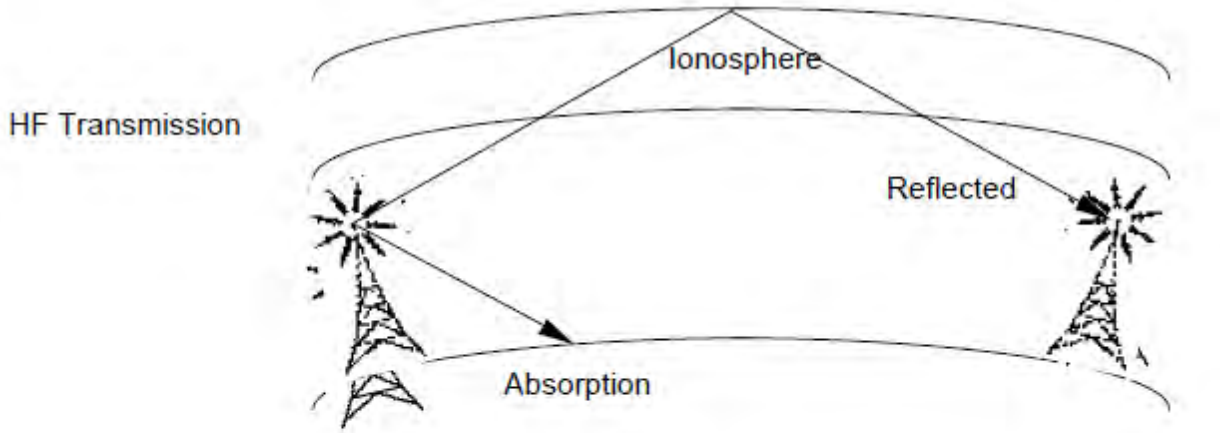
Wireless communication is the transfer of information over a distance without the use of enhanced electrical conductors or "wires". The distances involved may be short (a few meters as in television remote control) or long (thousands or millions of kilometers for radio communications). When the context is clear, the term is often shortened to "wireless". Wireless communication is generally considered to be a branch of telecommunications. It encompasses various types of fixed, mobile, and portable two-way radios, cellular telephones, personal digital assistants (PDAs), and wireless networking. Other examples of wireless technology include GPS units, garage door openers and or garage doors, wireless computer mice, keyboards and headsets, satellite television and cordless telephones.

The wireless communications technology covers all major wireless technologies: Wireless LAN (WLAN), WiMAX for metropolitan area networking (WMAN), Bluetooth, ZigBee and UWB for personal area networking (WPAN) and mobile wireless technologies (WWAN) from 1G to 3G. A comprehensive wireless communications technology guide for network and telecom professionals. Graphic illustration in diagrams and evolution tree of mobile wireless technologies from 1 G to 4G for both GSM/GPRS/WCDMA and cdmaOne/CDMA2000;. Comprehensive reference architecture framework to show all wireless technologies in one chart for quick understanding of the big picture. Designed by experts with decades of experience in wireless, packet and tele- communication industries.

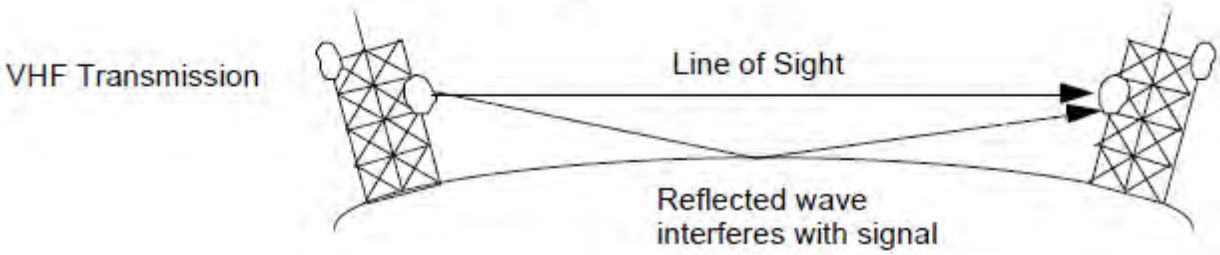
1.3. **Transmission Fundamentals Radio Basics**



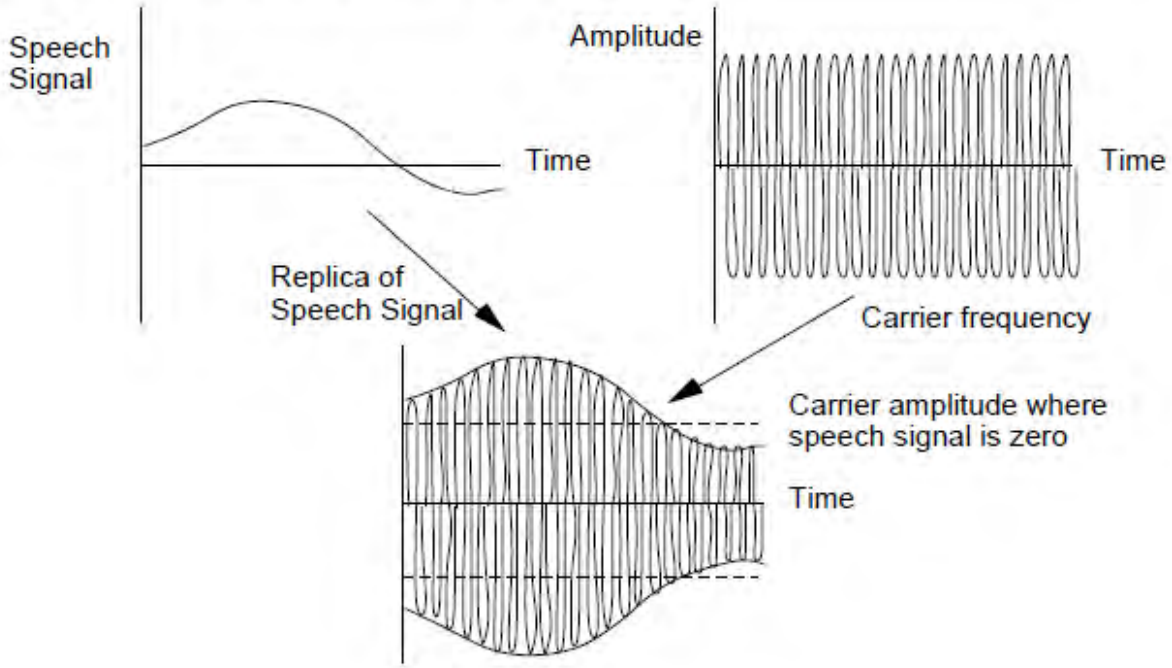
< 30 KHz	VLF
30 - 300 KHz	LF
300 KHz - 3 MHz	MF
3 - 30 MHz	HF
30 - 300 MHz	VHF
300 MHz - 3 GHz	UHF
3 - 30 GHz	SHF
> 30 GHz	EHF



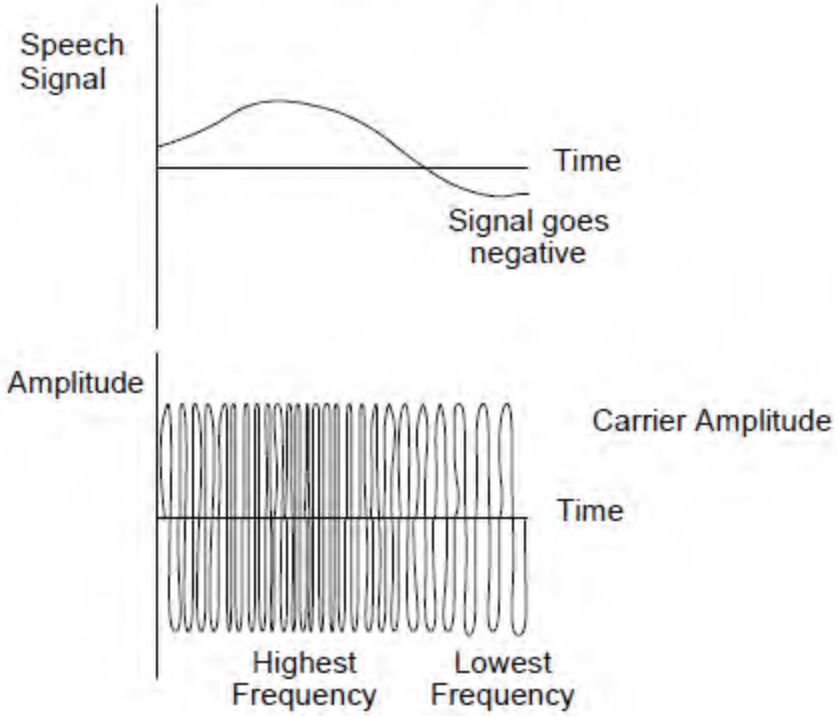
Directional Antenna



Amplitude Modulation (AM)



Frequency Modulation (FM)



Digital Modulation Techniques

• Carrier wave s:

– s(t) = A(t) * cos[(t)]

– Function of time varying amplitude A and time varying angle Θ

• Angle Θ rewritten as:

– Θ(t) = ω0 + φ (t)

– 0 radian frequency, phase (t)

• s(t) = A(t) cos[ω0t + φ (t)]

– ω radians per second

– relationship between radians per second and hertz

» $\omega = 2 \pi f$

Digital Modulation Techniques

• Demodulation

– Process of removing the carrier signal

• Detection

– Process of symbol decision

– Coherent detection

» Receiver users the carrier phase to detect signal

» Cross correlate with replica signals at receiver

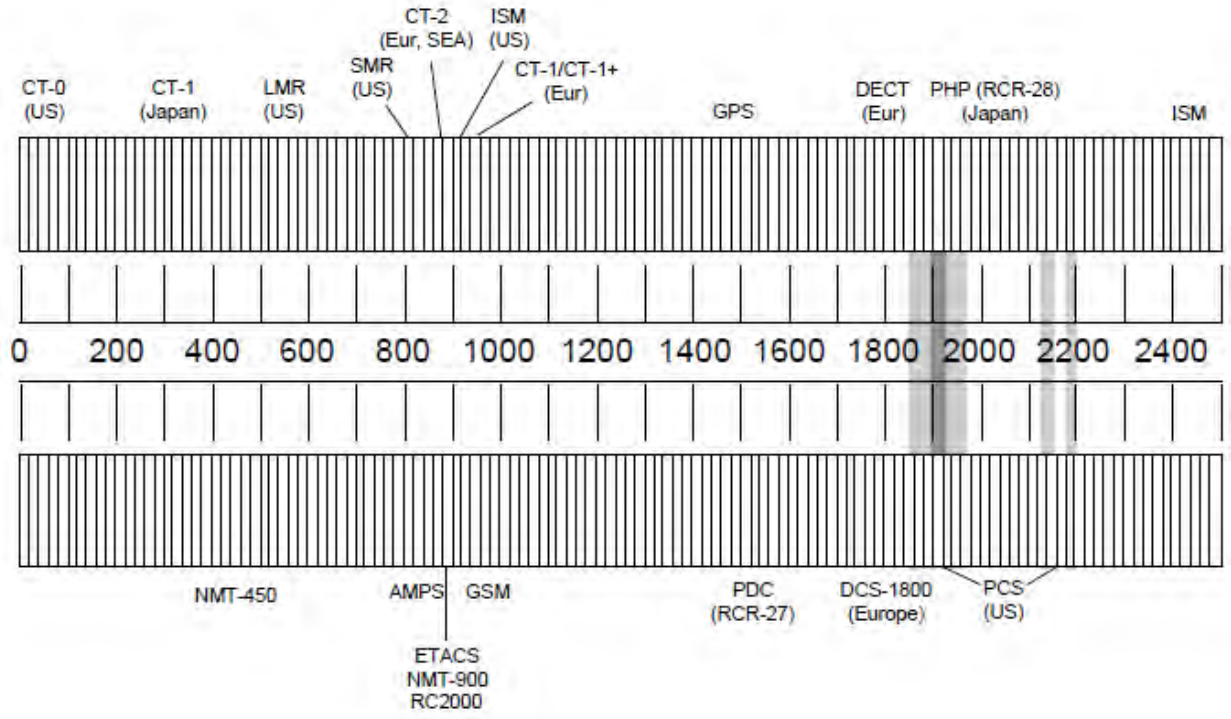
» Match within threshold to make decision

– Noncoherent detection

» Does not exploit phase reference information

» Less complex receiver, but worse performance

Wireless Spectrum



1.4 TCP/IP

Computer Communication Protocol

A computer communication protocol is a description of the rules computers must follow to communicate with each other.

What is TCP/IP?

TCP/IP is the communication protocol for communication between computers on the Internet.

TCP/IP stands for Transmission Control Protocol / Internet Protocol.

TCP/IP defines how electronic devices (like computers) should be connected to the Internet, and how data should be transmitted between them.

Inside TCP/IP

Inside the TCP/IP standard there are several protocols for handling data communication:

TCP (Transmission Control Protocol) communication between applications
UDP (User Datagram Protocol) simple communication between applications

IP (Internet Protocol) communication between computers
ICMP (Internet Control Message Protocol) for errors and statistics
DHCP (Dynamic Host Configuration Protocol) for dynamic addressing
TCP Uses a Fixed Connection

TCP is for communication between applications.

If one application wants to communicate with another via TCP, it sends a communication request. This request must be sent to an exact address. After a "handshake" between the two applications, TCP will set up a "full-duplex" communication between the two applications.

The "full-duplex" communication will occupy the communication line between the two computers until it is closed by one of the two applications.

UDP is very similar to TCP, but simpler and less reliable.

IP is Connection-Less

IP is for communication between computers.

IP is a "connection-less" communication protocol.

IP does not occupy the communication line between two computers. IP reduces the need for network lines. Each line can be used for communication between many different computers at the same time.

With IP, messages (or other data) are broken up into small independent "packets" and sent between computers via the Internet.

IP is responsible for "routing" each packet to the correct destination.

IP Routers

When an IP packet is sent from a computer, it arrives at an IP router.

The IP router is responsible for "routing" the packet to the correct destination, directly or via another router.

The path the packet will follow might be different from other packets of the same communication. The router is responsible for the right addressing, depending on traffic volume, errors in the network, or other parameters.

Connection-Less Analogy

Communicating via IP is like sending a long letter as a large number of small postcards, each finding its own (often different) way to the receiver.

TCP/IP

TCP/IP is TCP and IP working together.

TCP takes care of the communication between your application software (i.e. your browser) and your network software.

IP takes care of the communication with other computers.

TCP is responsible for breaking data down into IP packets before they are sent, and for assembling the packets when they arrive.

IP is responsible for sending the packets to the correct destination.

IP Addresses

Each computer must have an IP address before it can connect to the Internet.

Each IP packet must have an address before it can be sent to another computer.

This is an IP address: 192.68.20.50
This might be the same IP address: www.abc.com

An IP Address Contains 4 Numbers.

Each computer must have a unique IP address.

This is your IP address: 183.87.203.156

TCP/IP uses four numbers to address a computer. The numbers are always between 0 and 255.

IP addresses are normally written as four numbers separated by a period, like this: 192.168.1.50.

32 Bits = 4 Bytes

TCP/IP uses 32 bits addressing. One computer byte is 8 bits. So TCP/IP uses 4 computer bytes.

A computer byte can contain 256 different values:

00000000, 00000001, 00000010, 00000011, 00000100, 00000101, 00000110, 00000111, 00001000 .....and all the way up to 11111111.

Now you know why a TCP/IP address is four numbers between 0 and 255.

Domain Names

A name is much easier to remember than a 12 digit number.

Names used for TCP/IP addresses are called domain names.

abc.com is a domain name.

When you address a web site, like http://www.abc.com, the name is translated to a number by a Domain Name Server (DNS).

All over the world, DNS servers are connected to the Internet. DNS servers are responsible for translating domain names into TCP/IP addresses.

When a new domain name is registered together with a TCP/IP address, DNS servers all over the world are updated with this information.

TCP/IP is a large collection of different communication protocols.

Application Layer	**HTTP FTP Telnet Finger SSH DNS POP3/IMAP SMTP Gopher BGP Time/NTP Whois TACACS+ SSL**	**DNS SNMP RIP RADIUS Archie Traceroute tftp**	**Ping**		
Transport Layer	**TCP**	**UDP**	**ICMP**	**OSPF**	
Internet Layer	**IP**				**ARP**
Network Interface Layer	**Ethernet/802.3 Token Ring (802.5) SNAP/802.2 X.25 FDDI ISDN Frame Relay SMDS ATM Wireless (WAP, CDPD, 802.11) Fibre Channel DDS/DS0/T-carrier/E-carrier SONET/SDH DWDM PPP HDLC SLIP/CSLIP xDSL Cable Modem (DOCSIS)**				

FIGURE 2. Abbreviated TCP/IP protocol stack.

A Family of Protocols

TCP/IP is a large collection of different communication protocols based upon the two original protocols TCP and IP.

TCP - Transmission Control Protocol

TCP is used for transmission of data from an application to the network.

TCP is responsible for breaking data down into IP packets before they are sent, and for assembling the packets when they arrive.

IP - Internet Protocol

IP takes care of the communication with other computers.

IP is responsible for the sending and receiving data packets over the Internet.

HTTP - Hyper Text Transfer Protocol

HTTP takes care of the communication between a web server and a web browser.

HTTP is used for sending requests from a web client (a browser) to a web server, returning web content (web pages) from the server back to the client.

HTTPS - Secure HTTP

HTTPS takes care of secure communication between a web server and a web browser.

HTTPS typically handles credit card transactions and other sensitive data.

SSL - Secure Sockets Layer

The SSL protocol is used for encryption of data for secure data transmission.

SMTP - Simple Mail Transfer Protocol

SMTP is used for transmission of e-mails.

MIME - Multi-purpose Internet Mail Extensions

The MIME protocol lets SMTP transmit multimedia files including voice, audio, and binary data across TCP/IP networks.

IMAP - Internet Message Access Protocol

IMAP is used for storing and retrieving e-mails.

POP - Post Office Protocol

POP is used for downloading e-mails from an e-mail server to a personal computer.

FTP - File Transfer Protocol

FTP takes care of transmission of files between computers.

NTP - Network Time Protocol

NTP is used to synchronize the time (the clock) between computers.

DHCP - Dynamic Host Configuration Protocol

DHCP is used for allocation of dynamic IP addresses to computers in a network.

SNMP - Simple Network Management Protocol

SNMP is used for administration of computer networks.

LDAP - Lightweight Directory Access Protocol

LDAP is used for collecting information about users and e-mail addresses from the internet.

ICMP - Internet Control Message Protocol

ICMP takes care of error-handling in the network.

ARP - Address Resolution Protocol

ARP is used by IP to find the hardware address of a computer network card based on the IP address.

RARP - Reverse Address Resolution Protocol

RARP is used by IP to find the IP address based on the hardware address of a computer network card.

BOOTP - Boot Protocol

BOOTP is used for booting (starting) computers from the network.

PPTP - Point to Point Tunneling Protocol

PPTP is used for setting up a connection (tunnel) between private networks.

Chapter 2

Wireless Communication Technology

2.1 Frequency for Radio Transmission

2.2 Signal Antennas

2.3 Signal Propagation

2.4 Multiplexing

2.5 Modulation

2.6 Spread Spectrum

2.7 Error Control

2.1 Frequency for Radio Transmission

Radio waves are a type of electromagnetic wave. Since it is a wave, we can make an analogy between radio waves and water waves. When a water waves propagate from a single center point by the action of tossing a rock into the water. Radio waves cause electrons in an antenna to surge back and forth the same way that the ripples in a pond case the water to move up and down. Unlike water waves, radio waves travel at the speed of light (186,000 miles per second).

Just like light waves, radio waves travel in straight lines. Radio waves can reflect off of surfaces, such as layers of ionized gases in the ionosphere. Radio waves reflect off of surfaces the same way that water waves do. The reflection of radio waves is shown in the figure below. Here, radio waves are reflected off of the ionosphere and the Earth. Radio waves, however are not shaped like the circular water waves. Radio waves typically are domed shaped as shown in the picture to the right. These domed-shaped waves radiate from a transmitting antenna in all directions. Waves of different wavelenghts can cross or even travel along the same lines without interfering with the others. This allows several waves to exist at the same time - all at different frequencies. It is the job of the receiver to tune into a specific frequency and ignore all radio signals at other frequencies.

The sound that humans here is caused by vibrations in air. If a sound is loud, the vibration will be large, if the sound is soft, the vibration will be small. This is shown in the figure above. The changing magnitude and the frequency of the wave shown above is what actually carries sound information to the human ear. The human ear can hear sounds in the range from

20Hz to 20,000Hz, where the unit Hz is used to indicate the frequency of the sound wave. A really deep sound will have a low frequency and a really high-pitched sound will have a high frequency. These small frequency waves cannot be transmitted. These audio frequency waves must somehow be changed in order to be transmitted. The two methods of doing this are frequency modulation (FM) and amplitude modulation (AM).

Amplitude modulation was the first radio transmission method. Amplitude modulation works by creating a signal which is has a constant frequency at whatever radio frequency you want to broadcast at (1Mhz for example). The amplitude of the wave is then changed according to the strength of the signal being broadcasted.

Frequency Modulation is the most popular radio transmission technique used today. FM is so popular because it is able to transmit more of the sound that we want to here. AM has problems in transmitting sounds which are at higher frequencies, such as those created by a flute. Frequency modulation works by first creating a signal at the desired carrier frequency (107.7Mhz for example). Then the varying sound level causes frequency changes - The louder the sound, the frequency of the carrier wave increases. Frequency modulation is shown in the figure below. An important factor in FM is that the carrier signal has constant amplitude

2.2 Signal Antennas

The following diagram depicts a typical radio system:

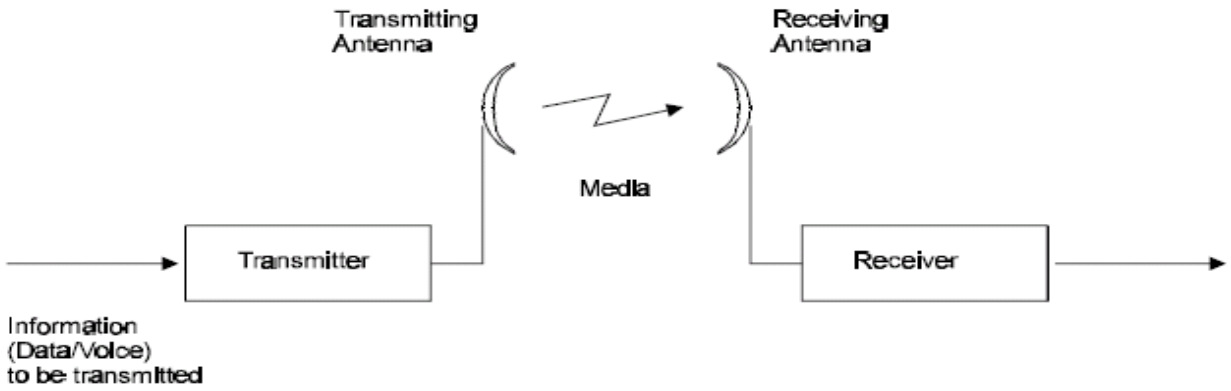


Fig: A Typical Radio System

An antenna (or aerial) is a transducer that transmits or receives electromagnetic waves. In other words, antennas convert electromagnetic radiation into electrical current, or vice versa. Antennas generally deal in the transmission and reception of radio waves, and are a necessary part of all radio equipment. Antennas are used in systems such as radio and television broadcasting, point-to-point radio communication, wireless LAN, cell phones, radar, and spacecraft communication. Antennas are most commonly employed in air or outer space, but can

also be operated under water or even through soil and rock at certain frequencies for short distances.

Physically, an antenna is an arrangement of one or more conductors, usually called elements in this context. In transmission, an alternating current is created in the elements by applying a voltage at the antenna terminals, causing the elements to radiate an electromagnetic field. In reception, the inverse occurs: an electromagnetic field from another source induces an alternating current in the elements and a corresponding voltage at the antenna's terminals. Some receiving antennas (such as parabolic and horn types) incorporate shaped reflective surfaces to collect the radio waves striking them and direct or focus them onto the actual conductive elements.

Antenna Characteristics

Isotropic Antenna

A hypothetical, lossless antenna having equal radiation intensity in all directions. Used as a zero dB gain reference in directivity calculation (gain).

Gain

Antenna gain is a measure of directivity. It is defined as the ratio of the radiation intensity in a given direction to the radiation intensity that would be obtained if the power accepted by the antenna was radiated equally in all directions (isotropically). Antenna gain is expressed in dBi.

Radiation Pattern

The radiation pattern is a graphical representation in either polar or rectangular coordinates of the spatial energy distribution of an antenna.

Side Lobes

The radiation lobes in any direction other than that of the main lobe.

Omni-directional Antenna

This antenna radiates and receives equally in all directions in azimuth. The following diagram shows the radiation pattern of an omni-directional antenna with its side lobes in polar form.

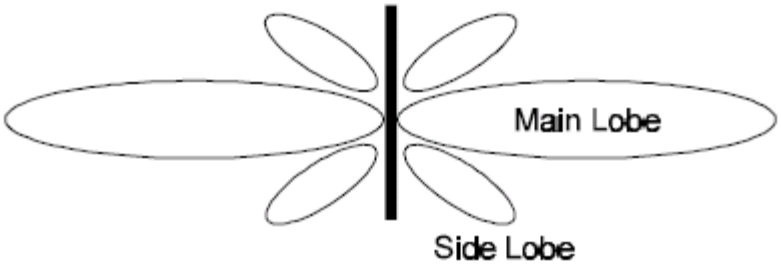


Fig: Side View

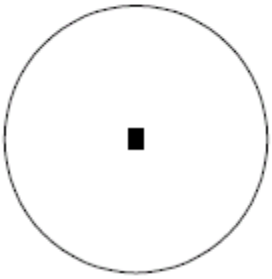


Fig: Top View

Directional Antenna
This antenna radiates and receives most of the signal power in one direction. The following diagram shows the radiation pattern of a directional antenna with its side lobes in polar form:

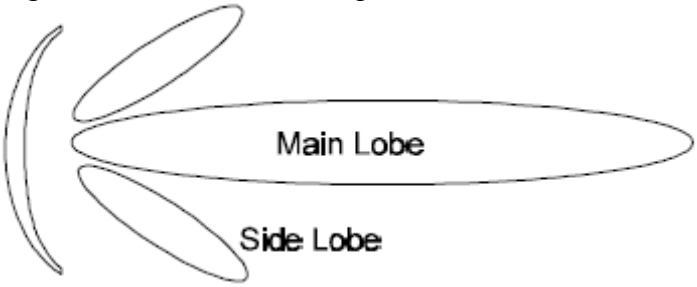


Fig: ***Radiation Pattern of Directional Antenna***

Antenna Beamwidth
The directiveness of a directional antenna. Defined as the angle between two half-power (-3 dB) points on either side of the main lobe of radiation.

2.3 Signal Propagation

Ground Wave Signal Propagation

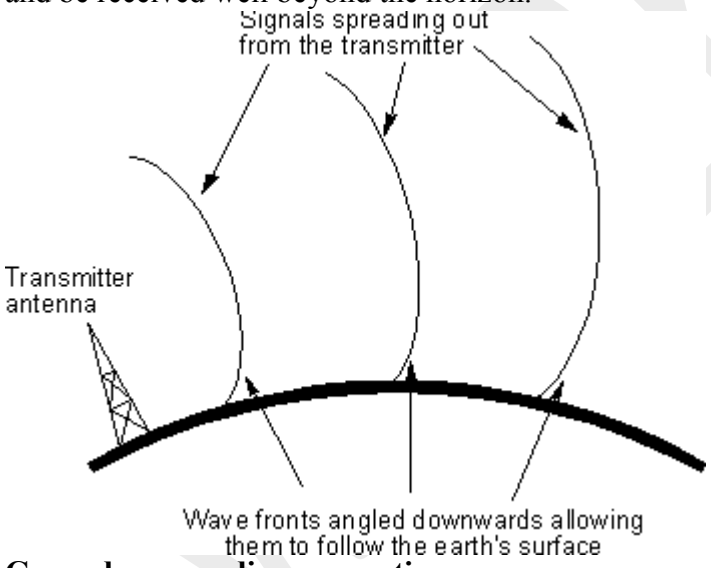
- the ground wave used for radio communications signal propagation on the long, and medium wave bands for local radio communications

Ground wave propagation is particularly important on the LF and MF portion of the radio spectrum. Ground wave radio propagation is used to provide relatively local radio communications coverage, especially by radio broadcast stations that require to cover a particular locality.
Ground wave radio signal propagation is ideal for relatively short distance propagation on these frequencies during the daytime. Sky-wave ionospheric propagation is not possible during the day because of the attenuation of the signals on these frequencies caused by the D region in the ionosphere. In view of this, radio communications stations need to rely on the ground-wave propagation to achieve their coverage.

A ground wave radio signal is made up from a number of constituents. If the antennas are in the line of sight then there will be a direct wave as well as a reflected signal. As the names suggest the direct signal is one that travels directly between the two antenna and is not affected by the locality. There will also be a reflected signal as the transmission will be reflected by a number of objects including the earth's surface and any hills, or large buildings. That may be present. In addition to this there is surface wave. This tends to follow the curvature of the Earth and enables coverage to be achieved beyond the horizon. It is the sum of all these components that is known as the ground wave. Beyond the horizon the direct and reflected waves are blocked by the curvature of the Earth, and the signal is purely made up from the diffracted surface wave. It is for this reason that surface wave is commonly called ground wave propagation.

Surface wave

The radio signal spreads out from the transmitter along the surface of the Earth. Instead of just travelling in a straight line the radio signals tend to follow the curvature of the Earth. This is because currents are induced in the surface of the earth and this action slows down the wave-front in this region, causing the wave-front of the radio communications signal to tilt downwards towards the Earth. With the wave-front tilted in this direction it is able to curve around the Earth and be received well beyond the horizon.



Ground wave radio propagation

Effect of frequency

As the wavefront of the ground wave travels along the Earth's surface it is attenuated. The degree of attenuation is dependent upon a variety of factors. Frequency of the radio signal is one of the major determining factor as losses rise with increasing frequency. As a result it makes this form of propagation impracticable above the bottom end of the HF portion of the spectrum (3 MHz). Typically a signal at 3.0 MHz will suffer an attenuation that may be in the region of 20 to 60 dB more than one at 0.5 MHz dependent upon a variety of factors in the signal path including the

distance. In view of this it can be seen why even high power HF radio broadcast stations may only be audible for a few miles from the transmitting site via the ground wave.

Effect of the ground

The surface wave is also very dependent upon the nature of the ground over which the signal travels. Ground conductivity, terrain roughness and the dielectric constant all affect the signal attenuation. In addition to this the ground penetration varies, becoming greater at lower frequencies, and this means that it is not just the surface conductivity that is of interest. At the higher frequencies this is not of great importance, but at lower frequencies penetration means that ground strata down to 100 metres may have an effect.
Despite all these variables, it is found that terrain with good conductivity gives the best result. Thus soil type and the moisture content are of importance. Salty sea water is the best, and rich agricultural, or marshy land is also good. Dry sandy terrain and city centres are by far the worst. This means sea paths are optimum, although even these are subject to variations due to the roughness of the sea, resulting on path losses being slightly dependent upon the weather! It should also be noted that in view of the fact that signal penetration has an effect, the water table may have an effect dependent upon the frequency in use.

HF Ionospheric Radio Signal Propagation

The basics of HF ionospheric radio propagation and how the ionosphere enables radio communications links to be established over large distances around the globe using what are termed sky waves or skywaves.

As electromagnetic waves, and in this case, radio signals travel, they interact with objects and the media in which they travel. As they do this the radio signals can be reflected, refracted or diffracted. These interactions cause the radio signals to change direction, and to reach areas which would not be possible if the radio signals travelled in a direct line. HF radio communications is dependent for most of its applications on the use of the ionosphere. This region in the atmosphere enables radio communications signals to be reflected, or more correctly refracted back to earth so that they can travel over great distances around the globe. Ionospheric propagation is normally though of as an HF propagation mode, although, it use can extend above and below the HF portion of the spectrum on many occasions. The fact that radio communications signals can travel all over the globe on the HF bands is widely used by many by broadcasters, news agencies, maritime, radio hams and many other users. Radio transmitters using relatively low powers can be used to communicate to the other side of the globe. Although radio propagation using the ionosphere may not be not as reliable as that provided by satellites, it nevertheless provides a very cost effective and efficient form of radio communication. To enable the most to be made of ionospheric propagation many radio users make extensive use of HF propagation programmes to predict the areas of the globe to which signals may travel, or the probability of them reaching a given area.
These HF propagation prediction programmes utilise a large amount of data, and many have been developed over many years, along with data about the prevailing conditions. However it is

still useful to gain a view of how signals travel when using ionospheric propagation and to understand why signal conditions change. In this way the best use can be made of ionospheric propagation.
Radio communications signals in the medium and short wave bands travel by two basic means. The first is known as a ground wave (covered on a separate page in this section), and the second a sky wave using the ionosphere.

Skywaves

When using ionospheric radio propagation, the radio signals leave the Earth's surface and travel towards the ionosphere where some of these are returned to Earth. These radio signals are termed sky waves for obvious reason. If they are returned to Earth, then the ionosphere may (very simply) be viewed as a vast reflecting surface encompassing the Earth that enables signals to travel over much greater distances than would otherwise be possible. Naturally this is a great over simplification because the frequency, time of day and many other parameters govern the reflection, or more correctly the refraction of signals back to Earth. There are in fact a number of layers, or more correctly regions within the ionosphere, and these act in different ways as described below.

D region

When a sky wave leaves the Earth's surface and travels upwards, the first region of interest that it reaches in the ionosphere is called the D region. This region attenuates the signals as they pass through. The level of attenuation depends on the frequency. Low frequencies are attenuated more than higher ones. In fact it is found that the attenuation varies as the inverse square of the frequency, i.e. doubling the frequency reduces the level of attenuation by a factor of four. This means that low frequency signals are often prevented from reaching the higher regions, except at night when the region disappears.
The D region attenuates signals because the radio signals cause the free electrons in the region to vibrate. As they vibrate the electrons collide with molecules, and at each collision there is a small loss of energy. With countless millions of electrons vibrating, the amount of energy loss becomes noticeable and manifests itself as a reduction in the overall signal level. The amount of signal loss is dependent upon a number of factors: One is the number of gas molecules that are present. The greater the number of gas molecules, the higher the number of collisions and hence the higher the attenuation. The level of ionisation is also very important. The higher the level of ionisation, the greater the number of electrons that vibrate and collide with molecules. The third main factor is the frequency of the signal. As the frequency increases, the wavelength of the vibration shortens, and the number of collisions between the free electrons and gas molecules decreases. As a result signals lower in the radio frequency spectrum are attenuated far more than

those which are higher in frequency. Even so high frequency signals still suffer some reduction in signal strength.

E and F Regions

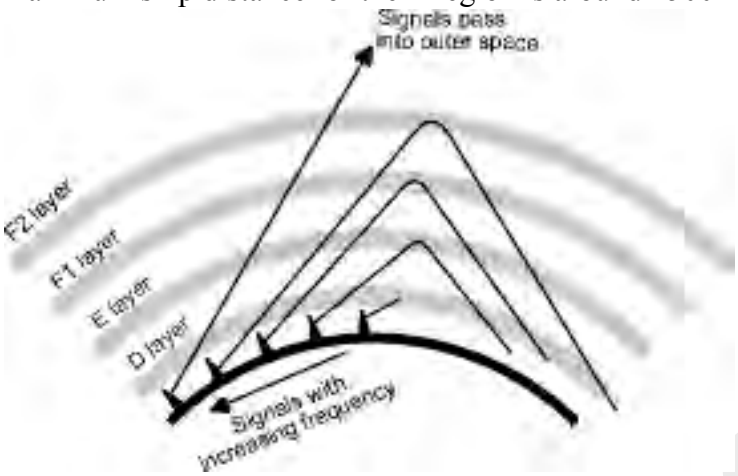
Once a signal passes through the D region, it travels on and reaches first the E, and next the F regions. At the altitude where these regions are found the air density is very much less, and this means that when the free electrons are excited by radio signals and vibrate, far fewer collisions occur. As a result the way in which these regions act is somewhat different. The electrons are again set in motion by the radio signal, but they tend to re-radiate it. As the signal is travelling in an area where the density of electrons is increasing, the further it progresses into the region, the signal is refracted away from the area of higher electron density. In the case of HF signals, this refraction is often sufficient to bend them back to earth. In effect it appears that the region has "reflected" the signal.
The tendency for this "reflection" is dependent upon the frequency and the angle of incidence. As the frequency increases, it is found that the amount of refraction decreases until a frequency is reached where the signals pass through the region and on to the next. Eventually a point is reached where the signal passes through all the regions and on into outer space.

Refraction of a radio signal as it enters an ionised region

Different frequencies

To gain a better idea of the characteristics of HF propagation using the ionosphere, it is worth viewing what happens to a radio communications signal if the frequency is increased across the frequency spectrum. First it starts with a signal in the medium wave broadcast band. During the day signals on these frequencies only propagate using the ground wave. Any signals that reach the D region are absorbed. However at night as the D region disappears signals reach the other regions and may be heard over much greater distances.
If the frequency of the signal is increased, a point is reached where the signal starts to penetrate the D region and signals reach the E region. Here it is reflected and will pass back through the D region and return to earth a considerable distance away from the transmitter.
As the frequency is increased further the signal is refracted less and less by the E region and eventually it passes right through. It then reaches the F1 region and here it may be reflected passing back through the D and E regions to reach the earth again. As the F1 region is higher than the E region the distance reached will be greater than that for an E region reflection.

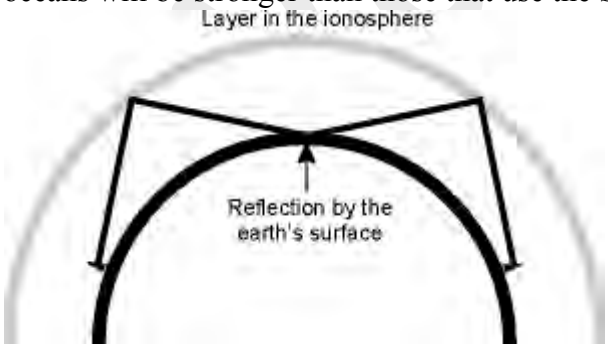
Finally as the frequency of the radio communications signal rises still further the it will eventually pass through the F1 region and onto the F2 region. This is the highest of the regions in the ionosphere and the distances reached using this are the greatest. As a rough guide the maximum skip distance for the E region is around 2500 km and 5000 km for the F2 region.



Signals reflected by the E and F regions

Multiple hops

Whilst it is possible to reach considerable distances using the F region as already described, on its own this does not explain the fact that radio signals are regularly heard from opposite sides of the globe using HF propagation with the ionosphere. This occurs because the signals are able to undergo several "reflections". Once the signals are returned to earth from the ionosphere, they are reflected back upwards by the earth's surface, and again they are able to undergo another "reflection" by the ionosphere. Naturally the signal is reduced in strength at each "reflection", and it is also found that different areas of the Earth reflect radio signals differently. As might be anticipated the surface of the sea is a very good reflector, whereas desert areas are very poor. This means that signals that are "reflected" back to the ionosphere by the Pacific or Atlantic oceans will be stronger than those that use the Sahara desert or the red centre of Australia.



Multiple reflections

It is not just the Earth's surface that introduces losses into the signal path. In fact the major cause of loss is the D region, even for frequencies high up into the HF portion of the spectrum. One of the reasons for this is that the signal has to pass through the D region twice for every reflection by the ionosphere. This means that to get the best signal strengths it is necessary signal paths

enable the minimum number of hops to be used. This is generally achieved using frequencies close to the maximum frequencies that can support communications using ionospheric propagation, and thereby using the highest regions in the ionosphere. In addition to this the level of attenuation introduced by the D region is also reduced. This means that a radio signal on 20 MHz for example will be stronger than one on 10 MHz if propagation can be supported at both frequencies.

HF propagation summary

HF propagation using the ionosphere is still a widely used as a form of radio communications. While not as reliable as satellite communications, it is not nearly as expensive, and can provide a useful back-up in case the satellite communications fail. It is also widely used as the primary form of radio communications by many organisations from radio broadcasters to radio amateurs, as well as ship to shore and many other forms of point to point communications. As a result HF propagation using the ionosphere is likely to remain in use indefinitely as a form of radio communications technology.

Line of Sight

Line-of-sight propagation refers to electro-magnetic radiation or acoustic wave propagation. Electromagnetic transmission includes light emissions traveling in a straight line. The rays or waves may be diffracted, refracted, reflected, or absorbed by atmosphere and obstructions with material and generally cannot travel over the horizon or behind obstacles.

Especially radio signals, like all electromagnetic radiation including light emissions, travel in straight lines. At low frequencies (below approximately 2 MHz or so) these signals travel as ground waves, which follow the Earth's curvature due to diffraction with the layers of atmosphere. This enables AM radio signals in low-noise environments to be received well after the transmitting antenna has dropped below the horizon. Additionally, frequencies between approximately 1 and 30 MHz, can be reflected by the F1/F2 Layer, thus giving radio transmissions in this range a potentially global reach (see shortwave radio), again along multiply deflected straight lines. The effects of multiple diffraction or reflection lead to macroscopically "quasi-curved paths".

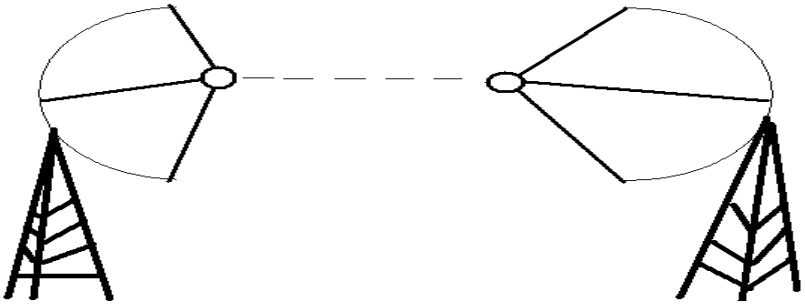


Fig: line of Sight

However, at higher frequencies and in lower levels of the atmosphere, neither of these effects apply. Thus any obstruction between the transmitting antenna and the receiving antenna will block the signal, just like the light that the eye may sense. Therefore, since the ability to visually see a transmitting antenna (disregarding the limitations of the eye's resolution) roughly corresponds to the ability to receive a radio signal from it, the propagation characteristic of high-frequency radio is called "line-of-sight". The farthest possible point of propagation is referred to as the "radio horizon".

In practice, the propagation characteristics of these radio waves vary substantially depending on the exact frequency and the strength of the transmitted signal (a function of both the transmitter and the antenna characteristics). Broadcast FM radio, at comparatively low frequencies of around 100 MHz, easily propagates through buildings and forests.

2.3 Multiplexing

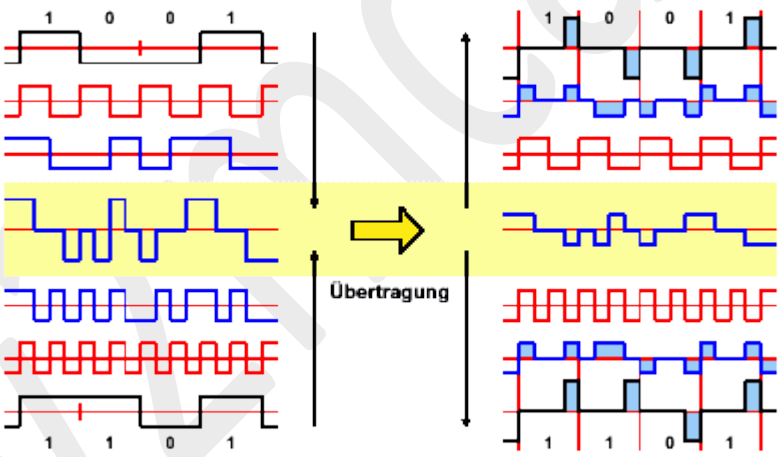
The following page gives a rough overview about several multiplexing techniques. These describe how several independent data channels have access to a single physical signal carrier medium like copper cable, fiber optical cable or the air for wireless transmission systems. In general the multiplexing schemes are based upon time, frequency and code. Depending upon the scheme multiplexing is done before and/or after the channel coding and/or modulation.

CDMA Code Division Multiplex Access
COFDM Coded Orthogonal Frequency Division Multiplexing
FDMA Frequency Division Multiplex Access
OFDM Orthogonal Frequency Division Multiplexing
TDMA Time Division Multiplex Access
WCDMA Wide Band Code Division Multiple Access
SDMA Spatial Division Multiplex Access

CDMA Code Division Multiplex Access (CDMA)

CDMA refers to any of several protocols used in so-called second-generation (2G) and third-generation (3G) wireless communications. CDMA is a form of multiplexing, allowing numerous signals to use a single transmission channel, optimizing the use of available bandwidth. The technology is used in ultra-high-frequency (UHF) cellular telephone systems in the 800-MHz and 1.9-GHz bands.

CDMA employs analog-to-digital conversion (ADC) in combination with spread spectrum technology. Audio input is first digitized into binary elements. The frequency of the transmitted signal is then made to vary according to a defined pattern (code), so it can be intercepted only by a receiver whose frequency response is programmed with the same code. The CDMA channel is nominally 1.23 MHz wide. CDMA networks use a scheme called soft hand-off, which minimizes signal breakup as a handset passes from one cell to another. The combination of digital and spread-spectrum modes supports several times as many signals per unit bandwidth as analog modes. CDMA is compatible with other cellular technologies which allows for nationwide Roaming.



The original CDMA standard, also known as CDMA One and still common in cellular telephones in the US, offers a transmission speed of only up to 14.4 Kbps in its single channel form and up to 115 Kbps in an eight-channel form. CDMA2000 and Wideband CDMA (WCDMA) deliver data many times faster.

COFDM (Coded OFDM, Coded Orthogonal Frequency Division Multiplexing)

COFDM is an expansion of the already available OFDM modulation technique. The special performance of COFDM with respect to multipath and interference, burst errors and fading are the reasons why COFDM is well-suited to the needs of terrestrial broadcasting channels. COFDM is resistant to multipath effects because it uses multiple carriers to transmit the same signal. Thus COFDM has been choosen for the two standards DAB (Digital Audio Broadcast) and DVB-T (Digital Video Broadcast-Terrestrial). COFDM is ideal for single frequency networks.

FDMA Frequency Division Multiplex Access

FDMA is the division of the frequency band allocated for wireless cellular telephone communication into 30 channels, each of which can carry a voice conversation or, carry data of a

digital service. FDMA is a basic technology in the analog Advanced Mobile Phone Service (AMPS), the most widely-installed cellular phone system installed in North America. With FDMA, each channel can be assigned to only one user at a time. The Digital-Advanced Mobile Phone Service (D-AMPS) also uses FDMA but adds time division multiple access (TDMA) to get three channels for each FDMA channel, tripling the number of calls that can be handled on a channel.

OFDM Orthogonal Frequency Division Multiplexing

OFDM is fundamentally different from other modulation schemes because it may be transmitted via AM, FM, QAM (Quadrature Amplitude Modulation), and so on. OFDM is defined as a mathematically technique for the generation and demodulation of radio waves.

Frequency division multiplexing (FDM) is a technology that transmits multiple signals simultaneously over a single transmission path, such as a cable or wireless system. Each signal travels within its own unique frequency range (carrier), which is modulated by the data (text, voice, video, etc.). Orthogonal FDM's (OFDM) spread spectrum technique distributes the data over a large number of carriers that are spaced apart at precise frequencies. This spacing provides the "orthogonality" in this technique which prevents the demodulators from seeing frequencies other than their own. The benefits of OFDM are high spectral efficiency, resiliency to RF interference, and lower multi-path distortion. This is useful because in a typical terrestrial broadcasting scenario there are multipath-channels (i.e. the transmitted signal arrives at the receiver using various paths of different length).

The multicarrier transmission techniques OFDM is seen as a key technology for high-rate communications and is already part of the IEEE 801.11 and ETSI BRAN standard for wireless local area networks. OFDM became a serious alternative by applying modern, digital signal processing methods based on the Fast Fourier Transform (FFT). The main advantages are the high spectral efficiency, the simple channel equalization and the suppression of intersymbol interference. However OFDM suffers from a high peak-to-average power ratio of the transmitting signal. The unfavorable PAR prohibits power efficient operation of the amplifier which is especially intolerable in portable systems. Thus OFDM is not confined to mobile communications but is used in Digital Audio Broadcast (DAB), Digital Video Broadcast (DVB) and xDSL systems.

PAR === (crest factor, feed forward (digital) predistortion)

TDMA Time Division Multiplex Access

TDMA is a technology used in digital cellular telephone communication that divides each cellular channel into time slots in order to increase the amount of data that can be carried. TDMA is used by Digital-American Mobile Phone Service (D-AMPS), Global System for Mobile communications (GSM), and Personal Digital Cellular (PDC). However, each of these systems implements TDMA in a somewhat different and incompatible way. An alternative multiplexing scheme to FDMA with TDMA is CDMA, which takes the entire allocated frequency range for a given service and multiplexes information for all users across the spectrum

range at the same time.
TDMA was first specified as a standard in EIA/TIA Interim Standard 54 (IS-54). IS-136, an evolved version of IS-54, is the United States standard for TDMA for both the cellular (850 MHz) and personal communications services (1.9 GHz) spectrums. TDMA is also used for Digital Enhanced Cordless Telecommunications (DECT).

WCDMA Wide Band Code Division Multiple Access

WCDMA is an ITU standard and was derived from the CDMA standard. WCDMA is a third-generation (3G) mobile wireless technology offering much higher data speeds to mobile and portable wireless devices than older technologies. WCDMA can support mobile/portable voice, images, data, and video communications at up to 2 Mbps (local area access, no movement) or 384 Kbps (wide area access, movement in trains, cars, ...). The input signals are digitized and transmitted in coded, spread-spectrum mode over a broad range of frequencies. A 5 MHz-wide carrier is used, compared with 200 kHz-wide carrier for narrowband CDMA.

SDMA Spatial Division Multiplex Access

SDMA has proven to be an interesting option for capacity increase of wireless communications systems. The idea is to allow several users to use the same frequency band (and time slot) simultaneously and to identify them from their positions. SDMA makes use of antenna-array processing and advanced digital signal processing techniques which rely heavily on concepts from linear algebra. The matrix decompositions involved are complex, mostly O^{n3}, with n the problem size. Moreover, data have to be processed at a high rate (e.g. GSM 270 kbits/s), so that the computational requirement is in the Mflops/s or even Gflops/s range. Hence there is a strong need for efficient algorithms which can be obtained by using adaptive matrix decomposition techniques.

2.5 Modulation

The following page gives a rough overview about several modulation techniques. The basics of modulation is to take a message bearing signal like an audio signal and superimpose it upon a carrier signal for transmission. For ease of transmission such carrier signals use generally high frequencies:

- For easy propagation as electromagnetic waves with low loss and low dispersion
- Simultaneous transmission without interference from other signals
- Enables the construction of small antennas (a fraction, usually a quarter of the wavelength)
- Enables the multiplexing (combining) multiple signals for transmission at the same time over the same carrier

In general the four different modulation schemes amplitude, frequency, phase and code modulation are possible. Well known examples of high frequency carrier signals are:

- AM radio is 550-1600 KHz
- FM radio is 88 MHz-108 MHz
- TV is 52-88 MHz (channels 1-6), 174-216 MHz (channels 7-12) and 470-900 MHz (UHF)
- microwave and satellite signals are of the order of several GHz
- infra red fiber optic signals are of the order of 200-300 THz.

The selection of the modulation scheme is done according to power and bandwidth efficiency and the channel capacity. The power efficiency is a measure of how much signal power should be increased to achieve a particular BER (Bit Error Rate) for a given modulation scheme. Bandwidth efficiency is the ability to accommodate data within a limited bandwidth of a channel. Its also a tradeoff between data rate and pulse width.

The classical analog modulation schemes are based on the continuous change of sinusodial signal carrier. Here the modulation can plays with one of 3 parameters: Change in frequency, amplitude or signal phase. Modern digital modulation schemes have already been developed and will replace the classical more and more.

Pulse Modulation Schemes
ADPCM Adaptive Differential Pulse Code Modulation
PAM Pulse Amplitude Modulation
PCM Pulse Code Modulation
PDM Pulse Duration Modulation
PFM Pulse Frequency Modulation
PPM Pulse Position Modulation
PSK Phase Shift Keying
PTM Pulse Time Modulation
PWM Pulse Width Modulation
ASK Amplitude Shift Keying
BPSK Binary Phase Shift Keying
FSK Frequency Shift Keying
QPSK Quadrature Phase Shift Keying
QAM Quadrature Amplitude Modulation

Other specified Quadrature Modulation (QM) schemes are:

- QAM Quantized Amplitude Modulation
- QFM Quantized Frequency Modulation
- QPAM Quadrature Phase and Amplitude Modulation
- QPAM Quantized Pulse Amplitude Modulation
- QPM Quantized Phase Modulation
- QPPM Quantized Pulse Position Modulation
- QSAM Quadrature Sideband Amplitude Modulation
- QUAM Quantized Amplitude Modulation

ADPCM Adaptive Differential Pulse Code Modulation; G.726

ADPCM is a technique defined by the International Telecommunication Union (ITU) for converting sound or analog information to binary information by taking frequent samples of the sound and expressing the value of the sampled sound modulation in binary terms. This produces a lower bit rate and is sometimes used to effectively compress a voice signal, allowing both voice and digital data to be sent where only one would normally be sent. ADPCM is used to send sound on fiber-optic long-distance lines as well as to store sound along with text, images, and code on a CD-ROM. It is also used in digital cordless telephones, radio/wireless local loop and pair-gain.
ADPCM is a variation of pulse code modulation (PCM) that only sends the difference between two adjacent samples. This produces a lower bit rate and is used to effectively compress a voice signal, allowing more voice channels or both, voice and digital data, to be sent within a 32k-kbit/s digital channel. 3 or 4 bits are used to describe each sample, which represents the difference between two adjacent samples. Sample rate is 8000 samples/second.
G.726 provides an outline description of the ADPCM transcoding algorithm including ADPCM encoding and decoding algorithms respectively. The coding scale can be dynamically changed to compensate for amplitude and frequency variations.

Pulse Modulation Schemes

Pulse modulation schemes incorporate the basic idea to use the carrier signal as a pulse train. Different choices of pulse formats are possible and have different aspects in terms of energy and spectral content consumption. Examples of pulse formats are square pulses, raised cosine pulses or sync function (Nyquist) pulses. The characteristics of the pulse train that can be varied are:

- Pulse Amplitude
- Pulse Width
- Position of leading edge

Modulation schemes varying amplitude and width are called Pulse Amplitude Modulation (PAM) and Pulse Width Modulation (PWM) respectively.
Note that this basic scheme can be made more sophisticated by using several amplitude levels, for example grouping signal bits into groups of 2, i.e. 00, 01, 10 and 11 and have four different amplitude levels for each of these groups. This scheme is known as Quadrature Pulse Amplitude Modulation (QPAM or QAM).

PAM Pulse Amplitude Modulation

Pulse Amplitude Modulation refers to a method of carrying information on a train of pulses, the information being encoded in the amplitude of the pulses.

PCM Pulse Code Modulation

PCM is a general scheme for transmitting analog data in a digital and binary way, independent of the complexity of the analog waveform. With PCM all forms of analog data like video, voice, music and telemetry can be transferred.
To obtain PCM from an analog waveform at the source (transmitter), the analog signal amplitude is sampled at regular time intervals. The sampling rate (number of samples per second), is several times the maximum frequency of the analog waveform. The amplitude of the analog signal at each sample is rounded off to the nearest binary level (quantization). The number of levels is always a power of 2 (4, 8, 16, 32, 64, ...). These numbers can be represented by two, three, four, five, six or more binary digits (bits) respectively.
At the destination (receiver), a pulse code demodulator converts the binary numbers back into pulses having the same quantum levels as those in the modulator. These pulses are further processed to restore the original analog waveform.

PDM Pulse Duration Modulation

PDM is a method of pulse modulation in which the duration of a the pulse train is used to transfer the binary signal information.

PFM Pulse Frequency Modulation

PFM is a method of pulse modulation in which the modulating wave is used to frequency modulate a pulse-generating circuit. For example, the pulse rate may be 8000 pulses per second (pps) when the signal voltage is 0. The pulse rate may step up to 9000 pps for maximum positive signal voltage, and down to 7000 pps for maximum negative signal voltage. This method of modulation is not used extensively because of complicated PFM generation circuitry. It requires a stable oscillator that is frequency modulated to drive a pulse generator.

PPM Pulse Position Modulation

The amplitude and width of the pulse is kept constant in the system. The position of each pulse, in relation to the position of a recurrent reference pulse, is varied by each instantaneous sampled

value of the modulating wave. PPM has the advantage of requiring constant transmitter power since the pulses are of constant amplitude and duration. It is widely used but has the big disadvantage that it needs a synchronization between transmitter and receiver.

PTM Pulse Time Modulation

The Time characteristics of pulses may also be modulated with intelligence information. Two time characteristics may be affected, the time duration of the pulses and the occurrence of the pulses. Thus PTM can be differentiated into the following 4 modulation sub-schemes:

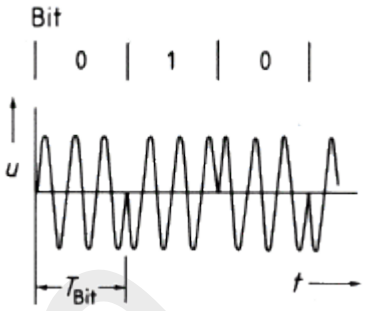
- PDM Pulse Duration Modulation
- PPM Pulse Position Modulation
- PFM Pulse Frequency Modulation
- PWM Pulse Width Modulation

PWM Pulse Width Modulation

Pulse Width Modulation refers to a method of carrying information on a train of pulses, the information being encoded in the width of the pulses.
In applications to motion control, it is not exactly information we are encoding, but a method of controlling power in motors without (significant) loss.
There are several schemes to accomplish this technique. One is to switch voltage on and off, and let the current recirculate through diodes when the transistors have switched off. Another technique is to switch voltage polarity back and forth with a full-bridge switch arrangement, with 4 transistors. This technique may have better linearity, since it can go right down to an effective 0% duty cycle by having the positive and negative voltage periods precisely equal. On/Off techniques may have trouble going down extremely close to 0% duty cycles, and may jitter between minimum duty cycles of positive and negative polarity.
In battery systems PWM is the most effective way to achieve a constant voltage for battery charging by switching the system controller's power devices on and off.
The generation of exact working PWM circuitry is complicated, but it is extremely conceptually important since there is good reason to believe that neurons transmit information using PWM spike trains.

ASK Amplitude Shift Keying

ASK is a scheme according to the old telegraph key. The transmission of the binary bit stream is done by switching the carrier signal on/off according to 1/0 bits.

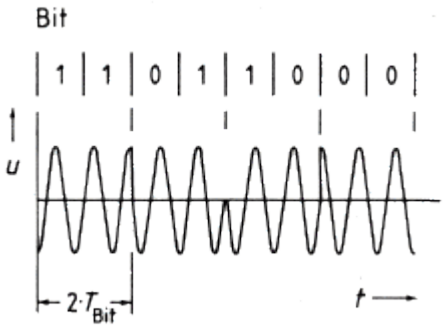


BPSK Binary Phase Shift Keying

BPSK shifts the phase of the carrier signal by 0 or by 180 degrees depending upon the bits to be transferred. This scheme is simple to implement, but it's inefficient in terms of using the available bandwidth. It is very robust, which is the reason that BPSK is extensively used in satellite communication systems.

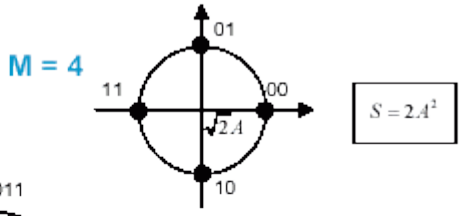
FSK Frequency Shift Keying

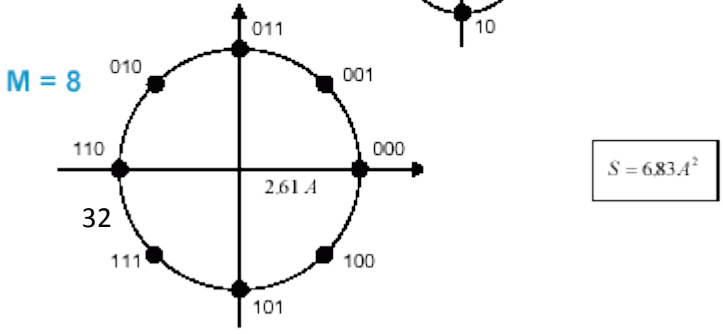
FSK is the frequency modulation of a carrier. For Simplex or Half Duplex operation, a single carrier (1170 Hz) is used - communication can only be transmitted in one direction at a time. A Mark or 1 is represented by 1270 Hz, and a Space or 0 is represented by 1070 Hz.
For Full Duplex (simultaneous data communication in both directions), the upper bandwidth of the Voice Channel is utilized. Another carrier is added at 2125 Hz. A Mark or 1 is represented by 2225 Hz, and a Space or 0 is represented by 2025 Hz. The originating modem uses the lower carrier (1170 Hz) and the answer modem uses the upper carrier (2125 Hz).



QPSK Quadrature Phase Shift Keying

Quadrature Phase Shift Keying employs shifting the phase of the carrier at a 600 baud rate plus an encoding technique. QPSK is used in Bell 212A-compatible modems and V.22 - both are 1200 bps Full Duplex standards. The originate modem transmits at 1200 Hz, and receives on 2400 Hz. The answer modem receives on 1200 Hz, and transmits on 2400 Hz. The digital information is encoded using 4 (Quad) level differential PSK at 600 baud, 00 for 90 degrees, 01 for 0 degrees, 10 for 180 degrees and 11 for 270 degrees.
The scheme of QPSK can be extended from 2-bit (M=4) symbols to 3-bit





symbols (M=8) or even more bits (m-QPSK).

QAM Quadrature Amplitude Modulation

QAM is a method of combining two amplitude-modulated (AM) signals into a single channel, thereby doubling the effective bandwidth. QAM is used with pulse amplitude modulation (PAM) in digital systems, especially in wireless applications.
In a QAM signal, there are two carriers, each using the same frequency band but differing in phase by 90 degrees (which is one quarter of a cycle, from which the term quadrature arises). One signal is called the *I signal*, and the other is called the *Q signal*. Mathematically the signals can be represented by a sine and a cosine wave. The two modulated carriers are combined at the source for transmission. At the destination, the carriers are separated, the data is extracted from each, and then the data is combined into the original modulating information. Because the orthogonal carriers occupy the same frequency band and differ by a 90 degree phase shift, each can be modulated independently, transmitted over the same frequency band, and separated by demodulation at the receiver. For a given available bandwidth, QAM enables data transmission at twice the rate of standard pulse amplitude modulation (PAM) without any degradation in the bit error rate (BER). QAM and its derivatives are used in both mobile radio and satellite communication systems.
QAM is actually is used for alphabets of size other than 4. 2400 Baud full duplex modems use 16-QAM (corresponding to grouping 4 bits together). 9600 Baud, 14,400 Baud, 28,800 Baud modems use 32-QAM, 128-QAM and 1024-QAM respectively (along with trellis coding scheme).

2.6 Spread Spectrum

In general, signal transmission is enabled through some means of modulation. In the past, systems have relied primarily on narrow-band modulation schemes. In these systems, all of the power in a transmitted signal is confined to a very narrow portion of the frequency bandwidth. As a result of these narrow frequencies, an interfering frequency at or near the transmitting frequency can cause interference, which render the signal unrecoverable. Amplitude Modulation (AM) is one example of a narrow-band modulation scheme in which the amplitude of the carrier signal is made stronger or weaker based on the information in the signal to be transmitted. The large amounts of power that are associated with Amplitude Modulation allow the signal to travel large distances before it attenuates to an undetectable level. A second popular form of modulation is Frequency Modulation (FM), in which the phase of the carrier frequency is adjusted in accordance with the signal being transmitted. Narrow-band modulation schemes are

not the only implementations available to broadcasters. Broadcasting entities may take advantage of the fact that a defined spectral power density may be achieved not only through high power over a very narrow frequency range, but also through lower powers spread over much larger frequency ranges

Spread spectrum is a class of modulation techniques developed over the past 50 years. In order to qualify as a spread spectrum signal, the following criteria must be met:

1. The transmitted signal bandwidth is greater than the minimal information bandwidth needed to successfully transmit the signal.
2. Some function other than the information itself if being employed to determine the resultant transmitted bandwidth.

Most commercial spread spectrum systems transmit an RF signal bandwidth in the neighborhood of one to two orders of magnitude greater than the bandwidth of the information that is being sent. Transmitted bandwidth can be as large as three orders of magnitude above the bandwidth of the information. There are a number of benefits that are obtained from spreading the transmitted signal bandwidth. First, because the spread spectrum signal is being spread over a large bandwidth, it can coexist with narrow-band signals with only a slight increase to the noise floor in a give slice of spectrum. This coexistence is possible because the spread-spectrum receiver is "looking" over such a large range of frequencies that it does not see the narrow-band frequency. Even if the spread-spectrum receiver does detect the narrow band signal, it does not recognize the signal because it is not being transmitted with the proper code sequence. There are a number of incarnations of spread spectrum modulations. We will concentrate our attention on two popular forms of spread spectrum modulation, Direct Sequence and Frequency Hopping, making note that a third hybrid form of the two presented here does exist in practice.

Direct Sequence is one of the most popular forms of spread spectrum. This is probably a result of the simplicity with which direct sequencing can be implemented. In this form of modulation, a pseudo-random noise generator creates a high-speed pseudo-noise code sequence. This sequence is transmitted at a maximum bit rate called the chip rate. The pseudo-random code sequence is used to directly modulate the narrow-band carrier signal; thus, it directly sets the transmitted radio frequency (RF) bandwidth. The chip rate has a direct correlation to the spread of the information. The information is demodulated at the receiving end by multiplying the signal by a locally generated version of the pseudo-random code sequence. While direct sequence is a very popular form of spread spectrum transmission, it is not by any means the only method available. Another popular from of implementing spread spectrum takes an entirely different approach to spreading then that of direct sequencing.

Frequency Hopping is a from of spread spectrum in which spreading takes place by hopping from frequency to frequency over a wide band. The specific order in which the hopping occurs is

determined by a hopping table generated with the help of a pseudo-random code sequence. The rate of hopping is a function of the information rate. The order of frequencies that is selected by the receiver is dictated by the pseudo-random noise sequence. While the transmitted spectrum of a frequency-hopping signal is quite different from that of a direct sequence signal, it is sufficient to note that the data is spread out over a signal band larger than is necessary to carry it. In both cases, the resultant signal appears noise-like and the receiver utilizes a similar technique to the one employed in transmitting in order to recover the original signal.

There are many advantages to using spread spectrum. Since spread-spectrum receivers can effectively ignore narrow-band transmissions, it is possible to share the same frequency band with other users. These users can weather a significant degree of overlap without interference effects. In both mechanisms discussed above, a pseudo-random noise sequence was employed—either to directly modulate the signal or to determine the order of frequencies in the hopping table. Since this pseudo-random signal makes the transmitted signal appear as noise, only receivers possessing the proper duplicate pseudo-random noise code sequence will be able to recover the signal. This fact has great implications for ensuring the privacy of point-to-point communications (or point to multi-point communications, as the case may be). In fact, the US military has for some years used the fact that the noise-like character of the transmitted signal drastically reduces the probability of signal detection and interception to ensure secure communications. The secure communications in and of itself is not sufficiently interesting as strong encryption and spoofing countermeasures can be added (perhaps at great cost) to existing narrow-band communications. The property of interest in spread spectrum transmission is the scheme's ability to provide point-to-point communications without explicit coordination of the speakers. A crude analogy can be made to the CB radios that truckers often employ: the speaker keeps switching the channel until a free spot is open. Spread spectrum's more sophisticated hopping sequence spreads the speaker's message over various channels at different points in time (in one incarnation of the system). This pseudo-random hopping behavior unseats the long-held assumption that signals from two or more speakers may not overlap in time and space in order for communication to occur. To the contrary, all spread-spectrum systems have a threshold or tolerance level below which useful communication continues unimpeded. The question that remains is coordination of users in a multiple access regime.

2.7 Error Control

Error Detection Probability:

- P_b : Probability of single bit error (BER)
- P_1 : Probability that a frame arrives with no bit errors
- P_2 : While using error detection, the probability that a frame arrives with one or more undetected errors

- P_3 : While using error detection, the probability that a frame arrives with one or more detected bit errors but no undetected bit errors

With no error detection

$$P_1 = (1 - P_b)^F$$
$$P_2 = 1 - P_1$$
$$P_3 = 0$$

Where F = Number of bits per frame

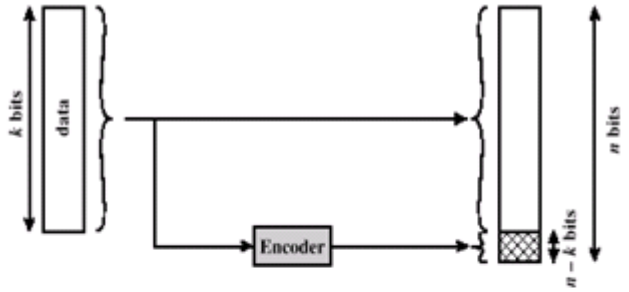
The process of Error Detection can be described in terms of transmitter and receiver

Transmitter

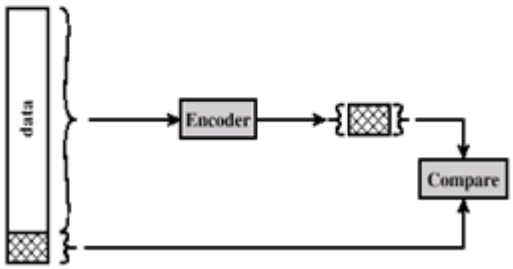
- For a given frame, an error-detecting code (check bits) is calculated from data bits
- Check bits are appended to data bits

Receiver

- Separates incoming frame into data bits and check bits
- Calculates check bits from received data bits
- Compares calculated check bits against received check bits
- Detected error occurs if mismatch



(a) Sender



(b) Receiver

Methods of detecting Error:

Parity Check:

A parity bit is a bit that is appended to a bit stream to make the total number of "1" bits in a given set of bits always even or odd. Parity checks are the simplest error detecting system.

For example, if we send some specific sequence of ones and zeros, and then count the number of ones that we sent and send an extra 1 if that count is odd (making the total now even) or an extra 0 if that count is already even, a single bit error can be detected: The reciever can count up the number of 1 bits they recieved, perform the same calculation, and if the result is not even, they will know that an error occured.

Parity checks have serious limitations: The reciever has no way of knowing which bit was wrong, and if two bits were changed in value (e.g. a pair of 1's are both changed to 0's) then the errors would pass undetected.

There are two types of parity: even and odd. An even parity bit is set to 1 if the number of ones in a given set of bits is odd (making the total number of ones, including the parity bit, even). An odd parity bit is set to 1 if the number of ones in a given set of bits is even (making the total number of ones, including the parity bit, odd).

Even parity is actually a special case of a cyclic redundancy check (CRC), where the 1-bit CRC is generated by the polynomial x+1.

Parity calculated vertically, along the columns of many sets of bits, is called an LRC or Longitudinal Redundancy Check

The job of summing up the ones in a stream of bits is commonly performed by the XOR logic function. In this case, the XOR can be thought of as counting the ones, although what it really does is produce a 1, when it's inputs are different and a 0 when it's inputs are the same. It turns out the result is the same. For example, If we take the bits 1 0 1 1 and add them up, we get 3, which is odd, so even parity would be 1. If we take the first bit and XOR it with the second, we get 1, because they are different. Then we take that result, 1, and XOR it with the next bit of data, also a 1 to get a 0, because they are the same. Finally, the 0 and the last 1 are different and so we end up with a 1. You can try that with any example you like, the end result is the same: XOR is even parity.

CRC:

A CRC is an error-detecting code. Its computation resembles a polynomial long division operation in which the quotient is discarded and the remainder becomes the result, with the important distinction that the polynomial coefficients are calculated according to the carry-less arithmetic of a finite field. The length of the remainder is always less than the length of the divisor (called the **generator polynomial**), which therefore determines how long the result can be. The definition of a particular CRC specifies the divisor to be used, among other things.

Although CRCs can be constructed using any finite field, all commonly used CRCs employ the finite field GF. This is the field of two elements, usually called 0 and 1, comfortably matching computer architecture. The rest of this article will discuss only these binary CRCs, but the principles are more general.

An important reason for the popularity of CRCs for detecting the accidental alteration of data is their efficiency guarantee. Typically, an n-bit CRC, applied to a data block of arbitrary length, will detect any single error burst not longer than n bits (in other words, any single alteration that spans no more than n bits of the data), and will detect a fraction $1-2^{-n}$ of all longer error bursts. Errors in both data transmission channels and magnetic storage media tend to be distributed non-randomly (i.e. are "bursty"), making CRCs' properties more useful than alternative schemes such as multiple parity checks.

The simplest error-detection system, the parity bit, is in fact a trivial 1-bit CRC: it uses the generator polynomial $x+1$.

To compute an n-bit binary CRC, line the bits representing the input in a row, and position the (n+1)-bit pattern representing the CRC's divisor (called a "polynomial") underneath the left-hand end of the row. Here is the first calculation for computing a 3-bit CRC:

```
11010011101100 <--- input
1011           <--- divisor (4 bits)
--------------
01100011101100 <--- result
```

If the input bit above the leftmost divisor bit is 0, do nothing and move the divisor to the right by one bit. If the input bit above the leftmost divisor bit is 1, the divisor is XORed into the input (in other words, the input bit above each 1-bit in the divisor is toggled). The divisor is then shifted one bit to the right, and the process is repeated until the divisor reaches the right-hand end of the input row. Here is the last calculation:

```
00000000001110 <--- result of previous step
         1011 <--- divisor
--------------
00000000000101 <--- remainder (3 bits)
```

Since the leftmost divisor bit zeroed every input bit it touched, when this process ends the only bits in the input row that can be nonzero are the n bits at the right-hand end of the row. These n bits are the remainder of the division step, and will also be the value of the CRC function (unless the chosen CRC specification calls for some post processing).

CRC using Modulo 2 Arithmetic

The division is done using Exclusive-OR operation.

- Parameters:
  - T = n-bit frame to be transmitted
  - D = k-bit block of data; the first k bits of T
  - F = $(n – k)$-bit FCS; the last $(n – k)$ bits of T
  - P = pattern of $n–k+1$ bits; this is the predetermined divisor
  - Q = Quotient
  - R = Remainder

- For T/P to have no remainder, start with

$$T = 2^{n-k}D + F$$

- Divide $2^{n-k}D$ by P gives quotient and remainder

$$\frac{2^{n-k}D}{P} = Q + \frac{R}{P}$$

- Use remainder as FCS

$$T = 2^{n-k}D + R$$

- Does R cause T/P have no remainder?

$$\frac{T}{P} = \frac{2^{n-k}D + R}{P} = \frac{2^{n-k}D}{P} + \frac{R}{P}$$

- Substituting,

$$\frac{T}{P} = Q + \frac{R}{P} + \frac{R}{P} = Q + \frac{R+R}{P} = Q$$

  - No remainder, so T is exactly divisible by P

CRC using Polynomial

- All values expressed as polynomials
  - Dummy variable X with binary coefficients

$$\frac{X^{n-k}D(X)}{P(X)} = Q(X) + \frac{R(X)}{P(X)}$$

$$T(X) = X^{n-k}D(X) + R(X)$$

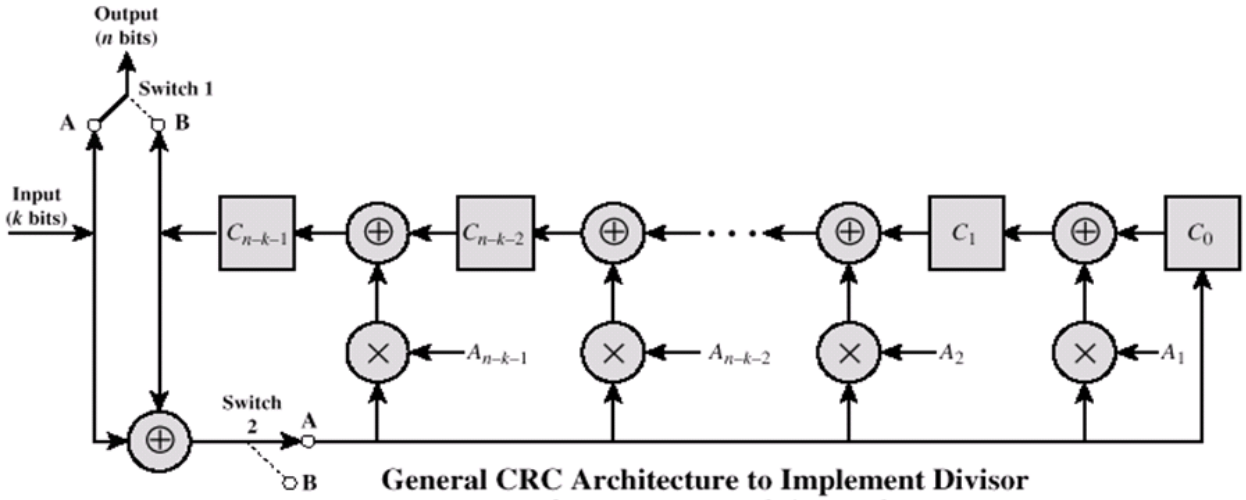
- Widely used versions of $P(X)$
  - CRC–12
    - $X^{12} + X^{11} + X^3 + X^2 + X + 1$
  - CRC–16
    - $X^{16} + X^{15} + X^2 + 1$
  - CRC – CCITT

- $X^{16} + X^{12} + X^5 + 1$

- CRC – 32
  - $X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + X + 1$

CRC using Digital Logic

- **Dividing circuit consisting of:**
  - **XOR gates**
    - **Up to *n* – *k* XOR gates**
    - **Presence of a gate corresponds to the presence of a term in the divisor polynomial *P*(*X*)**
  - **A shift register**
    - **String of 1-bit storage devices**
    - **Register contains *n* – *k* bits, equal to the length of the FCS**



General CRC Architecture to Implement Divisor
$1 + A_1X + A_2X^2 + \ldots + A_{n-1}X^{n-k-1} + X^{n-k}$

Chapter 3

Mobile Internet

3.1 WAE

3.2 WML

3.3 WAP 2.0 protocol

3.4 XHTML

3.5 CHTML

3.1 WAE

Wireless Application Protocol (WAP) is a result of continuous work to define an industry wide specification for developing applications that operate over wireless communication networks. The wireless market is growing very quickly and reaching new customers and services. To enable operators and manufacturers to meet the challenges in advanced services, differentiation and fast/flexible service creation, WAP defines a set of protocols in transport, session and application layers.

Wireless Application Environment (WAE) is part of the WAP Forum's effort to specify an application framework for wireless terminals such as mobile phones, pagers, and PDAs. The framework extends and leverages other WAP technologies, including WTP and WSP, as well as other Internet technologies such as XML, URLs, scripting, and various media types. The effort enables operators, manufacturers, and content developers to meet the challenges in building advanced and differentiating services and implementations in a fast and flexible manner.

- General WAE Features

The following sections provide a specification for the core components of Wireless Application Environment (WAE), including the Wireless Markup Language (WML), the Wireless Markup Scripting language (WMLScript), WAE User Agents and WAE supported media types.

- Session Layer Interface

The WML and the Wireless Telephony Application (WTA) user agents communicate using the Wireless Session Protocol (WSP) over one or more WSP sessions per user agent. This network communication is in the form of WSP/HTTP 1.1 headers and content. The WSP session is created and controlled by the Session Management Entity. The Session Management Entity is not defined within the WAP specification framework, and is implementation specific.

- Basic Authentication Scheme

  WAE user agents should implement Basic Authentication as specified in the HTTP 1.1 specification [RFC2068].
- URL Schemes
- The following standard URL scheme is defined for WAP User Agents:

http: This scheme identifies a particular URL syntax suitable for naming resources stored on HTTP origin servers (see[RFC2396]). The specification of an http scheme does not imply the use of a particular communication protocol between a phone and network gateway. The origin server specified by the URL may be accessed via a WSP-to-HTTP gateway (or proxy). Alternatively, the URL may specify an network server, which combines the function of WSP gateway and origin server into one entity. In this case, the resource is accessed directly across the WSP protocol. Additional, non-standard URL schemes are defined to access client/terminal specific content within the WTAI specification see □WTAI□ Since these schemes are specific to a particular WAE user agent, they are not included in this section.

- User Agent Characteristics
  In order to optimise the WAE client-server model, a number of characteristics are sent from the user agent to the WAP origin server. These characteristics allow the origin server to avoid sending inappropriate content to the user agent. They also provide the server and gateway with a means of customising the response for a particular user agent. The WSP layer provides typed data transfer for the WAE layer (see reference [WSP]). The WSP/HTTP 1.1 content headers are used to perform content negotiation and define character set encoding and language settings. The origin server or WAP gateway may need or want to modify responses based on characteristics of the user agent. For each WAP-defined media type included in the WSP/HTTP Accept header, the user agent should include a parameter, named uaprof, specifying the URI for a profile specifying the user agent characteristics.

- Wireless Markup Language
  The specification of the WML language is available in □WML□
- WMLScript
  The specification of the WMLScript language is available □WMLScript□

□□□□□ WAE User Agents

The WML User Agent is a fundamental component of the WAE. However, WAE is not limited to WML User Agents. WAE allows the integration of domain specific user agents with varying architectures and environments. In particular, a WTA (Wireless Telephony Application) User Agent and a WTAI (Wireless Telephony Application Interface) programming interface has been specified as part of the WAE specification for the mobile telephony environment. The WTAI functions allow authors to access and interact with mobile phone features (eg, call control) as well as other user agents such as phone book and calendar user agents not specified by WAE.

- **WTA User Agent**
  The WTA User Agent is not fully specified as part of the WAP Standards specification. A number of requirements and guidelines are specified in [WTA]. A specification of the Wireless Telephony Application Interface (WTAI), which is the WAP Telephony Value Added Service API, is available in □WTAI□

□□□□□□**WML User Agent**

The WML User Agent is not fully specified as part of the WAP Standards specification. A number of requirements and guidelines are provided as part of the WML language specification (see section 5.1.5) and the WMLScript language specification

- WAE Media Types

  WAE specifies or adopts a number of content formats that facilitate inter-operable exchange of data. The most important formats are the encoded WML and the WMLScript bytecode formats. The encoding of WML and WMLScript makes transmission of WML and WMLScript more efficient and minimises the computational efforts needed to execute them on the client. WAE adopts an additional class of media types to facilitate the exchange of data objects between client and server or between two clients. These are currently limited to electronic business cards and electronic calendar objects.Such objects may be exchanged using WDP datagrams or through a WSP session. In case of exchange over datagrams, a set of well-defined ports has been reserved for the exchange to allow interoperability between different implementations(see [WTP]).Other content formats include the WAE image exchange format and application specific formats. In general the method of data exchange depends on the data type and the user agent involved.

- **Encoded WML format**

  The WML content format is defined in □WML□□and [WBXML].

- **Encoded WMLScript format**

  The WMLScript content format is defined in □WMLScript□

- **The Electronic Business Card Format (vCard 2.1)**

The vCard format was defined by the Versit Consortium and is currently administered by the IMC.

- **The Electronic Calendar and Scheduling Exchange Format (vCalendar 1.0)**

  The vCalendar format was specified by the Versit Consortium and is currently administered by the IMC.

- **Images**

  WAE provides a visual environment that is designed to address several competing requirements, including support for multiple pixels depths, support for colourspace tables, small encoding, very low CPU and RAM decoding and presentation demands and allowance for commonly available tools and support. WAE meets these unique requirements by: □□ Supporting standard WSP/HTTP media types for commonly used image formats, eg, *image/png.*□□ Introducing an optimised bitmap format, the Wireless BitMaP (WBMP) (WSP/HTTP media type image/xwap.wbmp).

  WBMP is an encapsulation format, ie, a WBMP object is a wrapper object which maps the verbose headers of the full image format to an identification (or typing) of the contents. The actual image contents contain all other information, eg, colour table (if any), image bit planes, etc.

  The WBMP specification is thus divided into two parts:

  1. The generic header contains the following information, which is common to all image formats.

a) Type
b) Width and height
c) WBMP version number.

The type identifier denotes the format of the embedded image. Type 0 is currently specified.

2. The type-specific formats specification, indicating the data format for a particular WBMP type. The WBMP format supports the definition of compact image formats suitable for encoding a wide variety of image formats and provides the means for optimisation steps such as stripping of superfluous headers and special purpose compression schemes. This leads to efficient communication to and from the client and for efficient presentation in the client display.

A WBMP image has the following characteristics:

- Compact binary encoding
- Scaleability, ie, future support for all image qualities and types (colour depths, animations, stream data, etc.)
- Extensibility (unlimited type definition space)
- Optimised for low computational costs in the client.

- **Multipart Messages**

WAE includes a multipart encoding specification, suitable for exchanging multiple typed entities over WSP. WSP translates the MIME multipart entity (see [RFC2045]) into a compact binary form, which is optimised for narrowband environments. See WSP.

- **WTA Events**

WAE defines a separate content type and encoding for delivering events from the WTA server/gateway to the WTA User Agent. See WTA for details on the event content type.

3.2 WML

WML (Wireless Markup Language) 1.x is the markup language defined in the WAP 1.x specification. WAP is the standard created by the WAP Forum (now the Open Mobile Alliance [OMA]) that brings the World Wide Web to wireless devices. It specifies the protocol stack and application environment of mobile Internet browsing applications. The role of WML in mobile Internet applications is the same as that of HTML in web applications. WAP sites are written in WML, while web sites are written in HTML. WML 1.x is very similar to HTML. Both of them use tags and are written in plain text format. Some tags in WML 1.x are directly borrowed from HTML. If you have experience in using the HTML markup language, you should be able to learn WML 1.x quickly.

Some features of WML 1.x are specific to wireless devices. For example, WML 1.x provides a way for developers to program the softkeys of mobile phones. This feature is not supported in HTML since computers do not have any softkeys. The most up-to-date version of the WAP 1.x specification is WAP 1.2.1, which defines WML 1.3. WML files have the extension ".wml". The MIME type of WML is "text/vnd.wap.wml". WML supports client-side scripting. The scripting language supported is called WMLScript. Its syntax is based on JavaScript. If you want to learn it, our WMLScript tutorial will be a good starting point for you.

WML 2.0

WAP site developers need not to care about WML 2.0. WML 2.0 is created for backward compatibility purposes and it is not for use by WAP site developers. To develop a WAP site with the WAP 2.0 standard, use XHTML Mobile Profile.
Eg.

```
<?xml version="1.0"?>
<!DOCTYPE wml PUBLIC "-//WAPFORUM//DTD WML 1.3//EN"
"http://www.wapforum.org/DTD/wml13.dtd">

<wml>
  <card id="card1" title="WML Tutorial">
    <p>Hello World</p>
  </card>

  <card id="card2" title="WML Tutorial">
    <p>Welcome to the world of WML</p>
</card>
</wml>
```

3.3 WAP 2.0

WAP 1.x is an earlier version of the WAP standard. The most current version is WAP 2.0. The markup language defined in WAP 2.0 is XHTML MP (XHTML Mobile Profile). It is a subset of the XHTML used on the web. XHTML MP supports a mobile version of cascading style sheet called WCSS (WAP CSS). It is a subset of the CSS2 used on the web plus some WAP specific extensions.

Most of the new mobile phone models released are WAP 2.0-enabled. As WAP 2.0 is backward compatible to WAP 1.x, WAP 2.0-enabled mobile devices can display both XHTML MP and WML documents.

WML 1.x is an earlier technology. However, that does not mean it is of no use, since a lot of wireless devices that only supports WML 1.x are still being used. Besides, some useful features of WML are not available in XHTML MP. For example, XHTML MP does not support

events, variables and client-side scripting. The major WML features lost in XHTML MP are discussed in the "WML Features Lost in XHTML MP" section of our XHTML MP tutorial.

If you write the markup of your mobile Internet site in WML, both old and new WAP-enabled wireless devices can be used to view your mobile Internet site. The user base of your WAP application is maximized.

3.4 XHTML

XHTML (eXtensible HyperText Markup Language) is the reformulation of HTML in XML (eXtensible Markup Language). The tags in XHTML are the same as those in HTML. You can consider XHTML as HTML written with the syntax of XML. XHTML has a cleaner and stricter structure than HTML, which ease the parsing of a document. This is particularly important for wireless devices such as mobile phones, since they have limited processing power. All major web browsers support XHTML. XHTML will gradually replace HTML.

XHTML Basic

XHTML Basic is a simplified version of XHTML. It is designed for devices with limited processing power and capabilities such as mobile phones, PDAs, smart watches, pagers, etc. XHTML Basic does not contain XHTML features that are difficult to support on these devices. For example, cascading style sheets, frames, and scripting are not supported in XHTML Basic. XHTML Basic is defined by the W3C (World Wide Web Consortium).

XHTML Mobile Profile

XHTML Mobile Profile is the official markup language in the most recent WAP specification version 2.0 defined by the former WAP Forum. The WAP Forum created XHTML Mobile Profile based on XHTML Basic, with the addition of some elements and attributes from the full version of XHTML such as <i>, <b>, <small>, <big> and <hr>. XHTML Mobile Profile supports a simplified version of cascading style sheet called WCSS / WAP CSS.

XHTML MP

XHTML MP (eXtensible HyperText Markup Language Mobile Profile) is the markup language defined in WAP 2.0. WAP 2.0 is the most recent mobile services specification created by the WAP Forum (now the Open Mobile Alliance [OMA]). The specification of WAP CSS (WAP Cascading Style Sheet or WCSS) is also defined in WAP 2.0. WAP CSS is the companion of XHTML Mobile Profile and they are used together. With WAP CSS, you can easily change and format the presentation of XHTML MP pages.

XHTML Mobile Profile is a subset of XHTML, which is the stricter version of HTML. XHTML Mobile Profile is XHTML Basic (also a subset of XHTML) plus some additional elements and attributes from the full version of XHTML.

The goal of XHTML Mobile Profile is to bring together the technologies for mobile Internet browsing and that for the World Wide Web. Before the coming out of XHTML Mobile Profile, WAP developers make use of WML and WMLScript to create WAP sites, while web developers use HTML / XHTML and CSS style sheets to build web sites.

With the announcement of XHTML Mobile Profile, the markup language of the wireless world and the wired world finally converges. XHTML Mobile Profile and WAP CSS give wireless Internet application developers more and better presentation control. The greatest advantage, however, is that the same technologies can now be used to develop both the web and wireless version of your Internet site. You can use any web browsers to view your WAP2.0 application during the prototyping and development process.
The previous version of WAP is 1.2.1. WAP 1.2.1 sites are developed using WML and WMLScript. WAP 2.0 is backward compatible to WAP 1.x. So, a WAP 2.0 wireless device can be used to visit both XHTML MP / WCSS and WML / WMLScript sites. If you are interested in learning WML or WMLScript, you may want to read our WML tutorial and WMLScript tutorial.

WCSS / WAP CSS

CSS (Cascading Style Sheet) is widely used on the World Wide Web to define how web pages should be presented in browsers. WCSS / WAP CSS is a simplified version of CSS2 with the addition of some WAP specific extensions. WAP CSS is defined in the WAP 2.0 specification. Since WAP CSS is designed for use on wireless devices, CSS2 features that are unsuitable or unnecessary for wireless devices are not included in WAP CSS. WAP CSS enables the separation of the presentation from the content. If you want to change the presentation details of an XHTML MP page, you just need to modify the style sheet. With WAP CSS, you can easily change the layout and style of your XHTML MP pages to suit different user agents.

3.5 CHTML

C-HTML (short for Compact HTML), also called i-mode-HTML, is a subset of the HTML markup language for small information devices, such as smart phones and PDAs, such as DoCoMo's i-mode mobile phones used in Japan. C-HTML adds several features not found in standard HTML, notably accesskeys, phone number shortcuts for links, and emoji pictorial characters as locally extended Shift JIS, all concepts borrowed from HDML/WML.
Because small devices such as cellular phones have hardware restrictions such as lower memory, low power CPUs with limited or no storage capabilities, small mono-color display screens, single-character fonts and restricted input methods (the absence of a keyboard or a mouse), there is a need for a simpler form of HTML.
C-HTML does not support tables, image maps, multiple fonts and styling of fonts, background colors and images, frames, style sheets, and is limited to a monochromatic display.
The language is defined so that all the basic interactive operations can be done by a combination of four buttons and not by two-dimensional cursor movement: cursor forward, cursor backward, select, and back/stop. Functionality requiring two-dimensional cursor pointing, like image maps, are excluded from C-HTML.

Design Principles

The Compact HTML is designed to meet the requirements of small information appliances described above. It is designed based on the following four principles.
(1) Completely based on the current HTML W3C recommendations
Compact HTML is defined as a subset of HTML 2.0, HTML 3.2 and HTML 4.0 specifications. This means that Compact HTML inherits the flexibility and portability from the standard HTML.

(2) Lite Specification
Compact HTML has to be implemented with small memory and low power CPU. Frames and tables which require large memory are excluded from Compact HTML.
(3) Can be viewed on a small mono-color display
Compact HTML assumes a small display space of black and white color. However, it does not assume a fixed display space, but it is flexible for the display screen size. Compact HTML also assumes single character font.
(4) Can be easily operated by the users
Compact HTML is defined so that all the basic operations can be done by a combination of four buttons; Cursor forward, Cursor backward, Select, and Back/Stop(Return to the previous page). The functions which require two-dimensional focus pointing like "image map" and "table" are excluded from Compact HTML.

Features of Compact HTML

The Compact HTML is a subset of HTML 2.0, HTML 3.2 and HTML 4.0. We describe the major features which are excluded from Compact HTML, as follows.
JPEG image
Table
Image map
Multiple character fonts and styles
Background color and image
Frame
Style sheet
We define that Compact HTML includes GIF image support. It should be noted that this subset does not require two-dimensional cursor moving, that is, it can be operated by using only four buttons. We can also expect that well-designed pages for small display fit the screen space and the scrolling is not necessary. Actually the Compact HTML browser can display the pages like "deck of cards" by HDML.Since the memory capacity is the most important issue in implementing the Compact HTML browser, we recommend the buffer limit for some functions.

INPUT
The maximum buffer size is 512 bytes.
SELECT
The maximum buffer size is 4096 bytes.
Though such a limitation belongs to the implementation issues, the common criteria is useful while developing devices.
One recommended implementation for the browser is to support the direct selection of anchors by using number buttons. For example, when five anchors are contained in an HTML page, the third anchor can be selected just by pressing the "3" button. (The HTML 4.0 specification includes a new attribute "accesskey" for the similar purpose of direct key assignment.)

Examples

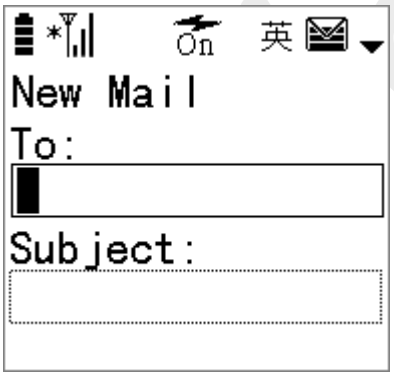
Here we describe the examples of applications by using Compact HTML. The following examples show the compact browser for cellular phones. The screen is the space of 7 text lines and 16 characters wide. The top line is used for displaying the status information.

(1) Compact HTML example: Simple Menu

In this example, the cursor focus point is expressed as the reverse text.

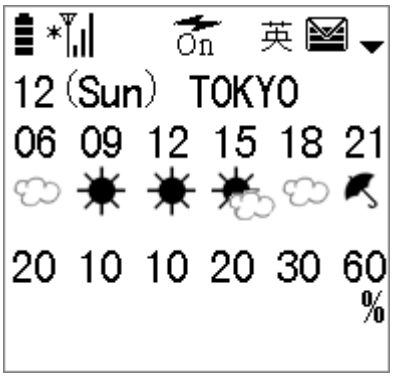
(2) Compact HTML example: Mail Send Form

This example shows the mail sending form using INPUT tags. The focused form is expressed as solid surrounding lines, and non-focused forms are expressed as dotted surrounding line. The cursor point for input characters is expressed as a reverse box.



(3) Compact HTML example: Image Contents

This example shows weather and rain information of the day. It uses mono-color GIF image.



Practical implementations and experiments show that Compact HTML is enough useful for small screen of 5-10 text lines and 10-20 characters wide.

Benefits of Compact HTML

The Compact HTML, an HTML-based approach, guarantees that small information appliances can connect to the open WWW world. Compact HTML keeps the advantage of HTML features and solves the problems arising from the restrictions of small information appliances.
The Compact HTML specification can be referred to by the tools like HTML authoring systems. In addition, the client-specific web services for such small devices can be realized by using user agent attributes. That is, the server can do the content filter for Compact HTML.

Chapter 4

Cellular Networks

4.1 Frequency reuse

4.2 First Generation systems

4.3 Second Generation systems

4.4 Third Generation systems

4.5 GSM

4.6 CDMA fundamentals

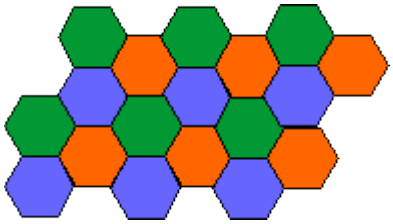
4.1 Frequency reuse

Frequency reuse is a technique of reusing frequencies and channels within a communications system to improve capacity and spectral efficiency. Frequency reuse is one of the fundamental concepts on which commercial wireless systems are based that involves the partitioning of an RF radiating area (cell) into segments of a cell. One segment of the cell uses a frequency that is far enough away from the frequency in the bordering segment that it does not provide interference problems. Frequency re-use in mobile cellular systems means that each cell has a frequency that is far enough away from the frequency in the bordering cell that it does not provide interference problems. The same frequency is used at least two cells apart from each other. This practice enables cellular providers to have many times more customers for a given site license.

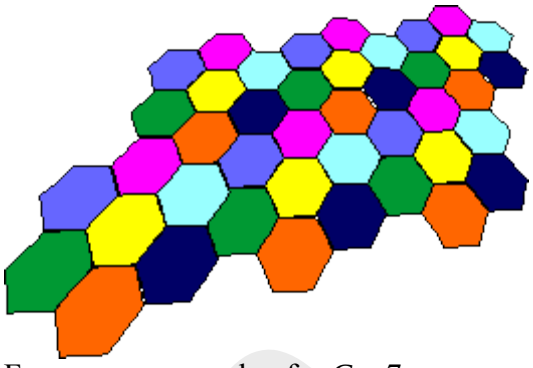
Principles of cellular frequency reuse

In the cellular concept, frequencies allocated to the service are re-used in a regular pattern of areas, called 'cells', each covered by one base station. In mobile-telephone nets these cells are usually hexagonal. In radio broadcasting, a similar concept has been developed based on rhombic cells.

To ensure that the mutual interference between users remains below a harmful level, adjacent cells use different frequencies. In fact, a set of C different frequencies$\{f_1, ..., f_C\}$ are used for each cluster of C adjacent cells. Cluster patterns and the corresponding frequencies are re-used in a regular pattern over the entire service area.



Frequency reuse plan for $C = 3$, with hexagonal cells. (i=1, j =1)



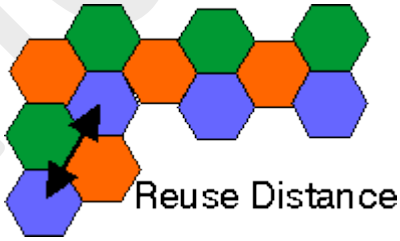
Frequency reuse plan for $C = 7$ (i=2, j =1).

The total bandwidth for the system is C times the bandwidth occupied by a single cell.

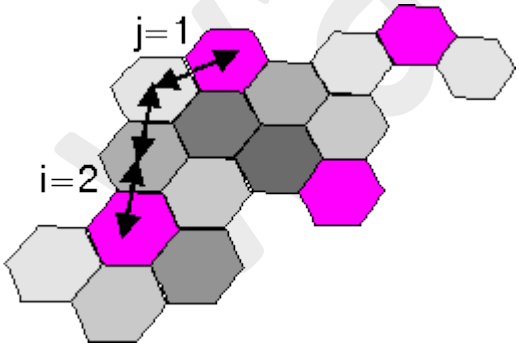
Reuse Distance

The closest distance between the centers of two cells using the same frequency (in different clusters) is determined by the choice of the cluster size C and the lay-out of the cell cluster. This distance is called the frequency "re-use" distance. It can be shown that the reuse distance r_u, normalized to the size of each hexagon, is

r_u = SQRT{3 C}



For hexagonal cells, i.e., with 'honeycomb' cell lay-outs commonly used in mobile radio, possible cluster sizes are $C = i^2 + ij + j^2$, with integer i and j (C = 1, 3, 4, 7, 9, ...). Integers i and j determine the relative location of co-channel cells.



7-cell reuse with i = 2 and j =1.

4.2. **First Generation systems**

First-generation mobile networks are analog systems. Some of the more widely deployed first-generation networks include AMPS and NMT. In this section we focus the discussion on AMPS.

The Advanced Mobile Phone Service (AMPS) is in wide use even today, almost 25 years after it was introduced. AMPS was conceived by Bell Labs in the 1970s, and improvements in the form of digital AMPS (D-AMPS) were made in the late 1980s. The AMPS air interface is specified in EIA/TIA-553. AMPS is based on FDMA.

The FCC allocated a total of 50 MHz (25 MHz on the A side and B side) in the 800-MHz spectrum for AMPS. Each voice channel is allocated a 30-KHz portion of the bandwidth within the AMPS frequency allocations. Because each carrier has 25 MHz of spectrum, this provides a total of 832 (25 MHz/30 KHz) cellular channels (forward and reverse). However, since the same frequency cannot be used in adjacent cells, the 416 duplex channels are a theoretical maximum (actual number of valid voice channels equals 312). An AMP uses the seven-cell frequency reuse method. Control channels are used to set up and clear calls as well as other control messages. Each band (25 MHz) contains 21 control channels. When a mobile station is not in session, it must monitor designated control channels. It tunes and locks into the strongest channel to receive system information. The forward control channel (FOCC) is a data stream from the base station to the mobile, and the reverse control channel (RECC) is from the mobile to the base station. Voice conversation is carried over the forward voice channel (FVC) and the reverse voice channel (RVC).

The identifiers used in AMPS are as follows:

- The mobile station's electronic serial number (ESN)
- The mobile operator's system identification (SID)
- The mobile station's mobile identification number (MIN)

The ESN for a mobile is a 32-bit number that uniquely identifies a mobile and is set up by the mobile manufacturer. System IDs (SIDs) are 15-bit binary numbers that are assigned to cellular systems. One of the uses of the SID is to determine a home network from a roaming network. The MIN is a 34-bit number that is derived from the mobile terminal's 10-digit telephone number.

The network utilizes the IS-41 protocol for mobility and authentication procedures. The MSC provides the capability for call processing, and the HLR and VLRs keep track of the mobile as it

moves. The mobile terminal is responsible for updating its location as it moves in the cellular network.

Data services in AMPS are straightforward and analogous to dial-up networking. Because AMPS is an analog technology, it is possible to make use of standard modems directly with AMPS. Data rates are at a maximum of 14.4 Kbps irrespective of the modem protocol (v.90 or others).

D-AMPS

D-AMPS, or digital AMPS, is a hybrid air interface that uses both first-generation and second-generation technology. The D-AMPS specification is detailed in IS-54-B. The primary reason for introducing D-AMPS in the early 1990s in North America was to overcome some of the shortcomings of AMPS technology. The co-channel interference problem of AMPS limited its capacity significantly, and the 30-KHz channel assigned to each user is excess capacity on a per user basis. The hybrid nature of D-AMPS comes from the fact that second-generation TDMA technology is placed on AMPS traffic channels.

The AMPS channels are still used, but the content and formats of the 30-KHz channels are modified. The channels defined for D-AMPS are as follows:

- FOCC— Forward analog control channel; direction: base station (BS) to mobile station (MS) control channel
- FVC— Forward voice channel; direction: BS to MS voice channel
- FDTC— Forward digital traffic channel; direction: BS to MS digital user and control channel
- RECC— Reverse analog control channel; direction: MS to BS control channel
- RVC— Reverse analog voice channel; direction: MS to BS voice channel
- RDTC— Reverse digital traffic channel; direction: MS to BS digital user and control channel

The FDTC and RDTC can be split up into fast associated control channel (FACCH) and slow associated control channel (SACCH), which are used for signaling. One of the improvements

that were made in the handoff process was the involvement of the mobile in the handoff procedure. Mobile assisted handoff was introduced in D-AMPS. The MS keeps measuring the quality of the forward channel and sends these measurements to the BS to allow the network to make a more informed decision.

First-generation AMPS and D-AMPS mobile networks continue to exist even today, especially in the United States. They complement coverage of second-generation digital networks such as GSM and IS-95. Most mobile terminals are dual mode (i.e., they incorporate a second-generation (2G) digital radio as well as the analog radio). With roaming agreements in place, 2G network operators can claim nationwide coverage. However, it is expected that the lifetime of these analog networks is coming to an end and will be decommissioned slowly in the next few years. One of the reasons for decommissioning these networks is to reclaim the spectrum for other uses.

4.3 Second Generation Systems

The second generation (2G) of the wireless mobile network was based on low-band digital data signalling. The most popular 2G wireless technology is known as Global Systems for Mobile Communications (GSM). GSM systems, first implemented in 1991, are now operating in about 140 countries and territories around the world. An estimated 248 million users now operate over GSM systems. GSM technology is a combination of Frequency Division Multiple Access (FDMA) and Time Division Multiple Access (TDMA). The first GSM systems used a 25MHz frequency spectrum in the 900MHz band. FDMA is used to divide the available 25MHz of bandwidth into 124 carrier frequencies of 200kHz each. Each frequency is then divided using a TDMA scheme into eight timeslots. The use of separate timeslots for transmission and reception simplifies the electronics in the mobile units. Today, GSM systems operate in the 900MHz and 1.8 GHz bands throughout the world with the exception of the Americas where they operate in the 1.9 GHz band.

In addition to GSM, a similar technology, called Personal Digital Communications (PDC), using TDMA-based technology, emerged in Japan. Since then, several other TDMA-based systems have been deployed worldwide and serve an estimated 89 million people worldwide. While GSM technology was developed in Europe, Code Division Multiple Access

(CDMA) technology was developed in North America. CDMA uses spread spectrum technology to break up speech into small, digitized segments and encodes them to identify each call. CDMA systems have been implemented worldwide in about 30 countries and serve an estimated 44 million subscribers. While GSM and other TDMA-based systems have become the dominant 2G wireless technologies, CDMA technology is recognized as providing clearer voice quality with less background noise, fewer dropped calls, enhanced security, greater reliability and greater network capacity.

The Second Generation (2G) wireless networks mentioned above are also mostly based on circuit-switched technology. 2G wireless networks are digital and expand the range of applications to more advanced voice services, such as Called Line Identification. 2G wireless technology can handle some data capabilities such as fax and short message service at the data rate of up to 9.6 kbps, but it is not suitable for web browsing and multimedia applications.

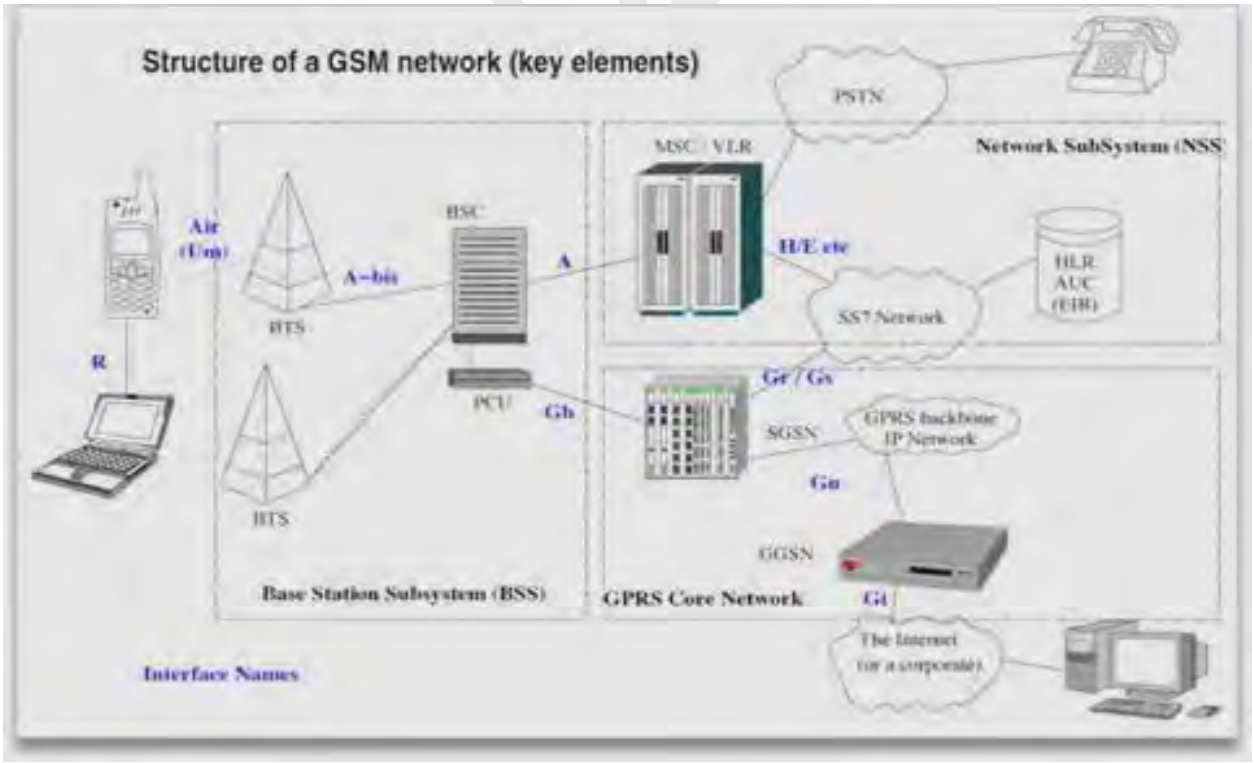
4.4. Third Generation Systems

3G wireless technology represents the convergence of various 2G wireless telecommunications systems into a single global system that includes both terrestrial and satellite components. One of the most important aspects of 3G wireless technologies is its ability to unify existing cellular standards, such as CDMA, GSM, and TDMA, under one umbrella. The following three air interface modes accomplish this result: wideband CDMA, CDMA2000 and the Universal Wireless Communication (UWC-136) interfaces. Wideband CDMA (W-CDMA) is compatible with the current 2G GSM networks prevalent in Europe and parts of Asia. W-CDMA will require bandwidth of between 5Mhz and 10 Mhz, making it a suitable platform for higher capacity applications. It can be overlaid onto existing GSM, TDMA (IS-36) and IS95 networks. Subscribers are likely to access 3G wireless services initially via dual band terminal devices. W-CDMA networks will be used for high-capacity applications and 2G digital wireless systems will be used for voice calls. The second radio interface is CDMA2000 which is backward compatible with the second generation CDMA IS-95 standard predominantly used in US. The third radio interface, Universal Wireless Communications – UWC-136, also called IS-136HS, was proposed by the TIA and designed to comply with ANSI-136, the North American TDMA standard. 3G wireless networks consist of a Radio Access Network (RAN) and a core network. The core network consists of a packet-switched domain, which includes 3G SGSNs and GGSNs, which provide the same functionality that they provide in a GPRS system, and a circuit-switched domain, which includes 3G MSC for switching of voice calls. Charging for services and access is done through the Charging Gateway Function (CGF), which is also part of the core network. RAN functionality is independent from the core network functionality. The access network provides a core network technology independent access for mobile terminals to different types of core networks and network services. Either core network domain can access any appropriate RAN service; e.g. it should be possible to access a “speech” radio access bearer from the packet switched domain.

The Radio Access Network consists of new network elements, known as Node B and Radio Network
Controllers (RNCs). Node B is comparable to the Base Transceiver Station in 2G wireless networks. RNC
replaces the Base Station Controller. It provides the radio resource management, handover control and
support for the connections to circuit-switched and packet-switched domains. The interconnection of the
network elements in RAN and between RAN and core network is over Iub, Iur and Iu interfaces based on
ATM as a layer 2 switching technology. Data services run from the terminal device over IP, which in turn
uses ATM as a reliable transport with QoS. Voice is embedded into ATM from the edge of the network
(Node B) and is transported over ATM out of the RNC. The Iu interface is split into 2 parts: circuit switched and packet-switched. The Iu interface is based on ATM with voice traffic embedded on virtual circuits using AAL2 technology and IP-over-ATM for data traffic using AAL5 technology. These traffic types are switched independently to either 3G SGSN for data or 3G MSC for voice.

4.5 GSM:

Global System for Mobile communications (GSM)



Global System for Mobile communications (GSM) originally known as Groupe Spécial Mobile is the most popular standard for mobile phones in the world. Its promoter, the GSM Association, estimates that 82% of the global mobile market uses the standard. GSM is used by over 2 billion people across more than 212 countries and territories. Its ubiquity makes international roaming very common between mobile phone operators, enabling subscribers to use their phones in many parts of the world. GSM differs from its predecessors in that both signaling and speech channels are digital call quality, and thus is considered a second generation (2G) mobile phone system. This has also meant that data communications were built into the system using the 3rd Generation Partnership Project (3GPP).

The ubiquity of the GSM standard has been advantageous to both consumers (who benefit from the ability to roam and switch carriers without switching phones) and also to network operators (who can choose equipment from any of the many vendors implementing GSM. GSM also pioneered a low-cost alternative to voice calls, the Short message service (SMS, also called "text messaging"), which is now supported on other mobile standards as well.

Newer versions of the standard were backward-compatible with the original GSM phones. For example, Release '97 of the standard added packet data capabilities, by means of General Packet Radio Service (GPRS). Release '99 introduced higher speed data transmission using Enhanced Data Rates for GSM Evolution (EDGE).

Technical Details

GSM is a cellular network, which means that mobile phones connect to it by searching for cells in the immediate vicinity. GSM networks operate in four different frequency ranges. Most GSM networks operate in the 900 MHz or 1800 MHz bands. Some countries in the Americas (including Canada and the United States) use the 850 MHz and 1900 MHz bands because the 900 and 1800 MHz frequency bands were already allocated.

The rarer 400 and 450 MHz frequency bands are assigned in some countries, notably Scandinavia, where these frequencies were previously used for first-generation systems.

Network Structure, Components and Security

The network behind the GSM system seen by the customer is large and complicated in order to provide all of the services which are required.

GSM Sections - It is divided into a number of sections including Base Station Subsystem (the base stations and their controllers), Network and Switching Subsystem (the part of the network most similar to a fixed network, sometimes also just called the core network) and GPRS Core Network (the optional part which allows packet based Internet connections). All of these elements in the system combine to produce many GSM services such as voice calls and data transfer over SMS and GPRS.

Subscriber Identity Module - One of the key features of GSM is the Subscriber Identity Module (SIM), commonly known as a SIM card. The SIM is a detachable smart card containing the user's subscription information and phonebook. This allows the user to retain his or her information after switching handsets. Alternatively, the user can also change operators while retaining the handset simply by changing the SIM.

GSM Security - GSM was designed with a moderate level of security. The system was designed to authenticate the subscriber using a pre-shared key and challenge-response. Communications between the subscriber and the base station can be encrypted. The development of UMTS introduces an optional USIM, that uses a longer authentication key to give greater security, as well as mutually authenticating the network and the user - whereas GSM only authenticated the user to the network (and not vice versa). The security model therefore offers confidentiality and authentication, but limited authorization capabilities, and no non-repudiation.

Data Communication

Short Message Service (SMS) - It is a communications protocol allowing the interchange of short text messages between mobile telephone devices. The SMS technology has facilitated the development and growth of text messaging. The connection between the phenomenon of text messaging and the underlying technology is so great that in parts of the world the term "SMS" is used colloquially as a synonym for a text message from another person or the act of sending a text message. Most SMS messages are mobile-to-mobile text messages, though the standard supports other types of broadcast messaging as well.

Typically a text message originating at a handset is sent to a Short Message Service Centre (SMSC). The SMSC then attempts to send the message to its recipient. If the recipient is not reachable, the SMSC queues the message for later retry. This mechanism is characterized as a store-and-forward delivery system. SMS message transmission is also characterized as best effort: there are no guarantees that a message will actually be delivered to its recipient, and delay or complete loss of a message is not ruled out by the specification of the protocol.

SMS messages can be used for text communication between mobile phone users, for carrying queries and responses between the phone user and computerized information services, or by the network operator for various management tasks. A variety of technologies have been developed to allow SMS messages to be interchanged with other networks, e.g. analogue phone lines or the Internet (websites, email etc).

General Packet Radio Service (GPRS) - It is a Mobile Data Service available to users of Global System for Mobile Communications (GSM) and IS-136 mobile phones. It provides data rates from 56 up to 114 kbps.

GPRS data transfer is typically charged per kilobyte of transferred data, while data communication via traditional circuit switching is billed per minute of connection time, independent of whether the user has actually transferred data or has been in an idle state. GPRS can be used for services such as Wireless Application Protocol (WAP) access, Short Message

Service (SMS), Multimedia Messaging Service (MMS), and for Internet communication services such as email and World Wide Web access.

Usually, GPRS data are billed per kilobyte of information transceived, while circuit-switched data connections are billed per second. The latter is inefficient because even when no data are being transferred, the bandwidth is unavailable to other potential users.

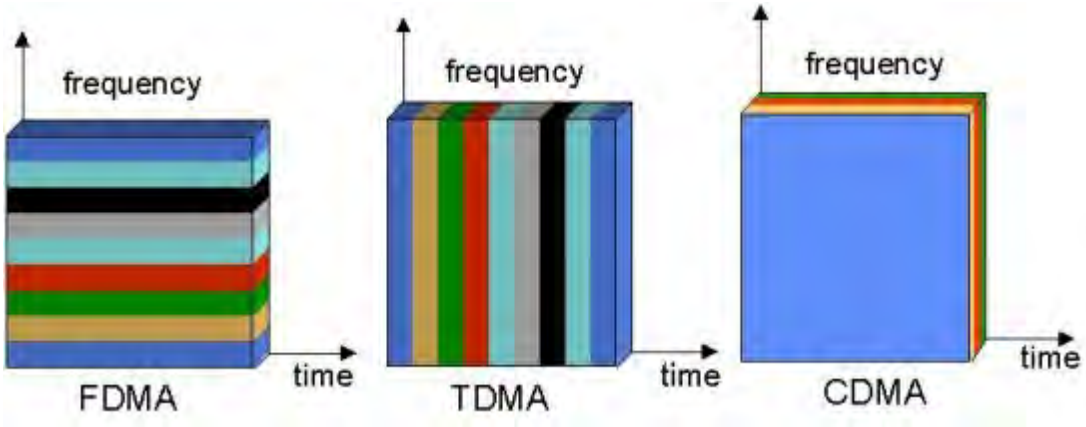
The multiple access methods used in GSM with GPRS are based on frequency division duplex (FDD) and TDMA. During a session, a user is assigned to one pair of up-link and down-link frequency channels. This is combined with time domain statistical multiplexing, i.e. packet mode communication, which makes it possible for several users to share the same frequency channel. The packets have constant length, corresponding to a GSM time slot. The down-link uses first-come first-served packet scheduling, while the up-link uses a scheme very similar to reservation ALOHA. This means that slotted Aloha (S-ALOHA) is used for reservation inquiries during a contention phase, and then the actual data is transferred using dynamic TDMA with first-come first-served scheduling.

GPRS originally supported (in theory) Internet Protocol (IP), Point-to-Point Protocol (PPP) and X.25 connections. The last has been typically used for applications like wireless payment terminals, although it has been removed from the standard. X.25 can still be supported over PPP, or even over IP, but doing this requires either a router to perform encapsulation or intelligence built in to the end-device/terminal e.g. UE(User Equipment). In practice, when the mobile built-in browser is used, IPv4 is being utilized. In this mode PPP is often not supported by the mobile phone operator, while IPv6 is not yet popular. But if the mobile is used as a modem to the connected computer, PPP is used to tunnel IP to the phone. This allows DHCP to assign an IP Address and then the use of IPv4 since IP addresses used by mobile equipment tend to be dynamic.

4.6 Fundamentals of CDMA:

ACCESS SCHEMES

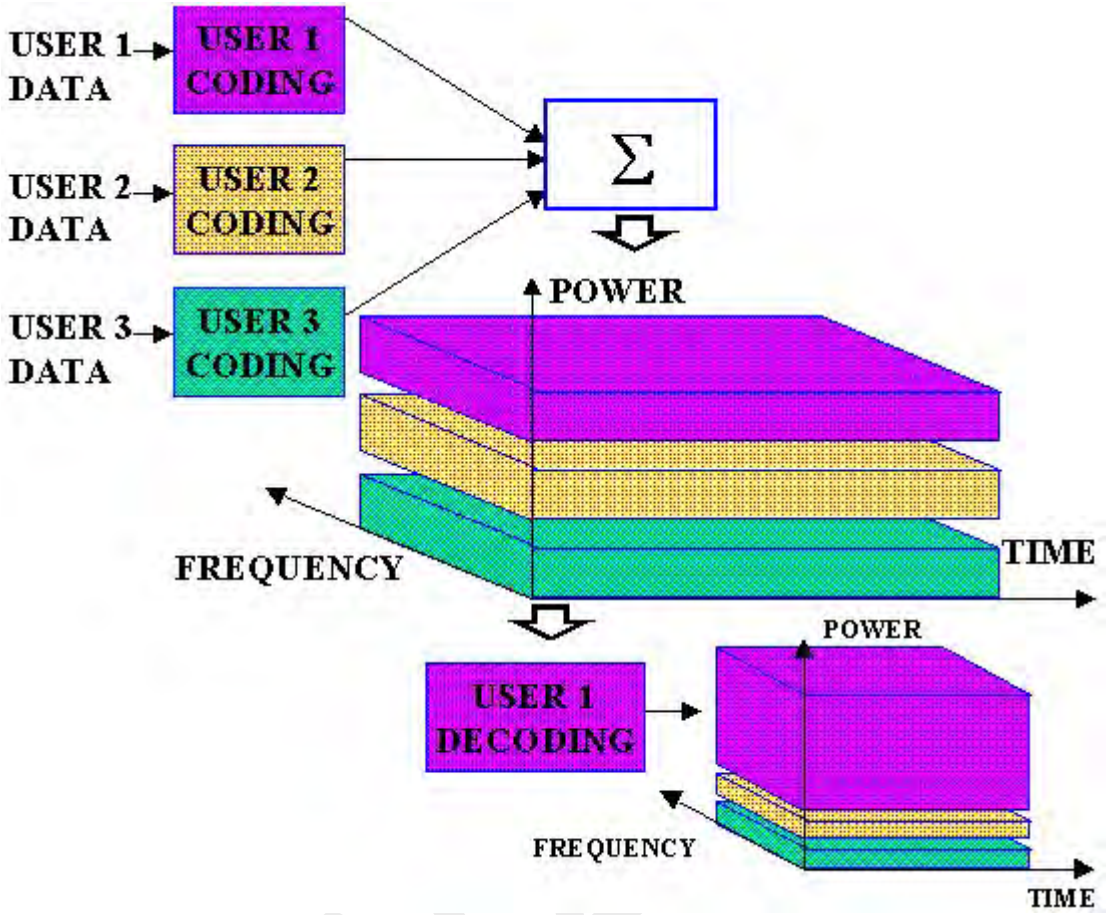
For radio systems there are two resources, frequency and time. Division by frequency, so that each pair of communicators is allocated part of the spectrum for all of the time, results in Frequency Division Multiple Access (FDMA). Division by time, so that each pair of communicators is allocated all (or at least a large part) of the spectrum for part of the time results in Time Division Multiple Access (TDMA). In Code Division Multiple Access (CDMA), every communicator will be allocated the entire spectrum all of the time. CDMA uses codes to identify connections.



Multiple Access Schemes

CODING

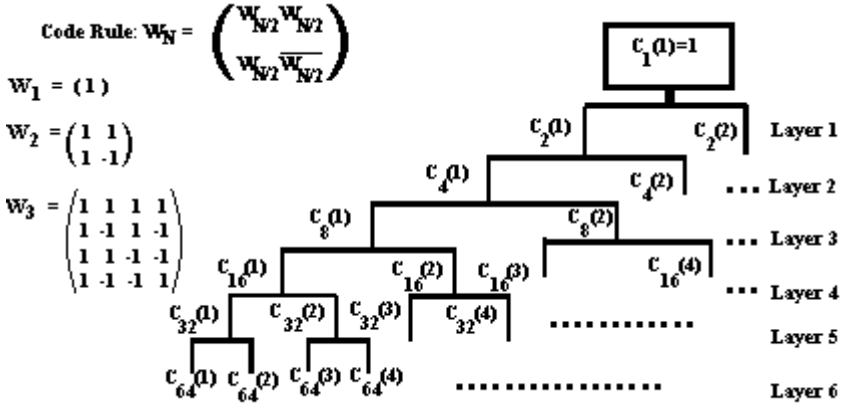
CDMA uses unique spreading codes to spread the baseband data before transmission. The signal is transmitted in a channel, which is below noise level. The receiver then uses a correlator to despread the wanted signal, which is passed through a narrow bandpass filter. Unwanted signals will not be despread and will not pass through the filter. Codes take the form of a carefully designed one/zero sequence produced at a much higher rate than that of the baseband data. The rate of a spreading code is referred to as chip rate rather than bit rate.



CDMA spreading

CODES

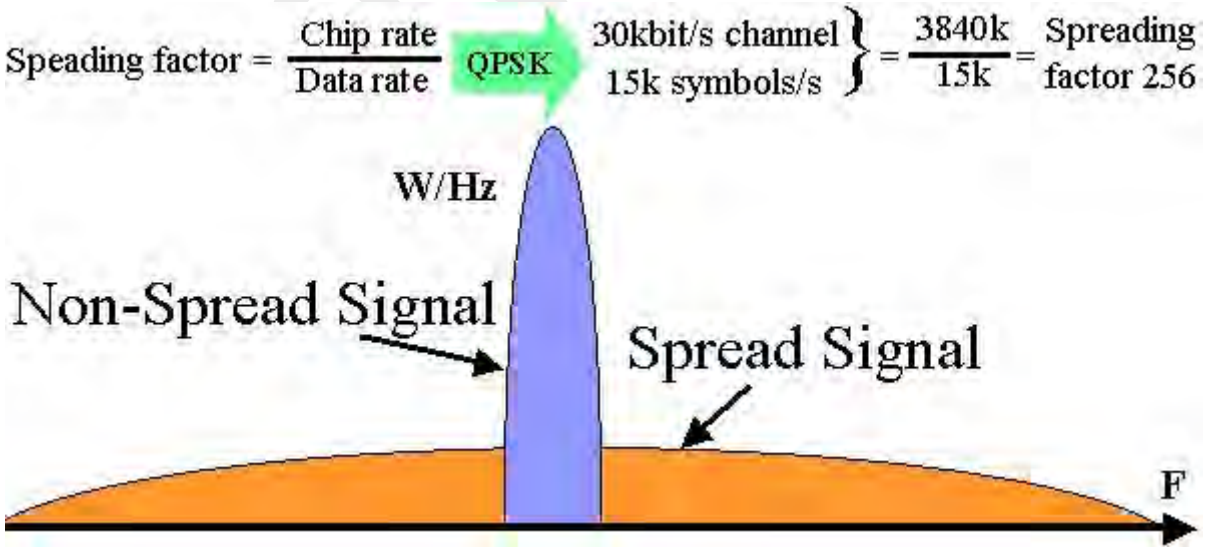
CDMA codes are not required to provide call security, but create a uniqueness to enable call identification. Codes should not correlate to other codes or time shifted version of itself. Spreading codes are noise like pseudo-random codes, channel codes are designed for maximum separation from each other and cell identification codes are balanced not to correlate to other codes of itself.



Example OVSF codes, used in channel coding

THE SPREADING PROCESS

WCDMA uses Direct Sequence spreading, where spreading process is done by directly combining the baseband information to high chip rate binary code. The Spreading Factor is the ratio of the chips (UMTS = 3.84Mchips/s) to baseband information rate. Spreading factors vary from 4 to 512 in FDD UMTS. Spreading process gain can in expressed in dBs (Spreading factor 128 = 21dB gain).



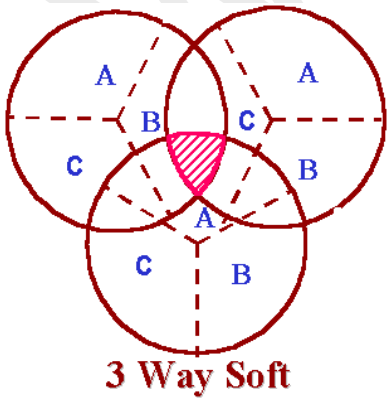
CDMA spreading

POWER CONTROL

CDMA is interference limited multiple access system. Because all users transmit on the same frequency, internal interference generated by the system is the most significant factor in determining system capacity and call quality. The transmit power for each user must be reduced to **limit interference**, however, the power should be enough to maintain the required Eb/No (signal to noise ratio) for a satisfactory **call quality**. Maximum **capacity** is achieved when Eb/No of every user is at the minimum level needed for the acceptable channel performance. As the MS moves around, the RF environment continuously changes due to fast and slow fading, external interference, shadowing , and other factors. The aim of the dynamic power control is to limit transmitted power on both the links while maintaining link quality under all conditions. Additional advantages are longer mobile battery life and longer life span of BTS power amplifiers

HANDOVER

Handover occurs when a call has to be passed from one cell to another as the user moves between cells. In a traditional "hard" handover, the connection to the current cell is broken, and then the connection to the new cell is made. This is known as a "break-before-make" handover. Since all cells in CDMA use the same frequency, it is possible to make the connection to the new cell before leaving the current cell. This is known as a "make-before-break" or "soft" handover. Soft handovers require less power, which reduces interference and increases capacity. Mobile can be connected to more that two BTS the handover. "Softer" handover is a special case of soft handover where the radio links that are added and removed belong to the same Node B.



CDMA soft handover

MULTIPATH AND RAKE RECEIVERS

One of the main advantages of CDMA systems is the capability of using signals that arrive in the receivers with different time delays. This phenomenon is called multipath. FDMA and TDMA, which are narrow band systems, cannot discriminate between the multipath arrivals, and resort to equalization to mitigate the negative effects of multipath. Due to its wide bandwidth and rake receivers, CDMA uses the multipath signals and combines them to make an even stronger signal at the receivers. CDMA subscriber units use rake receivers. This is essentially a set of several receivers. One of the receivers (fingers) constantly searches for different multipaths and feeds the information to the other three fingers. Each finger then demodulates the signal corresponding to a strong multipath. The results are then combined together to make the signal stronger.

Chapter 5

Fixed wireless networks and Wireless loops

5.1 Cordless Systems

5.2 WLL

5.3 IEEE 802.16

5.1 Cordless Systems

A communication system comprising:

1. cordless system units connected to a communication network each of said cordless system units respectively comprising at least one cordless base station and a check apparatus; a respective cordless terminal equipment having an identity that is stored in one of said cordless system units allocated to said cordless terminal equipment as a home system unit, and whereby an allowance of a connection setup from the respective cordless system unit to a cordless terminal equipment being checked by a check apparatus respectively allocated to the cordless system units; and a network offering a connectionless service for bidirectionally exchanging information between said cordless system units via said network, said cordless system units being connected to said network, and said information being exchanged dependent on check results of one of said check apparatuses.

2. The communication system according to Point 1, wherein the network offering a connectionless service is a datagram-oriented network.

3. The communication system according to Point 2, wherein the datagram-oriented network is exclusively provided for the information exchange between cordless system units.

4. The communication system according to Point 2, wherein the datagram-oriented network is a local network.

5. The communication system according to Point 2, wherein the datagram-oriented network is composed of at least two coupled local networks.

6. The communication system according to Point 2, wherein the datagram-oriented network has a single controller for monitoring access authorization of connected cordless system units onto said network and for rejecting unauthorized access attempts.

7. The communication system according to Point 6, wherein the datagram-oriented network has a central local network to which at least one of remote cordless system units and remote local networks of the datagram-oriented network are respectively coupled via a permanent point-to-point connection.

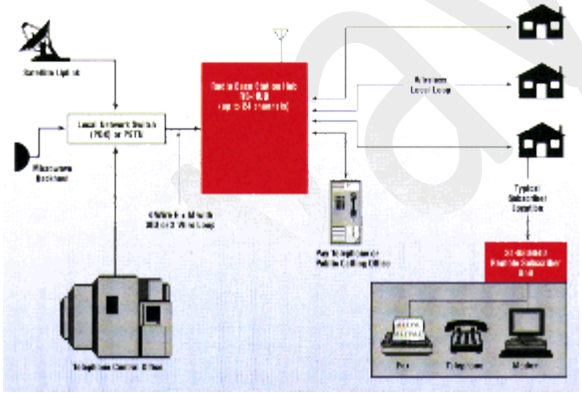
8. The communication system according to Point 3, wherein the datagram-oriented network is a local network.

9. The communication system according to Point 3, wherein the datagram-oriented network is composed of at least two coupled local networks.

10. The communication system according to Point 3, wherein the datagram-oriented network has a single controller for monitoring access authorization of connected cordless system units onto said network and for rejecting unauthorized access attempts.

11. The communication system according to Point 10, wherein the datagram-oriented network has a central local network to which at least one of remote cordless system units and remote local networks of the data-oriented network are respectively coupled via a permanent point-to-point connection.

5.2 WLL



Wireless Local Loop is an ideal application to provide telephone service to a remote rural area.

The system is based on a full-duplex radio network that provides local telephone-like service among a group of users in remote areas. These areas could be connected via radio links to the national telephone network, though allowing the WLL subscriber to call or be reached by any telephone in the world.

The WLL unit consists of a radio transceiver and the WLL interface assembled in one metal box. Two cables and a telephone connector are the only outlets from the box; one cable connects to a Yagi directional antenna and a phone receptacle to connect to a common telephone set. A fax or modem could also be connected for fax or computer communication.

The WLL is an enhancement of the SmarTrunk II system and for more information on the base/repeater station to operate the WLL, please review SmarTrunk under the product listing.

WLL solutions are particularly popular in remote or sparsely populated areas of developing countries, where cabled infrastructure is either too expensive to deploy or where speed of deployment is an issue.

Local Loops

The local loop is the physical link, or circuit, that connects the customer premises to the edge of the carrier, or service provider, network. Traditionally, the local loop was wireline in nature, specifically in the form an electrical circuit (i.e., loop) provisioned over UTP (Unshielded Twisted Pair) in support of voice communications.

Although it remains unusual in all but the most demanding, bandwidth-intensive applications, optical fiber local loops are a wireline alternative. The local loop takes the form of a trunk if it connects to a premises-based switching device such as a voice PBX (Private Branch Exchange), or a data switch or router.

Technically, a trunk connects switches, and the device at the edge of the service provider network generally is assumed to be a switch of some sort. The local loop takes the form of a line if it connects to a premises-based device other than a switch, with examples being a KTS (Key Telephone System), a single-line or multiline telephone set or a computer modem.

At the edge of the carrier network in a traditional PSTN (Public Switched Telephone Network) scenario, the local loop terminates in a circuit switch housed in an ILEC (Incumbent Local Exchange Carrier) CO (Central Office).

In another, more contemporary scenario, the local loop may terminate in a circuit switch owned by a CLEC (Competitive LEC) and housed in a POP (Point Of Presence), which typically is either an ILEC CO or a carrier hotel. Increasingly, the local loop may be provisioned in support of data communications applications, or combined voice and data.

The Electromagnetic Spectrum

The electromagnetic spectrum includes electricity - radio and light - all of which travel in the form of waves, or cycles. The frequency of the signal refers to the frequency of the sinusoidal waveforms, or sine waves.

The inverse of frequency is wavelength, which refers to the distance between the peaks or troughs of the waveforms, and which is the inverse of frequency.

Frequency generally is used in reference to electrical and radio signals, and is measured in Hz (Hertz). Radio signals are in the range from 3 KHz to 300 GHz. Wavelength generally is used to describe signals in the optical domain, where it is measured in mu (microns).

The higher the frequency of the signal, the greater the number of Hz that exist per unit of time (i.e., second), and the more raw bandwidth available to represent data. The specifics of the modulation technique determine the actual amount of data that can be impressed on the signal. So, there are clear advantages to high frequency radio signals.

On the downside, however, high frequency signals suffer to a greater extent from signal attenuation (i.e., loss of signal strength). The impact of signal attenuation dictates the extent to which the system will tolerate physical obstructions such as windows, doors and walls.

Therefore, line-of-sight is always desirable, and is required at the higher frequencies. Rain attenuation, or rain fade, is the term used to describe the phenomenon by which radio signals are affected by precipitation.

Not only rain, but also sleet, snow and hail can impact radio signals to a considerable extent. Fog and even high humidity have an effect, as do smog, agricultural haze and other environmental considerations. For that matter, radio signals suffer from the physical matter in the air, even on the clearest, purest day.

Therefore, the distance between the transmitter and the receivers is limited, sensitive to the signal strength (i.e., power level). In the vacuum of space, attenuation is not an issue, so we can receive clear electromagnetic signals from sources that are (or were) billions of light years away.

RF Issues

RF services generally are provided on the basis of licensed frequencies. This approach of licensing by the FCC (or other national or regional regulatory authority) provides assurances that the signal will be free from interference, as only a single entity is authorized to use a given frequency or frequency range within a given geographical area.

Increasingly, however, unlicensed frequencies are used with various access techniques, signal modulation methods and other mechanisms to mitigate issues of interference between competing signals. The unlicensed frequencies, which typically are in the ISM (Industrial, Scientific and Medical) bands, offer the advantages of no licensing cost and no licensing delays, both of which can be quite considerable.

A definite RF negative is the lack of security. Anyone with an antenna in proximity and tuned to the right frequency can capture the raw signal. Frequency hopping, a spread spectrum technique used in some networks, makes it extremely difficult to capture the signal.

However, the only real protection against security breaches involves authentication and encryption, and they're not all created equal.

Cellular Telephony

Cellular telephony has achieved WLL status, depending on how you want to define it. The statistics suggest that wireline local loops actually decreased in number in the U.S. in 2002. We can't blame it all on cellular growth, since DSL and CATV networks are at least partly to blame, but an increasing number of people have completely forsaken wireline service in favor of cellular.

Several people have done this for various reasons including budgetary. I know a number of people in South Africa (I teach a seminar there twice a year) who have done the same thing. And that's quite remarkable, given the current state of cellular networking.

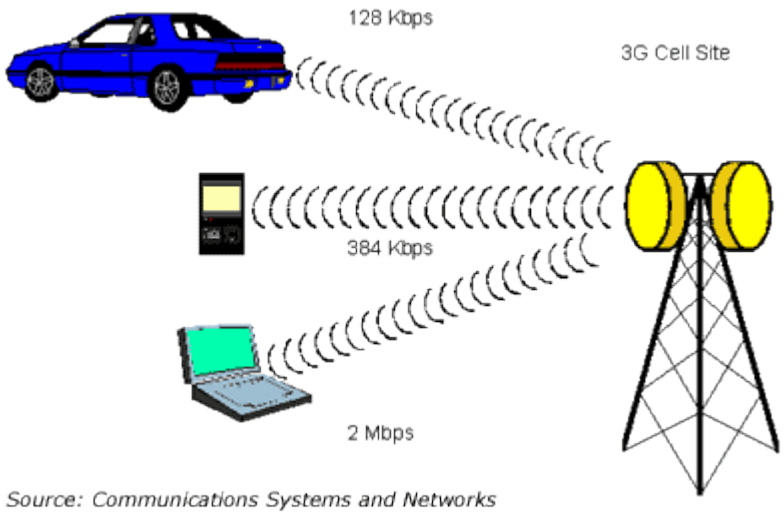
There are a number of contemporary digital cellular standards, including GSM, D-AMPS and PCS. While they are incompatible, they all share one thing in common - they are narrowband in nature.

The fixed wireless dimension of emerging 2.5G (Generation) and 3G cellular networking is quite another matter, however. All of these grew out of a failed attempt to create a single global standard, and fit under the umbrella of IMT-2000 (International Mobile Telecommunications-2000), an initiative of the ITU-T intended as an international wireless network architecture for the 21st century.

The various 3G specifications include the following speeds and intended applications:

- 128 Kbps for high mobility applications
- 384 Kbps for mobile applications at pedestrian speeds
- 2 Mbps for fixed WLL and in-building applications such as WLANs

Note: The theoretical data rates are best case hyperbole. Actual data rates usually are much lower due to factors such as EMI (ElectroMagnetic Interference), RFI (Radio Frequency Interference), signal attenuation and line-of-sight.



Source: Communications Systems and Networks

Now, an architectural umbrella is a fine thing, but it doesn't do much to resolve issues of incompatibility, which plague the contemporary wireless world. It has been said that the nice thing about wireless standards is that there are so many from which to choose.

And so it seems there will be into the foreseeable future, as well. Here's a short list of nextgen wireless standards, along with brief descriptions of them. *(Note: The data rates quoted are best case rates, and are not realistic in terms of actual throughputs in the real world.)*

HCSD (High-Speed Circuit Switched Data) is a 2+G approach that is considered to be an interim step toward 2.5G and 3G networks. HCSD improves on the current cellular data transmission rates of 9.6 Kbps by supporting packet data transmission over the circuit-switched GSM network through the linking of up to four GSM time slots at 14.4 Kbps for a total transmission rate of 57.6 Kbps. A small number of carriers have deployed HCSD.

GPRS (General Packet Radio Service) is the 2.5G enhancement to GSM. It is a packet-switched service that supports TCP/IP and X.25 packet protocols, with QoS (Quality of Service) differentiation. GPRS has been demonstrated to run at speeds up to 115 Kbps and has a theoretical transmission rate as high as 171.2 Kbps, although the actual data rate is generally much less. The applications include e-mail and mobile Internet browsing. In the U.S. AT&T Wireless, Cingular and T-Mobile are deploying GPRS networks.

EDGE (Enhanced Data rates for Global Evolution) is a 2.5G standard touted as the final stage in the GSM evolution in Europe. EDGE also is capable of running over IS-136 D-AMPS (Digital-Advanced Mobile Phone System) networks in the U.S. EDGE is an intermediate step in the evolution to 3G WCDMA (Wideband CDMA), although some carriers are expected to stop short of that final step. EDGE is planned to support data transmission at rates up to 384 Kbps. AT&T Wireless, Cingular and T-Mobile have made commitments to EDGE at various levels.

UMTS (Universal Mobile Telecommunications System), also known as ***WCDMA (Wideband Code Division Multiple Access)***, is a 3G technology intended to support data transmission rates

of 128 Kbps for high-mobility applications, 384 Kbps at pedestrian mobility speeds and 2 Mbps for fixed wireless applications. AT&T Wireless, Cingular and T-Mobile all have announced plans to deploy UMTS.

CDMA2000 (Code Division Multiple Access 2000), also known as ***IS-856***, is a 3G technology based on earlier versions of CDMA. The initial version, known as 1xRTT (1 times Radio Transmission Technology), is a 2.5G technology initially offering data speeds up to 153 Kbps, with throughput in the range of as much as 90 Kbps. The enhanced version, known as 1xEV-DO (1 times EVolution-Data Optimized), is an asymmetric technology offering peak data rates of up to 2.4 Mbps on the forward link and 153 Kbps on the reverse link. GSM1x is a version intended as a transition specification for GSM operators. Nextel, Sprint PCS and Verizon Wireless all have made commitments to CDMA2000.

As should now be abundantly clear, the confusion over cellular standards is not likely to abate anytime soon. Multimode terminal equipment resolves these issues of incompatibility in some cases.

Cingular Wireless, for example, has introduced a handset that allows its subscribers to roam freely between its GSM and TDMA networks. The AT&T multi-band phone will operate on GSM/GPRS networks in countries where AT&T has roaming agreements in place.

5.3 IEEE 802.16

In recent years there has been increasing interest shown in wireless technologies for subscriber access, as an alternative to traditional twisted-pair local loop.

These approaches are generally referred to as wireless local loop (WLL), or fixed-wireless access. To provide a standardized approach to WLL, the IEEE 802 committee set up the 802.16 working group in 1999 to develop broadband wireless standards.

IEEE 802.16 standardizes the air interface and related functions associated with WLL. Three working groups have been chartered to produce standards:

- IEEE 802.16.1 - Air interface for 10 to 66 GHz.
- IEEE 802.16.2 - Coexistence of broadband wireless access systems.
- IEEE 802.16.3 - Air interface for licensed frequencies, 2 to 11 GHz.

The work of 802.16.1 is the farthest along, and it's likely that it will generate the most interest in the industry, as it is targeted at available frequency bands.

An 802.16 wireless service provides a communications path between a subscriber site and a core network (the network to which 802.16 is providing access). Examples of a core network are the public telephone network and the Internet. IEEE 802.16 standards are concerned with the air interface between a subscriber's transceiver station and a base transceiver station.

Protocols defined specifically for wireless transmission address issues related to the transmission of blocks of data over a network. The standards are organized into a three-layer architecture.

☐ The lowest layer, the physical layer, specifies the frequency band, the modulation scheme, error-correction techniques, synchronization between transmitter and receiver, data rate and the time-division multiplexing (TDM) structure.

For transmission from subscribers to a base station, the standard uses the Demand Assignment Multiple Access-Time Division Multiple Access (DAMA-TDMA) technique. DAMA is a capacity assignment technique that adapts as needed to respond to demand changes among multiple stations. TDMA is the technique of dividing time on a channel into a sequence of frames, each consisting of a number of slots, and allocating one or more slots per frame to form a logical channel.

With DAMA-TDMA, the assignment of slots to channels varies dynamically. For transmission from a base station to subscribers, the standard specifies two modes of operation, one targeted to support a continuous transmission stream (mode A), such as audio or video, and one targeted to support a burst transmission stream (mode B), such as IP-based traffic. Both are TDM schemes.

☐ Above the physical layer are the functions associated with providing service to subscribers. These functions include transmitting data in frames and controlling access to the shared wireless medium, and are grouped into a media access control (MAC) layer. The MAC protocol defines how and when a base station or subscriber station may initiate transmission on the channel. Because some of the layers above the MAC layer, such as ATM, require quality of service, the MAC protocol must be able to allocate radio channel capacity to satisfy service demands.

In the downstream direction (base station to subscriber stations), there is only one transmitter, and the MAC protocol is relatively simple. In the upstream direction, multiple subscriber stations compete for access, resulting in a more complex MAC protocol. In both directions, a TDMA technique is used, in which the datastream is divided into a number of time slots.

The sequence of time slots across multiple TDMA frames that is dedicated to one subscriber forms a logical channel, and MAC frames are transmitted over that logical channel. IEEE 801.16.1 is intended to support individual channel data rates of from 2M to 155M bit/sec.

☐ Above the MAC layer is a convergence layer that provides functions specific to the service being provided. For IEEE 802.16.1, bearer services include digital audio/video multicast, digital telephony, ATM, Internet access, wireless trunks in telephone networks and frame relay.

Chapter 6

6.1 WiMAX

6.2 Rural wireless network

6.3 VSAT

6.4 Cellular Wireless Network – GPRS, 2.5G, 3G-WCDMA

6.1. WiMAX:

What is WiMAX?

WiMAX combines the familiarity of Wi-Fi with the mobility of cellular that will deliver personal mobile broadband that moves with you. It will let you get connected to the Internet, miles from the nearest Wi-Fi hotspot. Soon, Mobile WiMAX will blanket large areas—metropolitan, suburban, or rural—delivering mobile broadband Internet access at speeds similar to existing broadband. WiMAX is built for the future with advanced, efficient wireless technology that provides higher speeds than today's wide area wireless technologies. It will be able to completely transform your mobile Internet lifestyle, enabling you to connect in ways you've only dreamed about.

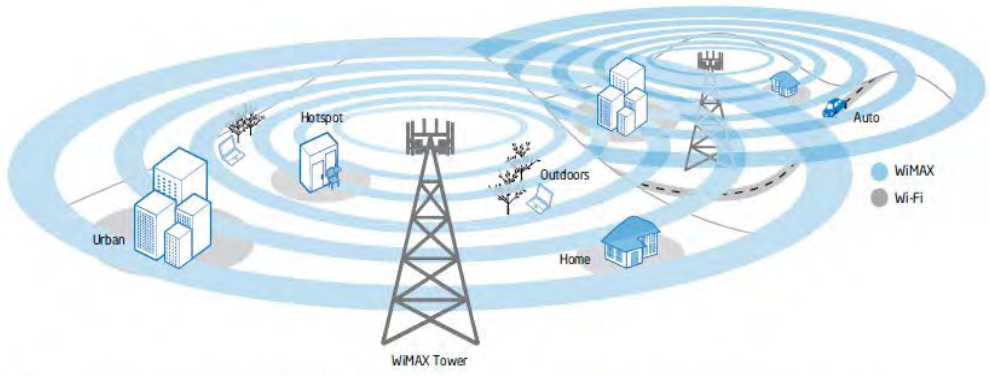


Figure 1. *WiMAX blankets large areas with broadband Internet.*

How Does WiMAX Work?

Think of WiMAX as taking the best part of cellular network access– the part that allows you to easily connect anywhere within your service provider's wide coverage area and taking the best part of your Wi-Fi experience—the fast speeds and a familiar broadband Internet experience. And combining them into a new wireless standard. Expanding Your Wireless World. You'll be able to get WiMAX as a subscription or pay-as-you-go service that lets you take your broadband with you, similar to the way you receive mobile phone service. WiMAX is a Wide

Area Network (WAN) technology. Service providers will deploy a network of towers that will enable access over many miles. Internet access is instantly available anywhere within coverage areas. And like Wi-Fi, WiMAX is a standards-based technology that will unleash the benefits of open markets and global economies of scale to deliver the devices and services that consumers want.

Why WiMAX?

WiMAX success stems from a robust vision incorporating four key strengths:

- Open standards-based, interoperable technology built from the ground up for the Internet fosters innovation and competition.
- Vibrant, growing ecosystem of industry leaders such as Intel, Sprint, Clearwire, Motorola, Samsung, Nokia, Cisco, and hundreds of other companies.
- Global economies of scale and more attractive intellectual property environment that enable lower costs compared to other wireless technologies.
- Advanced wireless technology that enables a faster wireless broadband solution for doing more on the go.

Intel's Role in WiMAX. Since very early in the development stages, Intel has played a large role in creating the WiMAX standard. The company has made significant capital investments worldwide to help meet regulatory and deployment challenges, in addition to the development of the technologies themselves.

Built for the Internet
WiMAX is built from the ground up for Internet applications, services, and security, with architecture specifically designed to seamlessly extend the Internet to mobile users. Because WiMAX is built on Internet protocol (IP) networking, it supports all the latest IP security and quality of service (QoS) standards. WiMAX support of QoS standards enables real-time media like Voice over IP (VoIP) and streaming video.

Advanced Wireless Security. WiMAX supports advanced security features to protect information as it travels across the airwaves. For example, it supports AES (AdvancedEncryption Standard), a state-of-the-art security technology that encrypts data as it passes between the client and the base station.

How WiMAX Works
So how does WiMAX transmit the Internet over the landscape? The WiMAX network uses an approach that is similar to that of cell phones. Coverage for a geographical area is divided into a series of overlapping areas called cells. Each cell provides coverage for users within that immediate vicinity. When you travel from one cell to another, the wireless connection is handed off from one cell to another. The WiMAX standard supports mobile, portable, and fixed service options. This enables wireless providers to offer broadband Internet access to areas underserved by telephone and cable companies. For fixed WiMAX deployments, service providers supply Customer Premises Equipment (CPE) that acts as a wireless "modem" to provide the interface to the WiMAX network for a specific location, such as a home, cafe, or office. WiMAX is also well

suited for emerging markets as a cost-effective way to deliver high-speed Internet. The WiMAX network includes two key components: a base station and a subscriber device. The WiMAX base station is mounted on a tower or tall building to broadcast the wireless signal. The subscriber receives the signals on a WiMAX enabled notebook, mobile Internet device (MID), or even a WiMAX modem.
Intel Enables Adoption. Intel is enabling mass market adoption of WiMAX in notebooks similar to the way it enabled Wi-Fi in notebooks. Intel has taken advantage of synergies between Wi-Fi and WiMAX with an integrated Wi-Fi/WiMAX solution that minimizes space used within portable devices.

WiMax Architecture

WiMax offers a rich feature set and flexibility, which also increases the complexity of service deployment and provisioning for fixed and mobile networks.

Let us take a look at the WiMax Management Information Base (MIB).

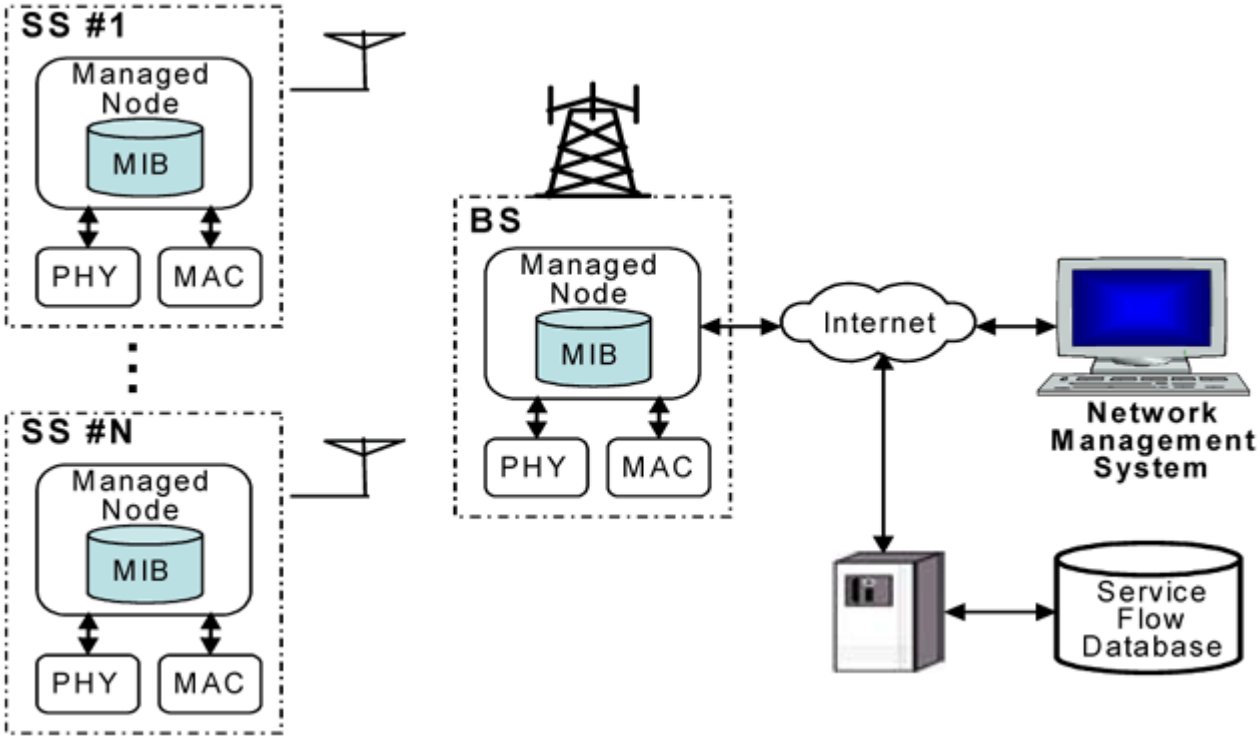


Figure 2: WiMax Management Information Base

The above figure shows the management reference model for BWA (Broadband Wireless Access) networks. This consists of a Network Management System (NMS), some nodes, and a database. BS and SS managed nodes collect and store the managed objects in an 802.16 MIB format. Managed objects are made available to NMS' using the Simple Network Management Protocol (SNMP).

When a customer subscribes to the WiMax service, the service provider asks the customer for the service flow information. This would include number of UL / DL connections with the data rates and QoS parameters. The customer also needs to tell the kind of applications that he proposes to run.

The service provider then proceeds to pre-provision the services and enters the information in the Service Flow Database.

WiMax Competing Technologies

As we learnt in the previous pages, that WiMax was formed to complement Bluetooth and WiFi technologies, let us look at the differences in each of these:

Parameters	**WiMax**	**WLAN**	**Bluetooth**
Frequency	2 ~V 11 GHz	2.4 GHz	Varies
Range	~31 miles	~100 metres	~10 metres
Data Transfer Rate	70mbps	11 ~V 55mbps	20 ~V 55mbps
Users	1000s	> 10	> 10

Table 1 ~V Difference between WiMax, WLAN and Bluetooth

The 802.11 is based on a distributed architecture, whereas, WiMax is based on a centrally controlled architecture. In this the scheduler residing in the Base station (BS) has the complete control of the wireless media access. WiMax can support multiple connections conforming to a set of QoS parameters and provides the packet classifier ability to map the connections to many user applications and interfaces.

WiFi and WiMax may end up complementing each other, but a new technology, IEEE 802.20, might give them both a run for their money. IEEE 802.20 standard like the 802.16 is aimed at wireless high-speed connectivity to mobile consumer devices like cellular phones, PDAs and laptops. It will operate in the 500 MHz ~V 3.5 GHz and is led by Flarion Technologies and ArrayComm.

6.2. Rural Wireless Network

Voice telephony has been the main option for providing access to telecommunications in rural areas. Today, a wide variety of new applications such as e-mail, e-commerce, tele-education, tele-health, and tele-medicine, among others, has made access to interactive multimedia services

as important as - maybe even more important than - voice connectivity alone. Since each rural district or community requires a different mix of voice, text, image, video and audio communications to best meet its needs, telecommunication network operators must be able to support the widest possible range of services and/or applications and different bandwidth levels at a reasonable cost.

The Internet (with the unavailability of IP network in rural area) is the most widely used platform used to deliver multimedia applications in rural areas of developing countries. Satellite broadcasting has also been widely adopted in distance education programs and other videoconferencing-based consultations in remote areas. These two platforms are expected to converge as Internet broadcasting and satellite-based Internet links continue to be developed. While much negative attention in developing countries has been focused on the use of the Internet as an illegal bypass mechanism in the international traffic arena, the long-term importance of the Internet for developing countries lies in its potential to improve the domestic flow of economic and educational resources between isolated rural communities and urban centers, until such technology IP networks are provided to the rural areas.

The following are basic requirements for communication systems deployed in rural areas of developing countries:

- Implementation and operation is possible at a low cost in areas where population density is low;

- The system can be easily installed, even in remote and inaccessible locations;

- System operation and maintenance may be carried out even where qualified technical personnel are scarce;

- Implementation is possible even when basic infrastructure such as mains: electricity, running water, paved road networks, etc., are absent.

An increasing number of technologies are available that can meet the above requirements at a reasonable cost to rural network operators.

6.3 VSAT

VSAT is an abbreviation for a Very Small Aperture Terminal. It is basically a two-way satellite ground station with a less than 3 meters tall (most of them are about 0.75 m to 1.2 m tall) dish antenna stationed. The transmission rates of VSATs are usually from very low and up to 4 Mbit/s. These VSATs' primary job is accessing the satellites in the geosynchronous orbit and relaying data from terminals in earth to other terminals and hubs. They will often transmit narrowband data, such as the transactions of credit cards, polling, RFID (radio frequency identification) data, and SCADA (Supervisory Control and Data Acquisition), or broadband data, such as satellite Internet, VoIP, and videos. However, the VSAT technology is also used for various types of communications.

Equatorial Communications first used the spread spectrum technology to commercialize the VSATs, which were at the time C band (6 GHz) receive only systems. This commercialization led to over 30,000 sales of the 60 cm antenna systems in the early 1980s. Equatorial Communications sold about 10,000 more units from 1984 to 1985 by developing a C band (4 and 6 GHz) two way system with 1 m x 0.5 m dimensions.

In 1985, the current world's most used VSATs, the Ku band (12 to 14 GHz) was co-developed by Schlumberger Oilfield Research and Hughes Aerospace. It is primarily used to provide portable network connection for exploration units, particularly doing oil field drilling.

Implementations of VSAT

Currently, the largest VSAT network consists of over 12,000 sites and is administered by Spacenet and MCI for the US Postal Service (USPS). Walgreens Pharmacy, Dollar General, CVS, Riteaid, Wal-Mart, Yum! Brands (such as Taco Bell, Pizza Hut, Long John Silver's, and other fast food chains), GTEC, SGI, and Intralot also utilizes large VSAT networks. Many huge car corporations such as Ford and General Motors also utilizes the VSAT technology, such as transmitting and receiving sales figures and orders, along with announcing international communications, service bulletins, and for distance learning courses. An example of this is the "FordStar Network."

Two way satellite Internet providers also use the VSAT technology. Companies like StarBand, WildBlue, and HughesNet in the United States and SatLynx, Bluestream, and Technologie Satelitarne in Europe, and many other broadband services around the world in rural areas where high speed Internet connections cannot be provided use it too. A statistic from December 2004 showed that over a million VSATs were in place.

VSAT Configurations

Most of the current VSAT networks use a topology:

Star topology: This topology uses a central uplink site (eg. Network operations center (NOC)), which transports the data to and from each of the VSAT terminals using satellites

Mesh topology: In this configuration, each VSAT terminal will relay data over to another terminal through the satellite, acting as a hub, which also minimizes the need for an uplink site

Star + Mesh topology: This combination can be achieved (as some VSAT networks do) by having multiple centralized uplink sites connected together in a multi-star topology which is in a bigger mesh topology. This topology does not cost so much in maintaining the network while also lessening the amount of data that needs to be relayed through one or more central uplink sites in the network.

VSAT's Strengths

VSAT technology has many advantages, which is the reason why it is used so widely today. One is availability. The service can basically be deployed anywhere around the world. Also, the VSAT is diverse in that it offers a completely independent wireless link from the local infrastructure, which is a good backup for potential disasters. Its deployability is also quite amazing as the VSAT services can be setup in a matter of minutes. The strength and the speed of the VSAT connection being homogenous anywhere within the boundaries is also a big plus. Not to forget, the connection is quite secure as they ar private layer-2 networks over the air. The pricing is also affordable, as the networks themselves do not have to pay a lot, as the broadcast download scheme (eg. DVB-S) allows them to serve the same content to thousands of locations at once without any additional costs. Last but not least, most of the VSAT systems today use onboard acceleration of protocols (eg. TCP, HTTP), which allows them to delivery high quality connections regardless of the latency.

VSAT's Drawbacks

As with everything, VSAT also has its downsides. Firstly, because the VSAT technology utilizes the satellites in geosynchronous orbit, it takes a minimum latency of about 500 milliseconds every trip around. Therefore, it is not the ideal technology to use with protocols that require a constant back and forth transmission, such as online games. Also, surprisingly, the environment can play a role in slowing down the VSATs. Although not as bad as one way TV systems like DirecTV and DISH Network, the VSAT still can have a dim signal, as it still relies on the antenna size, the transmitter's power, and the frequency band. Last but not least, although not that big of a concern, installation can be a problem as VSAT services require an outdoor antenna that has a clear view of the sky. An awkward roof, such as with skyscraper designs, can become problematic.

6.4 Cellular Wireless Network – GPRS, 2.5G, 3G-WCDMA

A **cellular network** is a radio network distributed over land areas called cells, each served by at least one fixed-location transceiver known as a cell site or base station. When joined together these cells provide radio coverage over a wide geographic area. This enables a large number of portable transceivers (e.g., mobile phones, pagers, etc.) to communicate with each other and with fixed transceivers and telephones anywhere in the network, via base stations, even if some of the transceivers are moving through more than one cell during transmission.

Cellular networks offer a number of advantages over alternative solutions:

- increased capacity
- reduced power use

- larger coverage area
- reduced interference from other signals

WCDMA

WCDMA (Wideband Code Division Multiple Access) is the radio access scheme used for third generation cellular systems that are being rolled out in various parts of the globe. The 3G systems to support wideband services like high-speed Internet access, video and high quality image transmission with the same quality as the fixed networks. In WCDMA systems the CDMA air interface is combined with GSM based networks. The WCDMA standard was evolved through the Third Generation Partnership Project (3GPP) which aims to ensure interoperability between different 3G networks.

The standard that has emerged through this partnership project is based on ETSI's Universal Mobile Telecommunication System (UMTS) and is commonly known as UMTS Terrestrial Radio Access (UTRA). The access scheme for UTRA is Direct Sequence Code Division Multiple Access (DS-CDMA). The information is spread over a band of approximately 5 MHz. This wide bandwidth has given rise to the name Wideband CDMA or WCDMA.

In WCDMA, there are two different modes of operation possible:

TDD: In this duplex method, uplink and downlink transmissions are carried over the same frequency band by using synchronized time intervals. Thus time slots in a physical channel are divided into transmission and reception part.

FDD: The uplink and downlink transmissions employ two separated frequency bands for this duplex method. A pair of frequency bands with specified separation is assigned for a connection. Since different regions have different frequency allocation schemes, the capability to operate in either FDD or TDD mode allows for efficient utilization of the available spectrum

Key Features of WCDMA

The key operational features of the WCDMA radio interface are listed below:

- Support of high data rate transmission: 384 Kbps with wide area coverage, 2 Mbps with local coverage.

- High service flexibility: support of multiple parallel variable rate services on each connection.
- Both Frequency Division Duplex (FDD) and Time Division Duplex (TDD).
- Built in support for future capacity and coverage enhancing technologies like adaptive antennas, advanced receiver structures and transmitter diversity.
- Support of inter frequency hand over and hand over to other systems, including hand over to GSM.
- Efficient packet access.

GPRS

Introduction

General Packet Radio Service (GPRS) is a 2.5 generation packet based network technology for GSM networks.

Data Speed

GPRS data speeds are expected to reach theoretical data speeds of up to 171.2 Kbps. However, this is based on optimal conditions in terms of available cell/sector capacity in terms of available time slots, maximum coding scheme (CS-4) as well as moible phone availability to support the maximum number of time slots - eight. More practical data rates are currently in the order of 40-60 Kbps.

3G technologies such as W-CDMA will theoreticaly provide up to 2 Mbps in a fixed location. There will, however, be some significant limitations to this theoretical capacity. While 3G (and beyond) is expected to usher in the advent of high-bandwidth, multi-media services, the real impetus for

2.5G and packet based mobile data lies elsewhere.

Impetus for GPRS

The major impetus for GPRS and other packet based mobile data technologies is the "always-on" capability. Being packet based, GPRS allows for the use of infrastructure and facilities only when a transaction is required, rather than maintaining facilities in a session-like manner. This provides tremendous infrastructure efficiency and service delivery improvements.

Using GPRS as a bearer for WAP, for instance, will allow for the use of WAP on a per-transaction rather than a per-minute-of-use basis. More importantly perhaps is the ability for GPRS to allow for autonomous service realization through the always-on capability. For example, a GPRS customer could receive content or services without actually manually invoking

a service or transaction. This has significant implications for mobile commerce and location based services.

GPRS Architecture and Issues
GPRS architecture consists of Gateway GPRS Support Node (GGSN) and a Serving GPRS Support Node (SGSN). The GGSN acts as the gateway to other packet data networks such as the Internet. The SGSN is the serving node that enables virtual connections to the GPRS enabled mobile device and delivery of data.

The blessing and curse of the SGSN is that it supports an attach state when a user is engaged in GPRS data usage and a detach state when idle. The idle state creates a particular challenge for attempting to position the unit for location based services. In addition, GPRS presents a challenge in terms of the ability to offer prepaid mobile data services, which may be overcome by the introduction of CAMEL and perhaps the use of Parlay.

The evolution from GPRS to W-CDMA entails upgrade of the Radio Access Network (RAN) to include two new network elements. The Node B replaces the BTS and the Radio Network Controller (RNC) replaces the BSC in the RAN. However, mobile network operators will maintain their GPRS assets for that service and thus maintain the existing network elements along with the new ones for 3G. W-CDMA continues to use the same Core Network (CN) elements as GPRS.

Deployment and Operational Issues

Beyond the scope of this white paper, there are several significant issues associated with deployment and operation of GPRS systems. Those issues include:

- Capacity and network optimization
- Handset availability and performance
- Quality of service
- Charging for services

2.5G Mobile Systems

The move into the 2.5G world will begin with General Packet Radio Service (GPRS). GPRS is a radio technology for GSM networks that adds packet-switching protocols, shorter setup time for ISP connections, and the possibility to charge by the amount of data sent, rather than connection time. Packet switching is a technique whereby the information (voice or data) to be sent is broken up into packets, of at most a few Kbytes each, which are then routed by the network between different destinations based on addressing data within each packet. Use of network resources is optimized as the resources are needed only during the handling of each packet.

The next generation of data heading towards third generation and personal multimedia environments builds on GPRS and is known as Enhanced Data rate for GSM Evolution (EDGE).

EDGE will also be a significant contributor in 2.5G. It will allow GSM operators to use existing GSM radio bands to offer wireless multimedia IP-based services and applications at theoretical maximum speeds of 384 kbps with a bit-rate of 48 kbps per timeslot and up to 69.2 kbps per timeslot in good radio conditions. EDGE will let operators function without a 3G license and compete with 3G networks offering similar data services. Implementing EDGE will be relatively painless and will require relatively small changes to network hardware and software as it uses the same TDMA (Time Division Multiple Access) frame structure, logic channel and 200 kHz carrier bandwidth as today's GSM networks. As EDGE progresses to coexistence with 3G WCDMA, data rates of up to ATM-like speeds of 2 Mbps could be available.

GPRS will support flexible data transmission rates as well as continuous connection to the network. GPRS is the most significant step towards 3G

Third-Generation Mobile Systems
Third generation (3G) networks were conceived from the Universal MobileTelecommunications Service (UMTS) concept for high speed networks for enabling a variety of data intensive applications. 3G systems consist of the two main standards, CDMA2000 and W-CDMA, as well as other 3G variants such as NTT DoCoMo's Freedom of Mobile Multimedia Access (FOMA) and Time Division Synchronous Code Division Multiple Access (TD-SCDMA) used primarily in China.

Data Speed

The data speed of 3G is determined based on a combination of factors including the chip rate, channel structure, power control, and synchronization.

An example of calculating the theoretical 3G data speed is as follows:

- W-CDMA assigned code 400-500 Kpbs/code. 6 codes X 400 > 2Mbps (UMTS target for 3G data speed in fixed location)

Actual data speeds will vary in accordance with several factors including:

- Number of users in cell/sector
- Distance of user from cell
- User is moving or stationary
- Network operator capacity and network optimization requirements

1xEV-DO is a data-only solution, supporting a theoretical data speed of up to 2.457 Mbps

1xEV-DV is a data and voice solution, supporting a theoretical data speed of up to 3.072 Mbps

FOMA has two operational modes, supporting a dedicated 64 Kbps connection or a 384 Kbps downlink/64 Kbps uplink best-effort connection.

TD-SCDMA can operate in 1.6 MHz or 5 MHz mode for 2 Mbps or 6 Mpbs respectively

Comparison of W-CDMA to CDMA2000

Both use a coding scheme that separates each subscriber from other subscribers

Both use control channels to manage the network

W-CDMA and CDMA2000 are not compatible from the perspective that they have different chip rates - 3.84 MCPS for W-CDMA vs. 1.2888 MCPS for CMDA2000. W-CDMA uses a 5 MHz channel. Initially, CDMA2000 uses only a 1.25 MHz channel, but with CDMA2000 3x, three 1.25 MHz channels can be combined to form a super channel structure.

W-CDMA is synchronous, relying on mobile station time measurements between two base stations, rather than using GPS as CDMA2000 does.

There are three modes of operation for W-CDMA/CDMA2000:

- Direct Sequence (DS) W-CDMA (UMTS) for Frequency Divsion Duplex (FDD)
- W-CDMA Time Division Duplex (TDD)
- CDMA2000 Multi-carrier FDD

Each of the three radio interface methods may be employed on either a GSM or ANSI-based Core Network (CN).

IS-833 is a standard, developed by the 3GPP2, to support CDMA2000 1xRTT Radio Access Network (RAN) to interface with a GSM CN. RAN upgrade required includes CDMA base station and BSC. CN upgrade required includes CDMA PDSN and AAA server.

Impetus for 3G

The major impetus for 3G is to provide for faster data speed for data-intensive applications such as video. In addition, 3G to providing faster data speeds on a per-user basis, 3G is also helpful to provide greater overall capacity for voice and data users. For example, NTT DoCoMo's plan to migrate iMode users from the 2G PDC network to FOM is driven by overall capacity concerns, as apposed to individual user data speed requirement.

3G Architecture

W-CDMA uses the same CN as GPRS, utilizing existing infrastruture such as the GGSN and SGSN. W-CDMA, however, does require new RAN infrastructure such as the Node B, which replaces the BTS, and the Radio Network Controller (RNC), which replaces the BSC.

Ultimately, the W-CDMA CN will evolve to comprise a full Mobile IP infrastructure including Media Gateway (MGW) and Media Gateway Controller (MGC) equipment for VoIP and other new equipment such as the HSS and CSDF.

CDMA2000 starts with new channel cards and then migrates to a full Mobile IP infrastructure requiring new Core Network (CN) infrastructure such as the **AAA (** Authentication, Authorization, and Accounting)server and Packet Data Server Node (PDSN).

Chatper 7

Wireless Lans

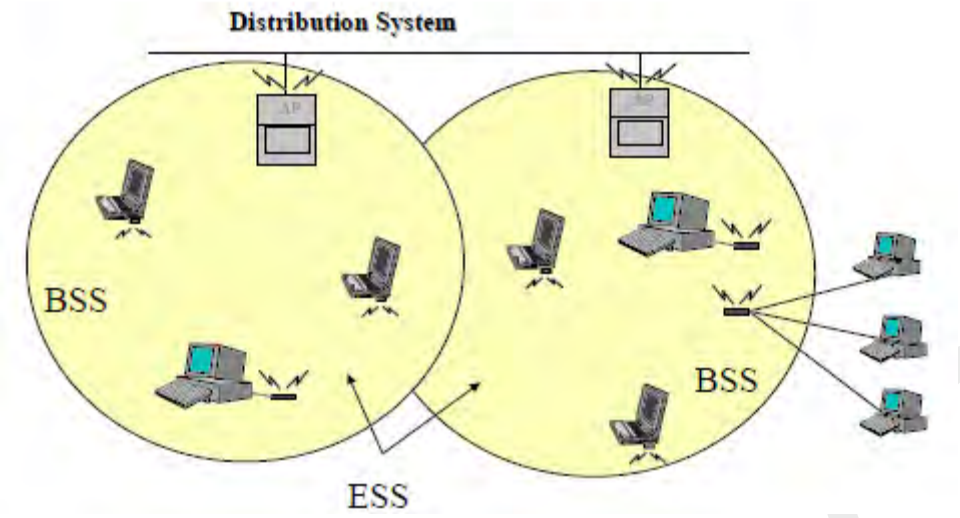
7.1 IEEE 802.11 Protocol Architecture

7.2 IEEE 802.11 Layers Description

7.3 WiFi security – WPA2

7.1 IEEE 802.11 Protocol Architecture

IEEE 802.11 Architecture
Architecture Components
An 802.11 LAN is based on a cellular architecture where the system is subdivided into cells, where each
cell (called **Basic Service Set** or **BSS**, in the 802.11 nomenclature) is controlled by a Base Station (called
Access Point, or in short **AP**).
Even though that a wireless LAN may be formed by a single cell, with a single Access Point, (and as will
be described later, it can also work without an Access Point), most installations will be formed by several
cells, where the Access Points are connected through some kind of backbone (called **Distribution System**
or **DS**), typically Ethernet, and in some cases wireless itself.
The whole interconnected Wireless LAN including the different cells, their respective Access Points and
the Distribution System, is seen to the upper layers of the OSI model, as a single 802 network, and is
called in the Standard as **Extended Service Set** (**ESS**).
The following picture shows a typical 802.11 LAN, with the components described previously:



The standard also defines the concept of a **Portal**, a Portal is a device that interconnects between an 802.11 and another 802 LAN. This concept is an abstract description of part of the functionality of a "translation bridge".
Even though the standard does not necessarily request so, typical installations will have the AP and the Portal on a single physical entity, and this is the case with BreezeCom's AP which provides both functions.

7.2 **IEEE 802.11 Layers Description**

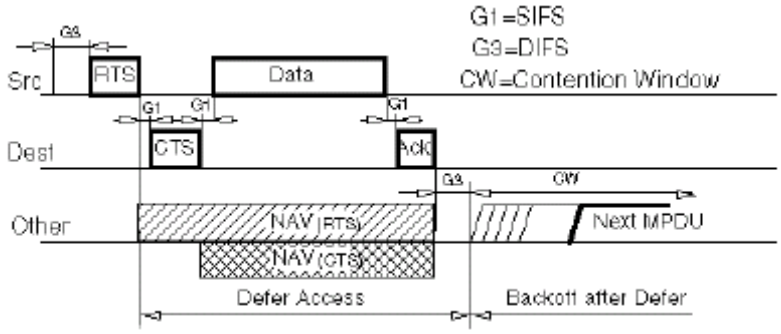
As any 802.x protocol, the 802.11 protocol covers the MAC and Physical Layer, the Standard currently
defines a single MAC which interacts with three PHYs (all of them running at 1 and 2 Mbit/s) :
· Frequency Hopping Spread Spectrum in the 2.4 GHz Band
· Direct Sequence Spread Spectrum in the 2.4 GHz Band, and
· InfraRed
802.2 Data Link
Layer
802.11 MAC
FH DS IR PHY Layer
Beyond the standard functionality usually performed by MAC Layers, the 802.11 MAC performs other
functions that are typically related to upper layer protocols, such as Fragmentation, Packet Retransmitions,
and Acknowledges.
The MAC Layer
The MAC Layer defines two different access methods, the Distributed Coordination Function and the Point Coordination Function:
The Basic Access Method: CSMA/CA
The basic access mechanism, called **Distributed Coordination Function**, is basically a Carrier Sense Multiple Access with Collision Avoidance mechanism (usually known as **CSMA/CA**).

CSMA protocols are well known in the industry, where the most popular is the Ethernet, which is a CSMA/CD protocol (CD standing for Collision Detection).
A CSMA protocol works as follows: A station desiring to transmit senses the medium, if the medium is busy (i.e. some other station is transmitting) then the station will defer its transmission to a later time, if the medium is sensed free then the station is allowed to transmit. These kind of protocols are very effective when the medium is not heavily loaded, since it allows stations to transmit with minimum delay, but there is always a chance of stations transmitting at the same time (collision), caused by the fact that the stations sensed the medium free and decided to transmit at once. These collision situations must be identified, so the MAC layer can retransmit the packet by itself and not by upper layers, which would cause significant delay. In the Ethernet case this collision is recognized by the transmitting stations which go to a retransmission phase based on an **exponential random backoff** algorithm.
While these Collision Detection mechanisms are a good idea on a wired LAN, they cannot be used on a
Wireless LAN environment, because of two main reasons:
1. Implementing a Collision Detection Mechanism would require the implementation of a Full Duplex radio, capable of transmitting and receiving at once, an approach that would increase the price significantly.
2. On a Wireless environment we cannot assume that all stations hear each other (which is the basic assumption of the Collision Detection scheme), and the fact that a station willing to transmit and senses the medium free, doesn't necessarily mean that the medium is free around the receiver area. In order to overcome these problems, the 802.11 uses a Collision Avoidance mechanism together with a
Positive Acknowledge scheme, as follows:
A station willing to transmit senses the medium, if the medium is busy then it defers. If the medium is free for a specified time (called DIFS, Distributed Inter Frame Space, in the standard) then the station is allowed to transmit, the receiving station will check the CRC of the received packet and send an acknowledgment packet (ACK). Receipt of the acknowledgment will indicate the transmitter that no collision occurred. If the sender does not receive the acknowledgment then it will retransmit the fragment until it gets acknowledged or thrown away after a given number of retransmissions.

Virtual Carrier Sense

In order to reduce the probability of two stations colliding because they cannot hear each other, the standard defines a Virtual Carrier Sense mechanism:
A station willing to transmit a packet will first transmit a short control packet called **RTS** (Request To Send), which will include the source, destination, and the duration of the following transaction (i.e. the packet and the respective **ACK**), the destination station will respond (if the medium is free) with a response control Packet called **CTS** (Clear to Send), which will include the same duration information. All stations receiving either the RTS and/or the CTS, will set their **Virtual Carrier Sense** indicator (called **NAV**, for **Network Allocation Vector**), for the given duration, and will use this information together with the Physical Carrier Sense when sensing the medium. This mechanism reduces the probability of a collision on the receiver area by a station that is "hidden" from the transmitter, to the short duration of the RTS transmission, because the station will hear the CTS and "reserve" the medium as busy until the end of the

transaction. The duration information on the RTS also protects the transmitter area from collisions during the ACK (by stations that are out of range from the acknowledging station).
It should also be noted that because of the fact that the RTS and CTS are short frames, it also reduces the overhead of collisions, since these are recognized faster than it would be recognized if the whole packet was to be transmitted, (this is true if the packet is significantly bigger than the RTS, so the standard allows for short packets to be transmitted without the RTS/CTS transaction, and this is controlled per station by a parameter called **RTSThreshold**).
The following diagrams show a transaction between two stations A and B, and the NAV setting of their neighbors:



The NAV State is combined with the physical carrier sense to indicate the busy state of the medium.

MAC Level Acknowledgments

As mentioned earlier in this document, the MAC layer performs the Collision Detection by expecting the reception of an acknowledge to any transmitted fragment (exception to these are packets that have more than one destination, such as Multicasts, which are not acknowledged).

Fragmentation and Reassembly

Typical LAN protocols use packets of several hundreds of bytes (e.g Ethernet longest packet could be up to 1518 bytes long), on a Wireless LAN environment there are some reasons why it would be preferable to use smaller packets:

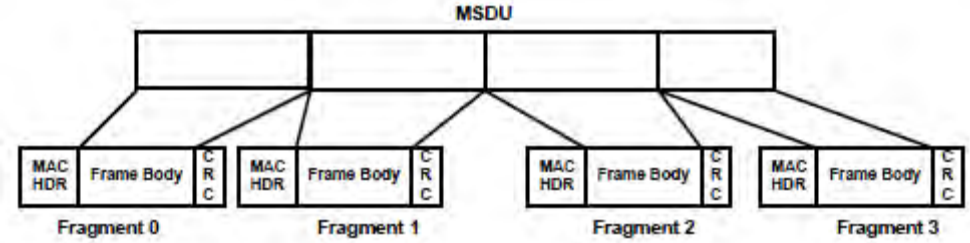
· Because of the higher Bit Error Rate of a radio link , the probability of a packet to get corrupted increases with the packet size.

· In case of packet corruption (either because of collision or noise), the smallest the packet the less overhead it causes to retransmit it.

· On a Frequency Hopping system, the medium is interrupted periodically for hopping (in our case every 20 milliseconds), so the smaller the packet, the smaller the chance that the transmission will be postponed to after the dwell time.

On the other hand, it doesn't make sense to introduce a new LAN protocol that cannot deal with packets of 1518 bytes which are used on Ethernet, so the committee decided to solve the problem by adding a simple fragmentation/reassembly mechanism at the MAC Layer.
The mechanism is a simple Send-and-Wait algorithm, where the transmitting station is not allowed to transmit a new fragment until one of the following happens:
1. Receives an ACK for the said fragment, or
2. Decides that the fragment was retransmitted too many times and drops the whole frame

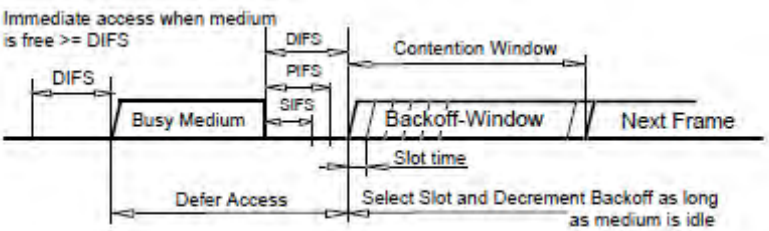
It should be noted that the standard does allow the station to transmit to a different address between retransmissions of a given fragment, this is particularly useful when an AP has several outstanding packets to different destinations and one of them does not respond.
The following diagram shows a frame (MSDU) being divided to several fragments (MPDUs):



Inter Frame Spaces
The Standard defines 4 types of Inter Frame Spaces, which are use to provide different priorities:
· **SIFS** - Which stands for **Short Inter Frame Space**, is used to separate transmissions belonging to a single dialog (e.g. Fragment-Ack), and is the minimum Inter Frame Space, and there is always atmost one single station to transmit at this given time, hence having priority over all other stations. This value is a fixed value per PHY and is calculated in such a way that the transmitting station will be able to switch back to receive mode and be capable of decoding the incomming packet, on the 802.11 FH PHY this value is set to 28 microseconds
· **PIFS - Point Cooordination IFS**, is used by the Access Point (or Point Coordinator, as called in this case), to gain access to the medium before any other station. This value is SIFS plus a Slot Time (defined in the following paragraph), i.e 78 microseconds.
· **DIFS - Distributed IFS**, is the Inter Frame Space used for a station willing to start a new transmission, which is calculated as PIFS plus one slot time, i.e. 128 microseconds and
· **EIFS - Extended IFS**, which is a longer IFS used by a station that has received a packet that could not understand, this is needed to prevent the station (who could not understand the duration information for the Virtual Carrier Sense) from colliding with a future packet belonging to the current dialog.

Exponential Backoff Algorithm
Backoff is a well known method to resolve contention between different stations willing to access the medium, the method requires each station to choose a Random Number (n) between 0 and a given number, and wait for this number of Slots before accessing the medium, always checking whether a different station has accessed the medium before. The **Slot Time** is defined in such a way that a station will always be capable of determining if other station has accessed the medium at the beginning of the previous slot. This reduces the collision probability by half. Exponential Backoff means that each time the station chooses a slot and happens to collide, it will increase the maximum number for the random selection exponentially. The 802.11 standard defines an **Exponential Backoff Algorithm**, that must be executed in the following cases:
□□□ If when the station senses the medium before the first transmission of a packet, and the medium is busy,
□□□ After each retransmission, and

□□□ After a successful transmission
The only case when this mechanism is not used is when the station decides to transmit a new packet and
the medium has been free for more than DIFS.
The following figure shows a schematic of the access mechanism:



7.3 WiFi security – WPA2

Until recently, wired equivalent privacy (WEP) was the premier 128-bit optional encryption standard used to protect a home wireless network. Although wireless routers, access points, and wireless computer adapters for residential use shipped with WEP capabilities, most manufacturers turned WEP off by default. Many home wireless users never bothered to turn it on, so they had no security and no protection from intruders. And those who sorted through the challenges of configuring WEP on home networks often failed to ever change the WEP key. But now there is good news, especially for Microsoft Windows XP users. Prior weaknesses in wireless security and tricky configuration issues for the home user have been eliminated thanks to a new security specification called Wi-Fi Protected Access**(WPA**).

WPA and Windows XP Service Pack 1 (including Windows XP Home Edition, Windows XP Professional, Tablet PC Edition, and Media Center Edition) coupled with a downloadable update now provide the home user with easier to configure and stronger security than was previously available. Now you don't need to hire a network consultant or find a friendly neighborhood geek to set up wireless security.
In this column, I'll show you how I configured my home network with a special WPA mode, called **WPA-PSK** (**Pre Shared Key**) and I'll explain how WPA-PSK works. I'll share some of the solutions I'm using to include older non-WPA capable 802.11b equipment on a WPA-enabled network. I'll also provide a brief update on the additional security procedures you can use to secure your wireless network.

How Do WPA and WPA-PSK Work?
WPA resolves the issue of weak WEP headers, which are called initialization vectors (IV), and provides a way of insuring the integrity of the messages passed through **MIC** (called Michael or message integrity check) using **TKIP** (the Temporal Key Integrity Protocol) to enhance data encryption. **WPA-PSK** is a special mode of WPA for home users without an enterprise authentication server and provides the same strong encryption protection.

In simple terms, WPA-PSK is extra-strong encryption where encryption keys are automatically changed (called **rekeying**) and authenticated between devices after a specified period of time, or after a specified number of packets has been transmitted. This is called the **rekey interval**. WPA-PSK is far superior to WEP and provides stronger protection for the home/SOHO user for two reasons. The process used to generate the encryption key is very rigorous and the rekeying (or key changing) is done very quickly. This stops even the most determined hacker from gathering enough data to break the encryption.
WEP was confusing to home users because of the various types of keys vendors used (HEX, ASCII, or passphrase) and because home users mix and match equipment from multiple vendors, all using different types of keys. But WPA-PSK employs a consistent, easy to use method to secure your network. This method uses a passphrase (also called a **shared secret**) that must be entered in both the wireless access point/router and the WPA clients. This shared secret can technically be between 8 and 63 characters and can include special characters and spaces. The WPA preshared key should be a random sequence of either keyboard characters (upper and lowercase letters, numbers, and punctuation) at least 20 characters long or hexadecimal digits (numbers 0-9 and letters A-F) at least 24 hexadecimal digits long. The more random your WPA preshared key, the safer it is to use.
The **Temporal Key Integrity Protocol** (**TKIP**) takes over after the initial shared secret is entered in your wireless devices and handles the encryption and automatic rekeying.
WPA is not an official IEEE standard, but is based on and is expected to be compatible with the upcoming 802.11i security standard, sometimes referred to as WPA2. WPA is designed to be a software upgrade. The 802.11i standard will likely require a hardware upgrade. However, wireless vendors and security professionals expect today's WPA and WPA-PSK to be useful for a very long time.

Features of WPA security

The following security features are included in the WPA standard:

WPA authentication

802.1x authentication is required in WPA. In the 802.11 standard, 802.1x authentication was optional.
For environments without a Remote Authentication Dial-In User Service (RADIUS) infrastructure, WPA supports the use of a preshared key. For environments with a RADIUS infrastructure, Extensible Authentication Protocol (EAP) and RADIUS is supported.

WPA key management

With 802.1x, the rekeying of unicast encryption keys is optional. Additionally, 802.11 and 802.1x provide no mechanism to change the global encryption key used for multicast and broadcast traffic. With WPA, rekeying of both unicast and global encryption keys is required. For the unicast encryption key, the Temporal Key Integrity Protocol (TKIP) changes the key for every frame, and the change is synchronized between the wireless client and the wireless access point (AP). For the global encryption key, WPA includes a facility for the wireless AP to advertise the changed key to the connected wireless clients.

Temporal Key Integrity Protocol (TKIP)

For 802.11, Wired Equivalent Privacy (WEP) encryption is optional. For WPA, encryption using TKIP is required. TKIP replaces WEP with a new encryption algorithm that is stronger than the WEP algorithm but that uses the calculation facilities present on existing wireless devices to perform encryption operations. TKIP also provides for the following:

- The verification of the security configuration after the encryption keys are determined.
- The synchronized changing of the unicast encryption key for each frame.
- The determination of a unique starting unicast encryption key for each preshared key authentication.

Michael

With 802.11 and WEP, data integrity is provided by a 32-bit *integrity check value* (ICV) that is appended to the 802.11 payload and encrypted with WEP. Although the ICV is encrypted, you can use cryptanalysis to change bits in the encrypted payload and update the encrypted ICV without being detected by the receiver.

With WPA, a method known as *Michael* specifies a new algorithm that calculates an 8-byte*message integrity code* (MIC) using the calculation facilities available on existing wireless devices. The MIC is placed between the data portion of the IEEE 802.11 frame and the 4-byte ICV. The MIC field is encrypted together with the frame data and the ICV.

Michael also helps provide replay protection. A new frame counter in the IEEE 802.11 frame helps prevent replay attacks.

AES support

WPA defines the use of Advanced Encryption Standard (AES) as an additional replacement for WEP encryption. Because you may not be able to add AES support through a firmware update to existing wireless equipment, support for AES is optional and is dependant on vendor driver support.

Supporting a mixture of WPA and WEP wireless clients

To support the gradual transition of WEP-based wireless networks to WPA, a wireless AP can support both WEP and WPA clients at the same time. During the association, the wireless AP determines which clients use WEP and which clients use WPA. The support of a mixture of WEP and WPA clients is problematic. The global encryption key is not dynamic because WEP-based clients cannot support it. All other benefits to the WPA clients are maintained, including integrity.

Changes required to support WPA

WPA requires software changes to the following:

- Wireless access points
- Wireless network adapters
- Wireless client programs

Changes to wireless access points

Wireless access points must have their firmware updated to support the following:

- **The new WPA information element**

  To advertise their support of WPA, wireless APs send the beacon frame with a new 802.11 WPA information element that contains the wireless AP's security configuration (encryption algorithms and wireless security configuration information).
- **The WPA two-phase authentication**

  Open **system**, and then open **802.1x** (EAP with RADIUS or preshared key).
- **TKIP**
- **Michael**
- **AES** (optional)

To upgrade your wireless access points to support WPA, obtain a WPA firmware update from your wireless AP vendor and upload it to your wireless AP.

Changes to wireless network adapters

Wireless network adapters must have their firmware updated to support the following:

- **The new WPA information element**
  Wireless clients must be able to process the WPA information element and respond with a specific security configuration.
- **The WPA two-phase authentication**
- Open **system**, and then open **802.1x** (EAP or preshared key).
- **TKIP**
- **Michael**
- **AES** (optional)

To upgrade your wireless network adapters to support WPA, obtain a WPA update from your wireless network adapter vendor and update the wireless network adapter driver.

For Windows wireless clients, you must obtain an updated network adapter driver that supports WPA. For wireless network adapter drivers that are compatible with Windows XP (Service Pack 1), the updated network adapter driver must be able to pass the adapter's WPA capabilities and security configuration to the Wireless Zero Configuration service.

Microsoft has worked with many wireless vendors to embed the WPA firmware update in the

wireless adapter driver. Therefore, to update your Windows wireless client, you just obtain the new WPA-compatible driver, and then install the driver. The firmware is automatically updated when the wireless network adapter driver is loaded in Windows.

Changes to wireless client programs
Wireless client programs must be updated to permit the configuration of WPA authentication (and preshared key) and the new WPA encryption algorithms (TKIP and the optional AES component).

For wireless clients that are running Windows XP Service Pack 1 (SP1) and that are using a wireless network adapter that supports the Wireless Zero Configuration service, you must obtain and install the Windows WPA Client. For wireless clients that are running Windows XP service pack 2 (SP2) and that are using a wireless network adapter that supports the Wireless Zero Configuration service, the Windows WPA Client is included in Windows XP SP2. Therefore, additional downloads are not needed. The Windows WPA Client updates the wireless network configuration dialog boxes to support new WPA options.

For more information and to obtain the WPA client program, click the following article number to view the article in the Microsoft Knowledge Base:

For wireless clients running Windows 2000 (or clients running Windows XP SP1 and using a wireless network adapter that does not support the Wireless Zero Configuration service), you must obtain and install a new WPA-compliant configuration tool from your wireless network adapter vendor.

In July 2004, the IEEE approved the full IEEE 802.11i specification, which was quickly followed by a new interoperability testing certification from the WiFi Alliance known asWPA2. WPA2 is based on the Robust Security Network (RSN) mechanism, which provided support for all of the mechanisms available in WPA, as well as:

* Strong encryption and authentication support for infrastructure and ad-hoc networks (WPA is limited to infrastructure networks);

* Reduced overhead in key derivation during the wireless LAN authentication exchange;

* Support for opportunistic key caching to reduce the overhead in roaming between access points;

* Support for pre-authentication, where a station completes the IEEE 802.1X authentication exchange before roaming;

* Support for the CCMP (Counter Mode with Cipher Block Chaining Message Authentication Code Protocol) encryption mechanism based on the Advanced Encryption Standard (AES) cipher as an alternative to the TKIP protocol.

As of March 2006, the WPA2 certification became mandatory for all new equipment certified by the Wi-Fi Alliance, ensuring that any reasonably modern hardware will support both WPA and WPA2.

By leveraging the RC4 cipher (also used in the WEP protocol), the IEEE 802.11i task group was able to improve the security of legacy networks with TKIP while the IEEE 802.11i amendment was completed. It is important to note, however, that TKIP was designed as an interim solution for wireless security, with the goal of providing sufficient security for 5 years while organizations transitioned to the full IEEE 802.11i security mechanism. While there have not been any catastrophic weaknesses reported in the TKIP protocol, organizations should take this design requirement into consideration and plan to transition WPA networks to WPA2 to take advantage of the benefits provided by the RSN architecture.

Chapter 8

Bluetooth

8.1 Overview of Bluetooth

8.2 Radio specification

8.3 Base band Specification

8.4 Link Manager Specification

8.5 Logical Link Control and Adaptation

8.1 Overview of Bluetooth

During the past two decades, the progress in microelectronics and VLSI technology drove the cost of many consumer electronic products down to an acceptable level for average people. Only in the 1st quarter of 2001, over 32.5 million PCs were sold. The number of cellular phones is predicted to reach 1 billion in 2005. With the increase of the number of these devices, so does the need of connecting them together. Today numerous kinds of special cables are used for interconnection. It's cumbersome, not interchangeable and expensive. Bluetooth is devised to replace these cables. Bluetooth is a low cost, low power, radio frequency technology for short-range communications. It can be used to replace the cables connecting portable/fixed electronic devices, build ad-hoc networks or provide data/voice access points.

Frequency	2.4GHz ISM band, Frequency hopping
Modulation	Gaussian shaped BFSK
Data rate	723Kbps
Operating range	10m~100m
Size	28mm x 15mm x 2mm (Mitsumi WML-C05)
Cost	Long term: $5/endpoint ($20 currently)
Power	0.1W (Active)
Security	Good. FHSS. Link layer authentication and encryption
Acceptance	SIG have about 2500 member companies

Table 1. Bluetooth Summary

Figure 1. A Bluetooth Module

The research on Bluetooth was initiated at Ericsson of Sweden in 1994. The idea of Bluetooth comes from the desire to connect cellular phones with other devices without a cable. It's named after the 10th century Viking king of Denmark Herald Bluetooth. The advancement in microelectronics makes it possible to integrate complex functions into one small chip and thus achieve a low cost. With its low cost, low power consumption and low profile, you can virtually put one anywhere you want. This will make many concepts like smart appliances and embedded Internet possible. The development gained support from many companies. Currently, there are about 2500 companies joined the Bluetooth Special Interest Group (SIG). There are some commercial products available, and much more are rolling out. A new standard for Wireless Personal Area Network (WPAN)-IEEE802.15 is being developed, and to a large extent, it's an extension of Bluetooth. Despite its advantages, one of its key limitations so far is its speed. With a maximum data rate of 720KBps, it cannot be used to connect DVD players or HDTV, and it takes a long time to transfer large picture files to a printer. New version of Bluetooth may address this issue and have much higher data rate.

8.2 Radio Specification

8.2.1 Frequency Bands and Channel Arrangement

The Bluetooth radio accomplishes spectrum spreading by frequency hopping in 79 hops displaced by 1 MHz, starting at 2.402GHz and finishing at 2.480GHz. In a few countries (i.e France) this frequency band range is (temporarily) reduced, and a 23-hop system is used. In order to comply with out of band regulations in each country. In both systems a guard band is used at the lower and upper band edge

8.2.2 Transmitter Characteristics

Power Classes: Each device is classified into 3 power classes, Power Class 1, 2 & 3.

- Power Class 1: is designed for long range (~100m) devices, with a max output power of 20 dBm,
- Power Class 2: for ordinary range devices (~10m) devices, with a max output power of 4 dBm,
- Power Class 3: for short range devices (~10cm) devices, with a max output power of 0 dBm.

The Bluetooth radio interface is based on a nominal antenna power of 0dBm. Each device can optionally vary its transmitted power. Equipment with power control capability optimizes the output power in a link with LMP commands. It is done by measuring RSSI and report back if the power should be increased or decreased.

Modulation Characteristics: The Bluetooth radio module uses GFSK (Gaussian Frequency Shift Keying) where a binary one is represented by a positive frequency deviation and a binary zero by a negative frequency deviation. BT is set to 0.5 and the modulation index must be between 0.28 and 0.35.

Spurious Emissions: The spurious emission, in-band and out-of-band, is measured with a frequency hopping transmitter hopping on a single frequency; this means that the synthesizer must change frequency between receive slot and transmit slot, but always returns to the same transmit frequency.

Radio Frequency Tolerance: The transmitted initial center frequency accuracy must be ±75 kHz from Fc. The initial frequency accuracy is defined as being the frequency accuracy before any information is transmitted. Note that the frequency drift requirement is not included in the ±75 kHz.

8.2.3 Receiver Characteristics

Sensitivity Level: The receiver must have a sensitivity level for which the bit error rate (BER) 0.1% is met. For Bluetooth this means an actual sensitivity level of -70dBm or better.

Interference Performance: The interference performance on Co-channel and adjacent 1 MHz and 2 MHz are measured with the wanted signal 10 dB over the reference sensitivity level. On all other frequencies the wanted signal shall be 3 dB over the reference sensitivity level.

Out-of-Band blocking: The Out of band blocking is measured with the wanted signal 3 dB over the reference sensitivity level. The interfering signal shall be a continuous wave signal. The BER shall be less than or equal to 0.1%.

Intermodulation Characteristics: The reference sensitivity performance, BER = 0.1%, shall be met under the following conditions.

- The wanted signal at frequency f 0 with a power level 6 dB over the reference sensitivity level.
- A static sine wave signal at f 1 with a power level of –39 dBm
- A Bluetooth modulated signal at f 2 with a power level of -39 dBm

Such that f 0 =2f 1 -f 2 and |f 2 -f 1| =n*1 MHz, where n can be 3, 4, or 5. The system must fulfil one of the three alternatives.

Maximum Usable Level: The maximum usable input level the receiver shall operate at shall be better than –20 dBm. The BER shall be less or equal to 0,1% at –20* dBm input power.

RSSI: Receiver Signal Strength Indicator (Optional): A transceiver that wishes to take part in a power-controlled link must be able to measure its own receiver signal strength and determine if the transmitter on the other side of the link should increase or decrease its output power level. A Receiver Signal Strength Indicator (RSSI) makes this possible. The way the power control is specified is to have a **golden receive power range**. This golden receive power is defined as a range with a lower and higher threshold levelsand a high limit. The lower threshold level corresponds to a received power between -56 dBm and 6 dB above the actual sensitivity of the receiver. The upper threshold level is 20 dB above the lower threshold level to an accuracy of +/- 6 dB. The instructions to alter the TX power are carried in the LMP link

8.3 Baseband Specifications

The Baseband is the physical layer of the Bluetooth. It manages physical channels and links apart from other services like error correction, data whitening, hop selection and Bluetooth security. The Baseband layer lies on top of the Bluetooth radio layer in the bluetooth stack. The baseband protocol is implemented as a Link Controller , which works with the link manager for carrying out link level routines like link connection and power control. The baseband also manages asynchronous and synchronous links, handles packets and does paging and inquiry to access and inquire Bluetooth devices in the area. The baseband transceiver applies a time-division duplex (TDD) scheme. (alternate transmit and receive). Therefore apart from different hopping frequency (frequency division), the time is also slotted.

Physical Characteristics

8.3.1.1 Physical Channel

Bluetooth operates in the **2.4 GHz ISM band**. In the US and Europe, a band of 83.5 MHz width is available; in this band, 79 RF channels spaced 1 MHz apart are defined. In France, a smaller band is available; in this band, 23 RF channels spaced 1 MHz apart are defined.

The channel is represented by a **pseudo-random hopping sequence** hopping through the 79 or 23 RF channels. Two or more Bluetooth devices using the same channel form a **piconet**. There is one **master** and one or more **slave**(s) in each piconet. The hopping sequence is unique for the piconet and is determined by the Bluetooth device address (BD_ADDR) of the master; the phase in the hopping sequence is determined by the Bluetooth clock of the master. The channel is divided into time slots where each slot corresponds to an RF hop frequency. Consecutive hops correspond to different RF hop frequencies.

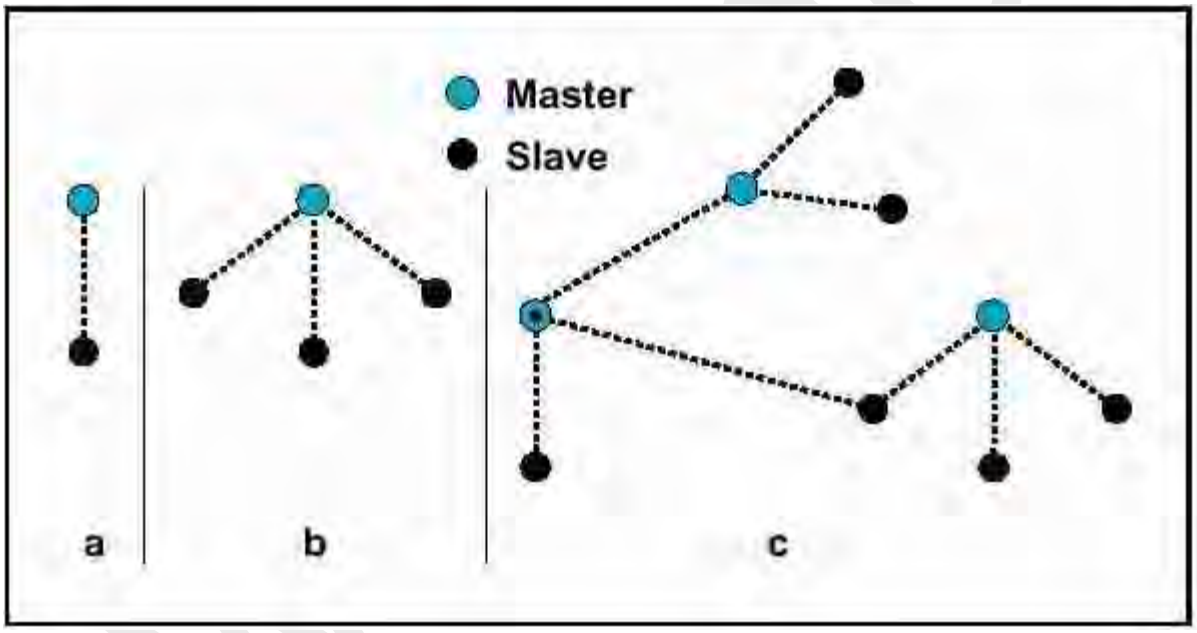


*Diagram Source: Courtesy of Bluetooth SIG, Baseband Spec, Figure 1.2 , p 42

The channel is divided into time slots, each 625 us in length. The time slots are numbered according to the Bluetooth clock of the piconet master.

A TDD scheme is used where master and slave alternatively transmit. The master shall start its transmission in even-numbered time slots only, and the slave shall start its transmission in odd-numbered time slots only. The packet start shall be aligned with the slot start.

8.3.1.2 Physical Links

The Baseband handles two types of links : SCO (Synchronous Connection-Oriented) and ACL (Asynchronous Connection-Less) link. The SCO link is a symmetric point-to-point link between a master and a single slave in the piconet. The master maintains the SCO link by using reserved slots at regular intervals (circuit switched type). The SCO link mainly carries voice information. The master can support up to three simultaneous SCO links while slaves can support two or three SCO links. SCO packets are never retransmitted. SCO packets are used for 64 kB/s speech transmission.

The ACL link is a point-to-multipoint link between the master and all the slaves participating on the piconet. In the slots not reserved for the SCO links, the master can establish an ACL link on a per-slot basis to any slave, including the slave already engaged in an SCO link (packet switched type). Only a single ACL link can exist. For most ACL packets, packet retransmission is applied.

8.3.1.3 Logical Channels

Bluetooth has five logical channels which can be used to transfer different types of information. LC (Control Channel) and LM (Link Manager) channels are used in the link level while UA, UI and US channels are used to carry asynchronous, isosynchronous and synchronous user information.

8.3.1.4 Device Addressing

4 possible types of addresses can be assigned to bluetooth units, BD_ADDR, AM_ADDR, PM_ADDR &AR_ADDR

BD_ADDR: Bluetooth Device Address.	Each Bluetooth transceiver is allocated a unique 48-bit device address. It is divided into a 24-bit LAP field, a 16-bit NAP field and a 8-bit UAP field.
AM_ADDR: Active Member Address	It is a 3-bit number. It is only valid as long as the slave is active on the channel. It is also sometimes called the MAC address of a Bluetooth unit.
PM_ADDR: Parked Member Address	It is a 8-bit member (master-local) address that separates the parked slaves. The PM_ADDR is only valid as long as the slave is parked.
AR_ADDR: Access Request Address	This is used by the parked slave to determine the slave-to-master half slot in the access window it is allowed to send

access request messages in. It is only valid as long as the slave is parked and is not necessarily unique.

8.3.2 Packets

All data on the piconet channel is conveyed in packets.

8.3.2.1 Packet Types

13 different packet types are defined for the baseband layer of the Bluetooth system. All higher layers use these packets to compose higher level PDU's. The packets are ID, NULL, POLL, FHS , DM1 ; these packets are defined for both SCO and ACL links. DH1, AUX1, DM3, DH3, DM5, DH5 are defined for ACL links only. HV1, HV2, HV3 ,DV are defined for SCO links only.

8.3.2.2 Packet Format

Each packet consists of 3 entities, the **access code** (68/72 bits), the **header** (54 bits) , and the **payload** (0-2745 bits).

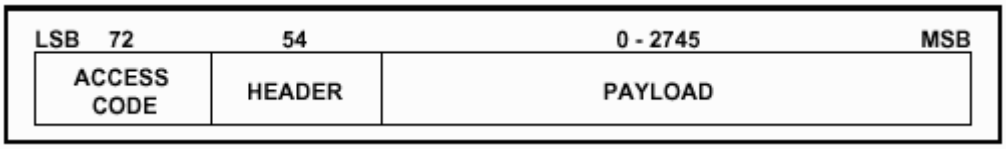


*Diagram Source: Courtesy of Bluetooth SIG, Baseband Specs, Fig 4.1 , p 47

- **Access Code**: Access code are used for timing synchronization, offset compensation, paging and inquiry. There are three different types of Access code: Channel Access Code (CAC), Device Access Code (DAC) and Inquiry Access Code (IAC). The channel access code identifies a unique piconet while the DAC is used for paging and its responses. IAC is used for inquiry purpose.
- **Header**:The header contains information for packet acknowledgement, packet numbering for out-of-order packet reordering, flow control, slave address and error check for header.
- **Payload**: The packet payload can contain either voice field, data field or both. It it has a data field, the payload will also contain a payload header.

8.3.2.3 Channel Control

8.3.2.3.1 Controller States

Bluetooth controller operates in two major states: **Standby** and **Connection** . There are seven substates which are used to add slaves or make connections in the piconet. These are **page, page scan, inquiry, inquiry scan, master response, slave response and inquiry response** .

The **Standby** state is the default low power state in the Bluetooth unit. Only the native clock is running and there is no interaction with any device whatsoever. In the **Connection** state, the master and slave can exchange packet , using the channel (master) access code and the master Bluetooth clock. The hopping scheme used is the channel hopping scheme.The other states (page, inquiry etc are described below)

8.3.2.3.2 Connection Setup (Inquiry/Paging)

Normally, a connection between two devices occur in the following fashion: If nothing is known about a remote device, both the inquiry(**1**) and page(**2**) procedure have to be followed. If some details are known about a remote device, only the paging procedure (**2**) is needed

Step 1:

The **inquiry procedure** enables a device to discover which devices are in range, and determine the addresses and clocks for the devices.

1.1: The inquiry procedure involve a unit (the source) sending out inquiry packets (**inquiry** state) and then receiving the inquiry reply

1.2: The unit that receives the inquiry packets (the destination), will hopefully be in the **inquiry scan** state to receive the inquiry packets.

1.3: The destination will then enter the **inquiry response** state and send an inquiry reply to the source.

After the inquiry procedure has completed, a connection can be established using the paging procedure.

Step 2:

With the **paging procedure**, an actual connection can be established. The paging procedure typically follows the inquiry procedure. Only the Bluetooth device address is required to set up a connection. Knowledge about the clock (clock estimate) will accelerate the setup procedure. A unit that establishes a connection will carry out a page procedure and will automatically be the master of the connection. The procedure occurs as follows:

2.1:	A device (the source) pages another device (the destination).	**Page** state
2.2:	The destination receives the page.	**Page Scan** state
2.3:	The destination sends a reply to the source.	**Slave Response** state : (Step 1)
2.4:	The source sends an FHS packet to the destination.	**Master Response** state : (Step 1)
2.5:	The destination sends it's second reply to the source.	**Slave Response** state : (Step 2)
2.6:	The destination & source then switch to the source channel parameters.	**Master Response** state: Step 2**& Slave Response** state: Step 3

The **Connection** state starts with a POLL packet sent by the master to verify that slave has switched to the master's timing and channel frequency hopping. The slave can respond with any type of packet.

8.3.2.3.3 Connection Modes

A Bluetooth device in the **Connection** state can be in any of the four following modes: **Active, Hold, Sniff** and**Park** mode.

- **Active Mode:** In the active mode, the Bluetooth unit actively participates on the channel. The master schedules the transmission based on traffic demands to and from the different

slaves. In addition, it supports regular transmissions to keep slaves synchronized to the channel. Active slaves listen in the master-to-slave slots for packets. If an active slave is not addressed, it may sleep until the next new master transmission.

- **Sniff Mode:** Devices synchronized to a piconet can enter power-saving modes in which device activity is lowered. In the SNIFF mode, a slave device listens to the piconet at reduced rate, thus reducing its duty cycle. The SNIFF interval is programmable and depends on the application. It has the highest duty cycle (least power efficient) of all 3 power saving modes (sniff, hold & park).

- **Hold Mode**: Devices synchronized to a piconet can enter power-saving modes in which device activity is lowered. The master unit can put slave units into HOLD mode, where only an internal timer is running. Slave units can also demand to be put into HOLD mode. Data transfer restarts instantly when units transition out of HOLD mode. It has an intermediate duty cycle (medium power efficient) of the 3 power saving modes (sniff, hold & park).

- **Park Mode:** In the PARK mode, a device is still synchronized to the piconet but does not participate in the traffic. Parked devices have given up their MAC (AM_ADDR) address and occasional listen to the traffic of the master to re-synchronize and check on broadcast messages. It has the lowest duty cycle (power efficiency) of all 3 power saving modes (sniff, hold & park).

8.3.2.3.4 Scatternet

Multiple piconets may cover the same area. Since each piconet has a different master, the piconets hop independently, each with their own channel hopping sequence and phase as determined by the respective master. In addition, the packets carried on the channels are preceded by different channel access codes as determined by the master device addresses. As more piconets are added, the probability of collisions increases; a graceful degradation of performance results as is common in frequency-hopping spread spectrum systems.

If multiple piconets cover the same area, a unit can participate in two or more overlaying piconets by applying time multiplexing. To participate on the proper channel, it should use the associated master device address and proper clock offset to obtain the correct phase. A Bluetooth unit can act as a slave in several piconets, but only as a master in a single piconet. A group of piconets in which connections consists between different piconets is called a **scatternet**.

Sometimes an existing master or slave may wish to swap roles (i.e a **master-slave switch**) , this can take place in two steps:

1. First a TDD switch of the considered master and slave, followed by a piconet switch of the both participants.

2. Then, if so desired, other slaves of the old piconet can be transferred to the new piconet.

When a unit have acknowledged the reception of the FHS packet, this unit uses the new piconet parameters defined by the new master and the piconet switch is completed.

8.3.2.4 Other Baseband Functions

8.3.2.4.1 Error Correction

There are three kinds of error correction schemes used in the baseband protocol: 1/3 rate FEC, 2/3 rate FEC and ARQ scheme.

- In **1/3 rate FEC** every bit is repeated three times for redundancy,
- In **2/3 rate FEC** a generator polynomial is used to encode 10 bit code to a 15 bit code,
- In the **ARQ scheme** , **DM, DH** and the data field of **DV** packets are retransmitted till an acknowledgement is received (or timeout is exceeded). Bluetooth uses fast, unnumbered acknowledgement in which it uses positive and negative acknowledgements by setting appropriate ARQN values. If the timeout value is exceeded, Bluetooth flushes the packet and proceeds with the next.

8.3.2.4.2 Flow Control

The Baseband protocol recommends using FIFO queues in ACL and SCO links for transmission and receive. The Link Manager fills these queues and link controller empties the queues automatically.

If these RX FIFO queues are full, flow control is used to avoid dropped packets and congestion. If data cannot be received, a **stop** indication is transmitted inserted by the Link Controller of the receiver into the header of the return packet. When the transmitter receives the **stop** indication, it freezes its FIFO queues. If receiver is ready it sends a**go** packet which resumes the flow again.

8.3.2.4.3 Synchronization

The Bluetooth transceiver uses a time-division duplex (TDD) scheme ,meaning that it alternately transmits and receives in a synchronous manner. The average timing of master packet

transmission should not drift faster than 20 ppm relative to the ideal slot timing of 625 us. Jitter from average timing should be less than 1 microsecond.

The piconet is synchronized by the system clock of the master.To transmit on the piconet channel you need 3 pieces of information, The (channel) hopping sequence, the phase of the sequence, and the CAC to place on the packets

1	**Channel Hopping Sequence**	The Bluetooth Device Address (BD_ADDR) of the master is used to derive this frequency hopping sequence.
2	**Phase**	The system clock of the master determines the phase in the hopping sequence.
3	**Channel Access Code**	This is derived from the Bluetooth Device Address (BD_ADDR) of the master.

The slaves adapt their native clocks with a timing offset in order to match the master clock, giving then anestimated clock value. The offset is zero for the master as it's native clock **is** the master clock. The Bluetooth clocks should have the LSB ticking in units of 312.5us, giving a clock rate of 3.2kHz.

A 20us uncertainty window is allowed around the exact receive time in order for the access correlator for the receiver to search for the correct channel access code and get synchronized with the transmitter. When a slave returns from the hold mode, it can correlate over a bigger uncertainty window till they don't overlap slots. A parkedslave periodically wakes up to listen to beacons from the master and re-synchronizes its clock offset.

8.3.2.4.4 Bluetooth Security

At the link layer, security is maintained by authentication of the peers and encryption of the information. For this basic security we need a public address which is unique for each device (BD_ADDR), two secret keys (authentication keys and encryption key) and a random number generator. First a device does the authentication by issuing a challenge and the other device has to then send a response to that challenge which is based on the challenge, it's BD_ADDR and a link key shared between them. After authentication, encryption may be used to communicate. See our Bluetooth Security page and Bluetooth article(s) for more details

8.4 Link Manager Specification

8.4.1 LMP PDUs

8.4.1.1 General Response

(M) LMP_accepted , LMP_not_accepted

These PDU's are used as response messages to other PDU's in a number of different procedures, containing the opcode of the message that is being responded to.

8.4.1.2 Authentication

(M) LMP_au_rand , LMP_sres

The authentication procedure is based on a challenge-response scheme. The verifier sends an LMP_au_rand PDU which contains a random number (the challenge) to the claimant. The claimant calculates a response, which is a function of the challenge, the claimant's BD_ADDR and a secret key. The response is sent back to the verifier, which checks if the response was correct or not. A successful calculation of the authentication response requires that two devices share a secret key. Both the master and the slave can be verifiers.

8.4.1.3 Pairing

(M) LMP_in_rand , LMP_au_rand , LMP_sres , LMP_comb_key , LMP_unit_key

When two devices do not have a common link key an initialization key (K init) is created based on a PIN and a random number. When both devices have calculated K init the link key is created, and finally a mutual authentication is made. The pairing procedure starts with a device sending LMP_in_rand; this device is referred to as "initiating LM" or "initiator" in. The other device is referred to as "responding LM" or "responder".

8.4.1.4 Change Link Key

(M) LMP_comb_key

If the link key is derived from combination keys and the current link is the semi-permanent link key, the link key can be changed. If the link key is a unit key, the units must go through the pairing procedure in order to change the link key. The contents of LMP_comb_key is protected by a bitwise XOR with the current link key.

8.4.1.5 Change the Current Link Key

(M) LMP_temp_rand , LMP_temp_key , LMP_use_semi_permanent_key

The current link key can be a semi-permanent link key or a temporary link key key. It can be changed temporarily, but the change is only valid for the session. Changing to a temporary link key is necessary if the piconet is to support encrypted broadcast.

8.4.1.6 Encryption

(O) LMP_encryption_mode_req , LMP_encryption_key_size_req , LMP_start_encryption_req, LMP_stop_encryption_req

If at least one authentication has been performed encryption may be used. If the master wants all slaves in the piconet to use the same encryption parameters it must issue a temporary key (K master) and make this key the current link key for all slaves in the piconet before encryption is started. This is necessary if broadcast packets should be encrypted.

8.4.1.7 Clock Offset Request

(M) LMP_clkoffset_req, LMP_clkoffset_res

When a slave receives the FHS packet, the difference is computed between its own clock and the master's clock included in the payload of the FHS packet. The clock offset is also updated each time a packet is received from the master. The master can request this clock offset anytime during the connection. By saving this clock offset the master knows on what RF channel the slave wakes up to PAGE SCAN after it has left the piconet. This can be used to speed up the paging time the next time the same device is paged.

8.4.1.8 Slot Offset Request

(O) LMP_slot_offset

With LMP_slot_offset the information about the difference between the slot boundaries in different piconets is transmitted. This PDU carries the parameters slot offset and BD_ADDR. The slot offset is the subtraction of time in us of the start of the master's TX slot in the piconet

where the PDU is transmitted from the time in us of the start of the master's TX slot in the piconet where the BD_ADDR device is master modulo 1250.

. Before doing a master-slave switch, this PDU shall be transmitted from the device that becomes master in the switch procedure. The PDU can also be useful in inter-piconet communications.

8.4.1.9 Timing Accuracy Information Request

(O) LMP_timing_accuracy_req , LMP_timing_accuracy_res

LMP supports requests for the timing accuracy. This information can be used to minimize the scan window for a given hold time when returning from hold and to extend the maximum hold time. It can also be used to minimize the scan window when scanning for the sniff mode slots or the park mode beacon packets. The timing accuracy parameters returned are the long term drift measured in ppm and the long term jitter measured in μs of the clock used during hold, sniff and park mode. These parameters are fixed for a certain device and must be identical when requested several times.

8.4.1.10 LMP Version

(M) LMP_version_req , LMP_version_res

The LMP layer supports requests for the version of the LM protocol. The requested device will send a response with three parameters: VersNr, CompId and Sub-VersNr. VersNr specifies the version of the Bluetooth LMP specification that the device supports. CompId is used to track possible problems with the lower Bluetooth layers. All companies that create a unique implementation of the Link Manager shall have their own CompId. The same company is also responsible for the administration and maintenance of the SubVersNr. It is recommended that each company has a unique SubVersNr for each RF/BB/LM implementation.

8.4.1.11 Supported Features

(M) LMP_features_req , LMP_features_res

The Bluetooth radio and link controller may support only a subset of the packet types and features described in Baseband Specification and Radio Specification. A device may not send any packets other than ID, FHS, NULL, POLL, DM1 or DH1 before it is aware of the supported features of the other device. After the features request has been carried out, the intersection of the supported packet types for both sides may also be transmitted. Whenever a request is issued, it

must be compatible with the supported features of the other device. For instance, when establishing an SCO link the initiator may not propose to use HV3 packets if that packet type is not supported by the other device.

8.4.1.12 Switch of Master-Slave Role

(O) LMP_switch_req , LMP_slot_offset

Since the paging device always becomes the master of the piconet, a switch of the master-slave role is sometimes needed. Suppose device A is slave and device B is master. The device that initiates the switch finalises the transmission of the current L2CAP message and then sends LMP_switch_req. Note: in a slave initiated master-slave switch the slave (A) will first send LMP_slot_offset, then LMP_switch. In a master initiated master-slave switch, the master (B) will first send LMP_switch, before receiving LMP_slot_offset from the slave (A). If the switch is accepted, the other device finalises the transmission of the current L2CAP message and then responds with LMP_accepted. The switch procedure then takes place, and afterwards device A is master and device B is slave.

8.4.1.13 Name Request

(M) LMP_name_req , LMP_name_res

LMP supports name request to another Bluetooth device. The name is a user-friendly name associated with the Bluetooth device and consists of a maximum of 248 bytes coded according to the UTF-8 standard. The name is fragmented over one or more DM1 packets.

8.4.1.14 Detach

(M) LMP_detach

The connection between two Bluetooth devices can be closed anytime by the master or the slave. A reason parameter is included in the message to inform the other party of why the connection is closed.

8.4.1.15 Hold Mode

(O) LMP_hold , LMP_hold_req

The ACL link of a connection between two Bluetooth devices can be placed in hold mode for a specified hold time. During this time no ACL packets will be transmitted from the master. The hold mode is typically entered when there is no need to send data for a relatively long time. The transceiver can then be turned off in order to save power. But the hold mode can also be used if a device wants to discover or be discovered by other Bluetooth devices, or wants to join other piconets. What a device actually does during the hold time is not controlled by the hold message, but it is up to each device to decide.

8.4.1.16 Sniff Mode

(O) LMP_sniff_req , LMP_unsniff_req

To enter sniff mode, master and slave negotiate a sniff interval T sniff and a sniff offset, D sniff , which specifies the timing of the sniff slots. The offset determines the time of the first sniff slot; after that the sniff slots follows periodically with the sniff interval T sniff . When the link is in sniff mode the master can only start a transmission in the sniff slot. Two parameters control the listening activity in the slave. The sniff attempt parameter determines for how many slots the slave must listen, beginning at the sniff slot, even if it does not receive a packet with its own AM address. The sniff timeout parameter determines for how many additional slots the slave must listen if it continues to receive only packets with its own AM address.

8.4.1.17 Park Mode

(O) LMP_park_req , LMP_unpark_PM_ADDR_req , LMP_unpark_BD_ADDR_req , LMP_set_broadcast_scan_window , LMP_modify_beacon

If a slave does not need to participate in the channel, but still should be FH-synchronized, it can be placed in park mode. In this mode the device gives up its AM_ADDR but still re-synchronizes to the channel by waking up at the beacon instants separated by the beacon interval. The beacon interval, a beacon offset and a flag indicating how the first beacon instant is calculated determine the first beacon instant. After this the beacon instants follow periodically at the predetermined beacon interval. At the beacon instant the parked slave can be activated again by the master, the master can change the park mode parameters, transmit broadcast information or let the parked slaves request access to the channel.

All PDUs sent from the master to the parked slaves are broadcast. These PDUs are the only PDUs that can be sent to a slave in park mode and the only PDUs that can be broadcast. When a slave is placed in park mode it is assigned a unique PM_ADDR, which can be used by the master to unpark that slave.

8.4.1.18 Power Control

(O) LMP_incr_power_req , LMP_decr_power_req , LMP_max_power , LMP_min_power

If the RSSI value differs too much from the preferred value of a Bluetooth device, it can request an increase or a decrease of the other device's TX power. Upon receipt of this message, the output power is increased or decreased one step. At the master side the TX power is completely independent for different slaves; a request from one slave can only effect the master's TX power for that same slave.The power adjustment requests can be made at anytime following a successful baseband paging procedure. If a device does not support power control requests this is indicated in the supported features list.

8.4.1.19 Channel Quality-Driven Change (between DM and DH)

(O) LMP_auto_rate , LMP_preferred_rate

The data throughput for a given packet type depends on the quality of the RF channel. Quality measurements in the receiver of one device can be used to dynamically control the packet type transmitted from the remote device for optimization of the data throughput. If a device A wants the remote device B to have this control it sends LMP_auto_rate once. The device B can then send back LMP_preferred_rate to device A whenever it wishes to change the packet type that A transmits.

This PDU has a parameter which determines the preferred coding (with or without 2/3FEC) and the preferred size (in slots) of the packets. Device A is not required to change to the packet type specified by this parameter and may never send a packet that is larger than the maximum allowed number of slots even if the preferred size is greater than this value.

8.4.1.20 Quality of Service

(M) LMP_quality_of_service , LMP_quality_of_service_req

The LM provides Quality of Service capabilities. A poll interval, which is defined as the maximum time between subsequent transmissions from the master to a particular slave, is used to support bandwidth allocation and latency control. In addition, master and slave negotiate the number of repetitions for broadcast packets (NBC).

8.4.1.21 SCO Links

(O) LMP_SCO_link_req , LMP_remove_SCO_link_req

When a connection has been established between two Bluetooth devices the connection consists of an ACL link. One or more SCO links can then be established. The SCO link reserves slots separated by the SCO interval, T sco . The first slot reserved for the SCO link is defined by T sco and the SCO delay, D sco . After that the SCO slots follows periodically with the SCO interval. Each SCO link is distinguished from all other SCO links by an SCO handle.

8.4.1.22 Control of Multi-Slot Packets

(M) LMP_max_slot , LMP_max_slot_req

The number of slots used by a device can be limited. A device allows the remote device to use a maximal number of slots by sending the PDU LMP_max_slot providing max slots as parameter. Each device can request to use a maximal number of slots by sending the PDU LMP_max_slot_req providing max slots as parameter. After a new connection, as a result of page, page scan, master-slave switch or unpark, the default value is 1 slot. Two PDUs are used for the control of multi-slot packets.These PDUs can be sent at anytime after connection setup is completed.

8.4.1.23 Paging Scheme

(O) LMP_page_mode_req , LMP_page_scan_mode_req

In addition to the mandatory paging scheme, the Bluetooth system defines optional paging schemes. LMP provides a means to negotiate the paging scheme, which is to be used the next time a unit is paged.

8.4.1.24 Link Supervision

(M) LMP_supervision_timeout

Each Bluetooth link has a timer that is used for link supervision. This timer is used to detect link loss caused by devices moving out of range, a device's power-down, or other similar failure cases. An LMP procedure is used to set the value of the supervision timeout.

8.4.1.25 Connection Establishment

(M) LMP_host_connection_req , LMP_setup_complete

When the paging device wishes to create a connection involving layers above LM, it sends LMP_host_connection_req. When the other side receives this message, the host is informed about the incoming connection. The remote device can accept or reject the connection request by sending LMP_accepted or LMP_not_accepted.

If LMP_host_connection_req is accepted, LMP security procedures (pairing, authentication and encryption) can be invoked. When a device is not going to initiate any more security procedures during connection establishment it sends LMP_setup_complete. When both devices have sent LMP_setup_complete the first packet on a logical channel different from LMP can then be transmitted.

8.4.1.26 Test Mode

(M) LMP_test_activate , LMP_test_control

LMP has PDUs to support different Bluetooth test modes, which are used for certification and compliance testing of the Bluetooth radio and baseband.

8.4.1.27 Error Handling

(M) LMP_not_accepted

If the Link Manager receives a PDU with unrecognised opcode, it responds with LMP_not_accepted with the reason code *unknown LMP PDU*. The opcode parameter that is echoed back is the unrecognised opcode. If the Link Manager receives a PDU with invalid parameters, it responds with LMP_not_accepted with the reason code*invalid LMP parameter*s. If

the maximum response time is exceeded or if a link loss is detected the party that waits for the response shall conclude that the procedure has terminated unsuccessfully.

Erroneous LMP messages can be caused by errors on the channel or systematic errors at the transmit side. To detect the latter case, the LM should monitor the number of erroneous messages and disconnect if it exceeds a threshold, which is implementation-dependent.

8.5 Logical Link Control and Adaptation

L2CAP provides connection-oriented and connectionless data services to upper layer protocols with protocol multiplexing capability, segmentation and reassembly operation, and group abstractions. L2CAP permits higher level protocols and applications to transmit and receive L2CAP packets up to 64 Kilobytes in length. L2CAP only supports ACL links. L2CAP uses the concept of channels to establish different pathways between different applications on Bluetooth devices. These channels are identified by Channel IDentifiers (CIDs) which represent a logical end point of a connection for each application on a device. CIDs are 16 bit numbers of which 0x0001 to 0x003F are reserved for specific L2CAP functions (0x0001 is a signalling channel, 0x0002 is a connectionless reception channel and rest are reserved or prohibited).

Connection Identifiers (CIDs)

The idea behind L2CAP is to provide an interface similar to TCP/IP function calls. In NSBLUE, the L2CAP connections and data are sent in the following manner:

```
cid = l2cap->openL2CAPConnection(ui->getContext());
l2cap->send(cid,data,len);
l2cap->recv(cid,data,len);
```

where `cid' is the channel identifier, `data' is the pointer to ``char'' and `len' is the length of the data. `getContext()' returns the local context like `Walmart', `JFK Airport' etc. CID is assigned from a pool of free CIDs (pool can be assigned in blocks to save memory). See `CONNECTION ESTABLISHMENT IN BLUETOOTH' for more information.

Protocol Multiplexing

L2CAP also does protocol multiplexing by using PSM field in the L2CAP Connection

Request command. L2CAP can multiplex connection requests to upper layer protocols like Service Discovery Protocol (PSM = 0x0001), RFCOMM (PSM = 0x0003) and Telephony Control (PSM = 0x0005).

Segmentation and Reassembly

Segmentation and Reassembly (SAR) operations are used to improve efficiency by supporting a maximum transmission unit (MTU) size larger than the largest Baseband packet. This reduces the overhead by spreading the network and transport packets used by higher layer protocols over several baseband packets. L2CAP segments higher layer packets into `chunks' that can be passed to the Link Manager for transmission and reassembles those chunks into L2CAP packets using information provided through HCI and from packet header. SAR is implemented using very little overhead in Baseband packets. The two L_CH bits defined in the first byte of Baseband payload (also called the frame header) are used to signal the start and continuation of L2CAP packets (L_CH shall be `10' for the first segment and `01' for a continuation segment. To avoid any reassembly problems due to out of order packets, all L2CAP segments associated with an L2CAP packet must be passed through to the Baseband before any other L2CAP packet destined to the same unit may be sent. Also `Stop and Wait' protocol used by Baseband makes sure that a packet is received correctly by the other unit before the next packet is sent. This avoids most of the out of the order packets, as received in wireline TCP/IP connections because of `window' based transmissions.

L2CAP Events and Actions

L2CAP operates using events and commands which it receives or transmits from/to upper or lower layers. These events can be, for instance, a connection request from the upper layer, a data write request or may be a disconnection request. The lower layers can tell L2CAP about incoming connection, disconnection or other requests. If L2CAP of this unit needs to talk to the L2CAP on the other unit, then it uses some special commands which are called signalling commands. These commands are generally used to establish connection-oriented channels after a link level connection is created or present. L2CAP has seven operational states: CLOSED, W4_L2CA_CONNECT_RSP, W4_L2CAP_CONNECT_RSP, CONFIG, OPEN, W4_L2CAP_DISCONNECT_RSP, and W4_L2CA_DISCONNECT_RSP. These states further make L2CAP connections look similar to TCP connections.

Signalling command codes

Code	Description
0x00	RESERVED
0x01	Command reject
0x02	Connection request

0x03	Connection response
0x04	Configure request
0x05	Configure response
0x06	Disconnection request
0x07	Disconnection response
0x08	Echo request
0x09	Echo response
0x0a	Information request
0x0b	Information response

Chapter 9

Mobile Computing platforms

9.1 TAPI

9.2 Mobile Operating Systems (Mobile OS)

9.1 TAPI

The Telephony Application Programming Interface (TAPI) is a Microsoft Windows API, which provides computer telephony integration and enables PCs running Microsoft Windows to use telephone services. Different versions of TAPI are available on different versions of Windows. TAPI allows applications to control telephony functions between a computer and telephone network for data, fax, and voice calls. It includes basic functions, such as dialing, answering, and hanging up a call. It also supports supplementary functions, such as hold, transfer, conference, and call park found in PBX, ISDN, and other telephone systems.

TAPI is used primarily to control either modems or, more recently, to control business telephone system (PBX) handsets. When controlling a PBX handset, the driver is provided by the manufacturer of the telephone system. Some manufacturers provide drivers that allow the control of multiple handsets. This is traditionally called "third-party control". Other manufacturers provide drivers that allow the control of a single handset. This is called "first-party control". Third-party drivers are designed to allow applications to see and/or control multiple extensions at the same time. Some telephone systems only permit one third-party connection at a time. First-party drivers are designed to allow applications to monitor and/or control one extension at a time.

Telephone systems naturally permit many of these connections simultaneously. Modem connections are by nature first-party.

TAPI 2.x vs TAPI 3.x

It is a common misconception that TAPI 3.0 (or TAPI 3.1) replaces TAPI 2.x.

TAPI 2.x (and all earlier versions) is written in C/C++ and requires applications to make heavy use of C style pointer arithmetic. This makes TAPI 2.x fast and easy to access from C/C++ applications, but it also makes it difficult to use from many other programming languages.

TAPI 3.x was designed with a COM (Component Object Model) interface. This was done with the intent of making it accessible from managed languages such as C# or other environments that provide easy access to COM but don't deal with C-style pointers.

TAPI 3.x has a slightly different set of functionality than TAPI 2.x. The addition of integrated media control was the most significant addition. But TAPI 3.x doesn't include all functionality that TAPI 2.x does, like support for the Phone class.

One very notable issue with TAPI 3.x is the lack of support for managed code (.NET environment). As documented in Microsoft KB Article 841712, Microsoft currently has no plans to support TAPI 3.x directly from .Net programming languages. However, Mark Smith has provided a Managed C++ library called ITAPI3.

One often overlooked reason an application developer might choose between TAPI 2.x and TAPI 3.x should be the hardware vendors recommendation. Even though TAPI provides an abstract model of phone lines, telephony applications are still heavily impacted by the specific behavior of the underlying hardware. Troubleshooting behavior issues usually requires both software and hardware vendors to collaborate. Because there is almost a 1:1 relationship between the TAPI Service Provider (TSP) interface and the TAPI 2.x interface, collaboration is often easier if the application is designed using TAPI 2.x. Experience with TAPI 3.x varies significantly between hardware vendors .

Fig: Example of TAPI

9.2 **Mobile Operating Systems (Mobile OS)**

Like a computer operating system, a mobile operating system is the software platform on top of which other programs run. When you purchase a mobile device, the manufacturer will have chosen the operating system for that specific device. The operating system is responsible for determining the functions and features available on your device, such as thumbwheel, keyboards, WAP, synchronization with applications, e-mail, text messaging and more. The mobile operating system will also determine which third-party applications can be used on your device.

Mobile computing has never been so popular. Now with Microsoft Windows Mobile you can run multimedia, manage storage, communicate and run office applications via the mobile platform. Windows Mobile also delivers interfacing with such outstanding services as Windows Live and Opera Web Browsing.

Since the early 1990's technology has evolved to such an degree that computing and communications are now possible from the smallest, pocket sized devices, while delivering fantastic visuals and speed. In order to embrace this shift in technological capability Microsoft has been at the forefront of software development, creating applications and operating systems that fit the specific requirements of the mobile device. One such operating system is

Windows Mobile. In essence a basic os, based on the Microsoft Win32 API, Windows Mobile was designed specifically for handheld devices, originally the Pocket PC, then Smartphone's and Portable Media Centers. In its earliest form the windows mobile operating system was similar in look, feel and functionality to Windows 98, delivering basic application delivery and minimal third-party support for common hardware and software components.

Today however Windows Mobile is more in tune with the 'Live' suite Microsoft are pushing, which has a dazzling array of onscreen informatics - e-mail messages, tasks, apointments and ownership details. As with Windows XP, the taskbar holds the current time, volume connectivity status, and resource processing. The notification bar holds all the standard desktop iconic representations such as running programs, connectivity, etc. In terms of application productivity the Mobile Office suite delivers impressively similar functionality to its big brother Microsoft Office. Word, Excel, Outlook and PowerPoint all deliver excellent functionality and while there are some gaps in functionality, for example table and image insertion are two areas that are missing, you won't see much of a drop off. Windows Media Player also has extensive operability providing users with the capability to play WMA, WMV, MP3 and AVI files, however MPEG's are not supported and WAV files require a separate player. There is also an abundance of personalization settings, from configuring background images and themes. Most importantly however is the development of ActiveSync and improved server interfacing for fast and secure synchronization of data to any desktop application.

The Windows Mobile operating system is now available on multiple platforms, but the first devices to use Windows Mobile were Pocket PC's. Driven by Pocket PC 2000 operating system and powered by Windows CE 3.0. It was soon realized that such an operating system brought real value, and with successful integration into handheld devices it became the os of choice on many Pocket PCs. Being the first release brought the obvious limited

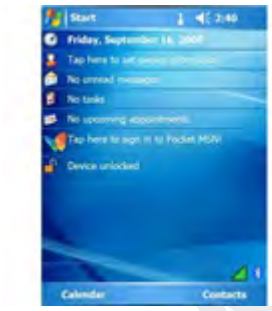
functionality and problems, the most limiting being Pocket PC 2000's ability to only support screens with a resolution of 240 x 320. Other than that, the operating system gave great encouragement for the future of mobile computing os delivery with such functionality as Pocket Office, Pocket Internet Explorer, Windows Media Player, Microsoft Money, Microsoft Reader, Notes, Character Recognition and Infrared file beaming.

Following Pocket PC 2000 came the enhanced operability and Smartphone compatibility of Pocket PC 2002. These Smartphone devices were developed to aid user interactions via one hand, bringing mobile computing to handheld communications devices for the first time. While Pocket PC 2002 still didn't resolve the issue of only being able to run at 240 x 320 resolution it did have improved navigability and slicker operability. This enhanced user interface also incorporated greater configurability, GSM communications and enhanced Pocket Office tools, bringing it more in-line with desktop versions. There was also an extension to the services on Pocket PC 2000, including virtual private networking, synchronization, MSN messenger and digital rights management. All of which were available for the first time via Pocket PC and now Smartphone's.

As Smartphone's advanced to incorporate greater multimedia capabilities, advancements were made in the area of Portable Media Centers. As such Microsoft focused its integration with Windows Media Center, Windows Media Player, and developed the Windows Mobile 2.3 operating system architecture. This enabled users to carry extensive media libraries and manage multimedia via a simple operating system, on one unit. Windows Mobile 2003 introduced the first multi-version consumer choice mobile operating system platforms. These were Windows Mobile 2003 for Pocket PC Premium Edition, Phone Edition, and Professional Edition, and Windows Mobile 2003 for Smart Phones. Catering for such a wide range of handheld devices, the Windows Mobile operating system was making considerable strides into the handheld market. With such a range of operating systems flexibility came enhanced communication support, keyboard facilities, games, image delivery, MIDI file support, messaging and further enhancements to the Pocket Office suite.

With the added diversity of hardware manufacturers, Microsoft then released the 2nd edition of Windows Mobile 2003. This brought enhanced viewing on VGA screens of 240 x 240 and 480 x 480 resolution, Wi-Fi security and an enhanced look and feel which included for the first time portrait or landscape viewing options.

As more hardware devices came to market, offering greater multimedia, communications and office type productivity there became a greater need for operating system flexibility for working (and playing) on the go. Thus Microsoft developed Windows Mobile Automotive. Enabling communications, entertainment and information systems to be interlaced for seamless operability on one sophisticated device. Having a need for such functionality in transit dictated the need for two different versions of the Mobile Automobile, the basic version had USBinterfacing and Bluetooth connectivity, whereas the standard version also had built-in GPS, GSM and greater security. Windows Mobile did require a higher spec processor, memory and microphone but it was a considerable step towards mobile computing for the roaming user.



In 2005 Microsoft launched Windows Mobile 5.0. Driven be Windows CE 5.0 and incorporating a .Net compact framework it delivered further compatibility with communications infrastructures and the first breakthrough cooperation with Microsoft Exchange Server architecture. This created tremendous potential for data synchronization, compatibility and greater capacity for storage. Which in turn saw major upgrades in improved battery life through Persistent Storage capacity where flash memory is used for primary storage as opposed to the previous memory intensive volatile RAM which had been used in the past.

Also available via Windows Mobile 2005 were major improvements to application software, such as the introduction of ‘Mobile Office’. Along with Word, Excel and Outlook, the Mobile Office suite also incorporated PowerPoint, enhanced graphics and tables. Windows Mobile 2005 also offered photo caller ID, greater communication support for Bluetooth and GPS, default

keyboard support, error reporting, ActiveSync and the aforementioned persistent storage for enhanced battery life.
Now in 2008 the current version is Windows Mobile 2006. Coming in three different iterations there is Windows Mobile 6 Standard for the Smartphone market, Windows Mobile 6 Professional for the Pocket PC with mobile phone capabilities, and Windows Mobile 6 Classic for the Pocket PC without mobile phone capabilities. In the next part we will look closely at Windows Mobile 6 and its development.

Some of the more common and well-known Mobile operating systems include the following:

Symbian OS
Symbian OS has become a standard operating system for smartphones, and is licensed by more than 85 percent of the world's handset manufacturers. The Symbian OS is designed for the specific requirements of 2.5G and 3G mobile phones.

Windows Mobile
The Windows Mobile platform is available on a variety of devices from a variety of wireless operators. You will find Windows Mobile software on Dell, HP, Motorola, Palm and i-mate products. Windows Mobile powered devices are available on GSM or CDMA networks.

Palm OS
Since the introduction of the first Palm Pilot in 1996, the Palm OS platform has provided mobile devices with essential business tools, as well as capability to access the Internet or a central corporate database via a wireless connection.

Mobile Linux:
The first company to launch phones with Linux as its OS was Motorola in 2003. Linux is seen as a suitable option for higher-end phones with powerful processors and larger amounts of memory.

MXI
MXI is a universal mobile operating system that allows existing full-fledged desktop and mobile applications written for Windows, Linux, Java, Palm be enabled immediately on mobile devices without any redevelopment. MXI allows for interoperability between various platforms, networks, software and hardware components.

Chapter 10

10.1 J2ME

10.2 MIDP

10.1 J2ME

What's J2ME?

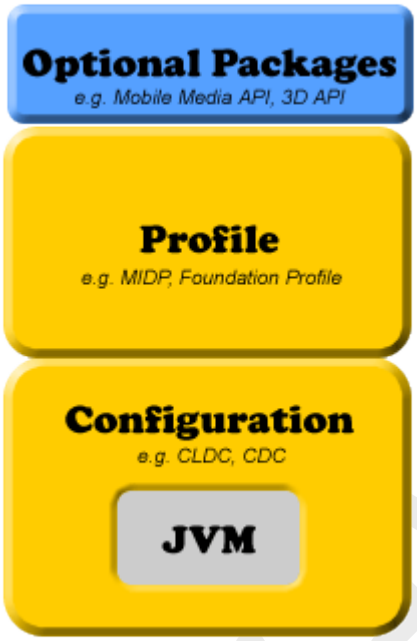


Figure 1: J2ME Stack

Java 2 Micro Edition (J2ME) is Sun's version of Java aimed at machines with limited hardware resources such as PDAs, cell phones, and other consumer electronic and embedded devices. J2ME is aimed at machines with as little as 128KB of RAM and with processors a lot less powerful than those used on typical desktop and server machines. J2ME actually consists of a set of profiles. Each profile is defined for a particular type of device -- cell phones, PDAs, microwave ovens, etc. -- and consists of a minimum set of class libraries required for the particular type of device and a specification of a Java virtual machine required to support the device. The virtual machine specified in any profile is not necessarily the same as the virtual machine used in Java 2 Standard Edition (J2SE) and Java 2 Enterprise Edition (J2EE). You'll see that the profile we'll use to develop a Palm OS device application is a subset of the Java Virtual Machine you already know.

To date, Sun has released the following profiles:

The Foundation Profile -- A profile for next generation consumer electronic devices

The Mobile Information Device Profile (MIDP) -- A profile for mobile information devices, such as cellular phones and two-way pagers, and PDAs

A profile in itself does not do anything; it just defines the specification. Profiles are implemented with a configuration. You can think of a configuration as an implementation of a J2ME profile for a particular type of device such as a PDA. Some of the configurations currently available are

Connected Device Configuration (CDC)

An implementation of the Foundation Profile for next-generation, consumer electronic and embedded devices

Connected Limited Device Configuration (CLDC)

An implementation of MIDP for small, resource-constrained devices such as Palm OS devices.

Since each profile defines a different set of Java class libraries, you cannot take a Java application written for one profile and run it on a machine that supports another profile. Likewise, you cannot take an application written for Java 2 Standard Edition (J2SE) or Java 2 Enterprise Edition (J2EE) and run it on a machine that supports J2ME. You can only use the Java classes provide in the Java class library included in your target device's profile. Restricting yourself from using all of the Java classes you've grown to rely on is one of the hardest parts of writing Java applications for small devices.

The rest of this article focuses on using the MIDP profile and CLDC configuration to create an application for a PDA running the Palm operating system.

Getting J2ME

The J2ME CLDC comes as two files that you can download from the Sun Community site. The first file is j2me_cldc_-1_0-src.winsol.zip and contains all of the class libraries needed to develop CLDC applications on your computer. In addition to containing the class libraries, it also includes both source code and Windows and Unix binaries for a reference implementation of the Java Virtual Machine specified by MIDP. This Java Virtual Machine is referred to as the K Virtual Machine (KVM), and its inclusion saves you the trouble downloading your applications to a handheld device during development. You can test them right on the same desktop machine you use for development.

The second file available from the Sun Community site is j2me_cldc-1_0-src-palm_overlay.zip, which contains an implementation of the KVM for devices running the Palm operating system, along with the tools to turn the .class file generated by a Java compiler into an executable Palm file. Note that this is a particular implementation of the CLDC for the Palm operating system. As other implementations become available, your Java application will be able to run on those other machines as well.

You'll need to install both of these ZIPs on your computer to develop applications for Palm OS devices. For the rest of this article, I assume you have these installed in the CLDC folder on the C: drive.

Note that you do not need a special Java compiler in order to develop CLDC applications. You use the compiler that comes with J2SE. The only difference is that you'll specify an alternative location for the class library during compile. You'll also need to process the resulting .class file with some special software included with the CLDC download.

Running the Sample Apps

The files you downloaded from Sun include both source and compiled versions of many sample applications. You can run these sample applications on your desktop computer using the KVM for Windows or Unix. Running these applications will let you get a feel of what a CLDC application is like. Before you can run these programs, you need to start a command shell and enter the following commands.

```
set cldc_classpath = c:\cldc\bin\api\classes;.

set bin = %bin%;c:\cldc\bin

cd \cldc\bin\samples\classes
```

Now that you have set up the environment variables and changed into the directory that holds the sample applications, you can start to run them. To run the UITest application, enter

```
kvm -classpath %cldc_classpath% UITest
```

Once you're done looking at the UITest application, you can take a look at the Pong game application by entering, at the the command line, the command

```
kvm -classpath %cldc_classpath% Pong
```

You can now go through and run each of the sample applications. Later you'll use the same method to run the sample application we'll develop on your desktop machine.

Understanding the Process of MIDlet Creation--Without the Toolkit

There are seven steps in the creation of a MIDlet. These steps are: designing, coding, compiling, preverification, packaging, testing, and deployment. Some of these steps are not strictly MIDlet-centric (for example, every application needs to be designed, coded, and compiled), but we will cover them here because there are MIDlet-centric differences. The Toolkit abstracts a lot of these steps so that it is easier for you in the overall scheme of things. This is fine and dandy once you know the process, but when you are only starting out, you really should be coding by hand, rather than using a sugar-coated abstraction.

To ensure that we get a hands-on understanding of these steps, let us take the help of a simple example. We will create a MIDlet that, when executed, will print the current date and time on a

mobile device for a short time. Along with this in mind, keep Figure 2 handy to understand the sequence of these steps. Also, note that I will explain the lifecycle of MIDlets later. For the moment, let's get a simple MIDlet up and running, which will illustrate these steps.

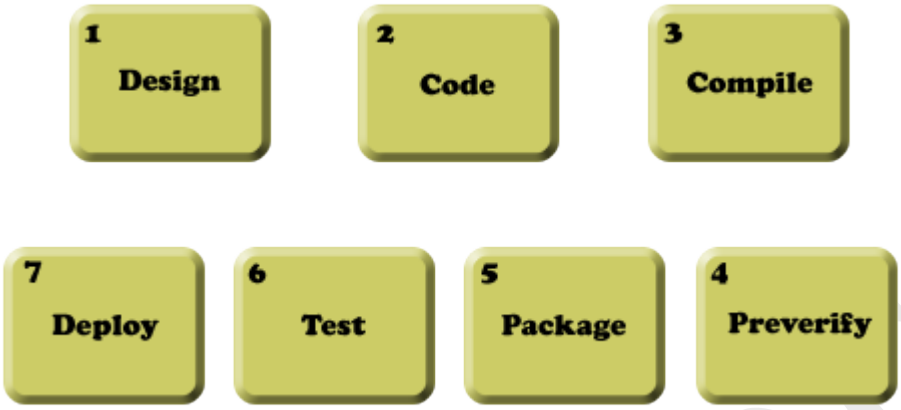


Figure 2. Steps to MIDlet creation

Step 1: Design

MIDlets are different from other applications that you may have created, simply because MIDlets run in an environment that is very different. There are several issues, not just those that are most visible (for example, the interactivity of your MIDlet with the user), but others that impact its usability.

For the example application, our Date-Time MIDlet does not need user interactivity. It needs to display the current date and time for a few seconds when the user executes the MIDlet. For simple cases like this, it is perhaps sufficient to mimic the design of the MIDlet by drawing it on a piece of paper. For more complex designs with multiple screens, it is best to design the screens professionally before starting the actual coding process.

Step 2: Code

Each MIDlet must extend the abstract MIDlet class found in the javax.microedition.midlet package, much like creating an applet by extending the java.applet.Applet class. At the minimum, your MIDlet must override three methods of this abstract class, startApp(), pauseApp(), anddestroyApp(boolean unconditional). Here is the DateTimeApp class:

```
package com.j2me.part1;

import java.util.Date;

import javax.microedition.lcdui.Alert;
import javax.microedition.lcdui.Display;
```

```
import javax.microedition.midlet.MIDlet;

public class DateTimeApp extends MIDlet {

  Alert timeAlert;

  public DateTimeApp() {
    timeAlert = new Alert("Alert!");
    timeAlert.setString(new Date().toString());
  }

  public void startApp() {
    Display.getDisplay(this).setCurrent(timeAlert);
  }

  public void pauseApp() {
  }

  public void destroyApp(boolean unconditional) {
  }
}
```

In this example, DateTimeApp's constructor creates the element that is necessary to display the time on a device's screen and the startApp method does the actual task of displaying this element. Don't worry if you don't understand how the Alert element works, or when the constructor or the other methods are called. I will cover the former in the next part, when we look at the GUI elements of MIDP 2.0, and the latter later in this article in the MIDlet Lifecycle section.

Copy this code into a file called *DateTimeApp.java* and save it in a folder that mimics its package structure (*com\j2me\part1*). You can save it anywhere you want on your machine; as far as this article is concerned, we will save it in the folder *C:\WTK22\article\com\j2me\part1*.

Step 3: Compile

With this simple code in place, you now need to know how to compile it so that it is ready for mobile devices. Compiling MIDlets is not very much different from compiling normal Java applications. You still use javac as the compiler, except you need to change the boot CLASSPATH while compiling MIDlets. This has the effect of changing the base Java classes that the Java compiler uses to compile your MIDlet against, thereby ensuring that compilation is targeted towards the narrow set of Java's APIs for the J2ME platform. So instead of compiling against the java.lang.Date in "normal" Java, you actually want compilation done for J2ME's java.lang.Date. This is done by pointing to the CLDC and MIDP classes for javac's -bootclasspath option while compiling. This is shown below for the DateTimeApp MIDlet compilation. To do this compilation, make sure you that you enter the command by navigating to the directory *C:\WTK22\article* via the command prompt.

```
C:\WTK22\article>javac -bootclasspath ..\lib\cldcapi11.jar;..\lib\midpapi20.jar
com\j2me\part1\DateTimeApp.java
```

Notice that I have done the compilation against the CLDC API's 1.1 and MIDP API's 2.0 versions, respectively, by including these libraries in thebootclasspath option. I could have done the compilation against other versions if it was required, by simply pointing to their respective libraries.

Step 4: Preverify

Before you can deploy your MIDlet class, it needs to be preverified. Verification of byte code is a step performed by the JVM before it runs any class file to ensure that the class file is structurally and conceptually correct as per the JVM specification. If the class file fails this check, it is rejected and the JVM shuts down, indicating either security or integrity violation of the class file. This verification is done by all JVMs, even the tiny JVM contained in a CLDC configuration for a J2ME device. Although this is not a problem for "normal" applications, verification in J2ME devices is a resource and memory constraint that they simply cannot handle (or should not handle). Therefore, the need for preverification.

Preverification is one part of a special two-step verification process, especially designed for constrained devices, such as the ones running CLDC-based JVMs. The idea is to let a developer preverify his classes, which limits the amount of work needed to be performed when the classes are verified in the device. This preverification process adds special information to the classes that identifies them as preverified and makes the process on the device much more efficient.

Keeping this in mind, preverify your Date-Time MIDlet. The Wireless Toolkit comes with a preverification tool in the bin folder of its installation (*C:\WTK22\bin*). The following command, when executed from *C:\WTK22\article*, will preverify the *DateTimeApp.class* created in the previous step.

```
C:\WTK22\article>..\bin\preverify.exe -classpath ..\lib\cldcapi11.jar;..\lib\midpapi20.jar
com.j2me.part1.DateTimeApp
```

By default, the preverifier will create the *preverified* version of your *DateTimeApp.class* file in a folder called *output* in the current directory. It will preserve the package structure, so your preverified class will now be in the folder *C:\WTK22\article\output\com\j2me\part1*. You can, of course, point the output to another folder, using the -d option for the preverify tool, but for the moment, use the default output folder.

Step 5: Package

Packaging your MIDlet so that it ready for testing and deployment is a fairly involved process, with several steps. Although each step is straightforward, they must be followed in proper sequence.

The first step is to create a Manifest file. This Manifest file describes the contents of the Java Archive (JAR) file that we will be creating in the next step. There are several attributes that can go in this file, but for your Date-Time MIDlet, stick to only the ones that are required. This file's contents are shown here:

MIDlet-Name: DateTimeApp
MIDlet-Version: 1.0.0
MIDlet-Vendor: Vikram Goyal

Save this file as *Manifest.mf* in the *C:\WTK22\article\output* folder. (Note the newline after the last attribute, MIDlet-Vendor. It **must** be present, otherwise this attribute will not be recognized.)

Next, create the JAR file that packages up the preverified *DateTimeApp.class* file and the Manifest file. To create this JAR file, navigate to the*C:\WTK22\article\output* directory and issue the following command:

C:\WTK22\article\output>jar cvfm DateTimeApp.jar Manifest.mf .\com

This will create the *DateTimeApp.jar* file in the current (*C:\WTK22\article\output*) folder.

The second-to-last step is to create a file that has an extension of .jad. A Java Application Descriptor (JAD) file points to the location of the MIDlet it describes so that a J2ME device can install it. Again, this file can contain several attributes for a single MIDlet (or for several MIDlets), but for your Date-Time MIDlet, you will stick with the ones that are required.

MIDlet-1: DateTimeApp, , com.j2me.part1.DateTimeApp
MIDlet-Name: DateTimeApp
MIDlet-Version: 1.0.0
MIDlet-Vendor: Vikram Goyal
MIDlet-Jar-URL: DateTimeApp.jar
MIDlet-Jar-Size:
MicroEdition-Profile: MIDP-2.0
MicroEdition-Configuration: CLDC-1.1

Save this file as *DateTimeApp.jad* in the same folder as the JAR file (*C:\WTK22\article\output*). I will explain the attributes in this file later, but for now, note that the value of the MIDlet-Jar-Size attribute is missing. This missing value brings you to the last step of the packaging step, where you determine the size of the *DateTimeApp.jar* file, and put that value in this JAD file, in actual bytes. It is very important to get this exactly right, as the installation of this MIDlet will

fail if this value is different from the actual size. On my machine, this value is 1469 bytes, and therefore, this is what this attribute looks like on my machine:

MIDlet-Jar-Size: 1469

This completes the packaging part. Well, actually, there are other steps in the packaging that I could talk about (for example, signing and obfuscation), but to keep things simple, I will leave those steps for later discussion. For now, you will move on to testing of your MIDlet.

Step 6: Test

Before deploying your MIDlets, they must be tested by using a base common emulator device that mimics the functionality of an actual device on your computer. This emulator is part of the Wireless Toolkit and provides functionality that is sure to be present in the majority of devices for which the MIDlet is targeted. This emulator is present in the bin folder of the Toolkit.

From the output directory created in the preverify step earlier, and where we now have a packaged MIDlet in the form of JAR and JAD files, issue the following command to run the emulator with this JAD file as an option.

C:\WTK22\article\output>..\..\bin\emulator.exe -Xdescriptor DateTimeApp.jad

You should see the emulator pop up on your screen as shown in Figure 3, with the DateTimeApp MIDlet selected. If it doesn't, the most likely error at this point would be incorrect JAR size information. Make sure you have the exact size listed in the JAD file.

Figure 3. Testing the DateTimeApp

At the lower right-hand corner of the emulated device's screen, you can see the "Launch" menu item listed. The emulator has installed the MIDlet and is ready to launch it. Click on the phone button just underneath that menu item and the MIDlet should display the current date time for a

few seconds and then disappear. Note that the MIDlet is still running even after the date and time disappear, because in code, you did not destroy it.

Step 7: Deploy

This is it! Now you have reached the stage where you can deploy the MIDlet directly on your mobile device. There are two ways to do this. The first is via a network connection between your computer and your handset. This can either be via a USB cable or a Bluetooth wireless connection, depending on your device. Most Java-enabled devices will allow you to install J2ME applications via this connection.

Second, and the one that is more interesting, because it opens up your MIDlet to the outside world, is via the Internet. After all, what good is your MIDlet if the rest of the world cannot see it? Of course, this means that your device should be able to connect to the Internet using its internal browser.

Before you proceed further, recall that when you created the JAD file, you entered two attributes in it that specified the version of CLDC (1.1) and MIDP (2.0) for which the MIDlet was created. Since the DateTimeApp MIDlet does not use any of the features of these versions, it should theoretically run on devices that support the lower versions of these attributes, as well. Therefore, the DateTimeApp MIDlet should run on CLDC 1.0 and MIDP 1.0, but because the JAD file restricts these versions to the newer ones, devices will fail to install this MIDlet if they do not support these new versions. If this is the case with your device, fear not! As I said before, because we are not using any MIDP-2.0- or CLDC-1.1-specific features, you can simply change these version numbers in the JAD file, and this will be sufficient to install this device on all Java-enabled devices. If this is the case with your device, or the device that you are going to test this MIDlet on, simply change these values in the JAD file and you are good to go.

To be able to deploy your MIDlet via the Internet, you need to have access to a web server with a real-world IP address or domain name. You also need to have administrative privileges to be able to modify the configuration files of your web server to add some Multipurpose Internet Mail Exchange (MIME) types for the JAD and JAR extensions. If you are using Jakarta/Tomcat as your web server, you don't need to do this, as it already has these MIME types. For the Apache web server, modify the *mime.types* file and add the following extension types.

text/vnd.sun.j2me.app-descriptor jad

application/java-archive jar

By adding these MIME types, you are informing the browser, or any client accessing these files from the server, how to handle these files when they are downloaded into the device.

Next, create an HTML file that will become the point of reference. Strictly, this is not necessary, because a device that can access an HTML page can also access a JAD file. But an HTML page provides a point of reference, and therefore, let's create one for your Date-Time MIDlet. The

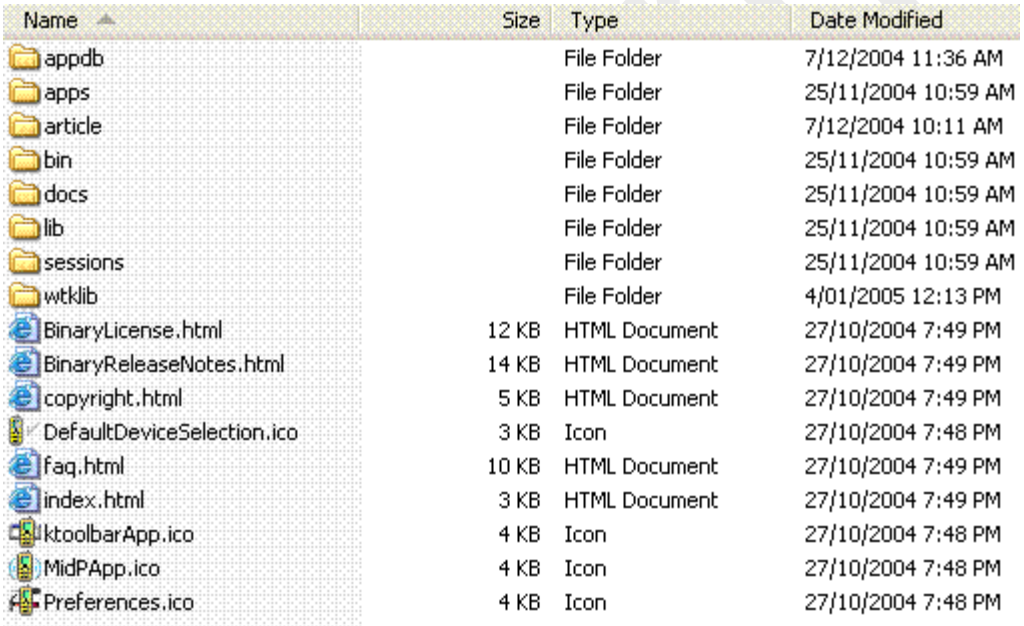
HTML doesn't need to be anything fancy. Don't forget that users will be accessing this page via a mobile device, so it is prudent to keep the size of this page to the minimum. This is shown in Listing 2.

```
<HTML>
  Click <a href="DateTimeApp.jad">here</a> to download DateTimeApp MIDlet!
</HTML>
```

Listing 2. DateTimeApp.html page for accessing the DateTimeApp MIDlet

Understanding the Process of MIDlet Creation--Using the Toolkit

In the section Acquiring and Installing J2ME Development Kit above, you downloaded the Toolkit and installed it in the folder *C:\WTK22* (as far as this article series is concerned; you may have downloaded and installed it in a different folder). Let's explore the contents of this folder. Figure 4 shows these contents as they should now look on your machine.



Name	Size	Type	Date Modified
appdb		File Folder	7/12/2004 11:36 AM
apps		File Folder	25/11/2004 10:59 AM
article		File Folder	7/12/2004 10:11 AM
bin		File Folder	25/11/2004 10:59 AM
docs		File Folder	25/11/2004 10:59 AM
lib		File Folder	25/11/2004 10:59 AM
sessions		File Folder	25/11/2004 10:59 AM
wtklib		File Folder	4/01/2005 12:13 PM
BinaryLicense.html	12 KB	HTML Document	27/10/2004 7:49 PM
BinaryReleaseNotes.html	14 KB	HTML Document	27/10/2004 7:49 PM
copyright.html	5 KB	HTML Document	27/10/2004 7:49 PM
DefaultDeviceSelection.ico	3 KB	Icon	27/10/2004 7:48 PM
faq.html	10 KB	HTML Document	27/10/2004 7:49 PM
index.html	3 KB	HTML Document	27/10/2004 7:49 PM
ktoolbarApp.ico	4 KB	Icon	27/10/2004 7:48 PM
MidPApp.ico	4 KB	Icon	27/10/2004 7:48 PM
Preferences.ico	4 KB	Icon	27/10/2004 7:48 PM

Figure 4. Wireless Toolkit folder contents

Note that the default installation of the Toolkit would **not** have created the *article* folder, and that you created it in the previous section.

As far as a MIDlet developer is concerned, the most important folders are the *apps* and *bin* folders, but here is a short summary of each of these folders.

Folder Name	Folder Description
appdb	Directory that acts as a simulation for mobile device file system
apps	MIDlets created using the Toolkit reside in this directory
bin	Contains executables for the various tools, including the Toolkit itself, and various other tools like the preverifier and the emulator
docs	The Wireless Toolkit documentation including API documentation for MIDP 2.0 and MIDP 1.1
lib	Contains the JAR files for MIDP (both 2.0 and 1.1), CLDC (both 1.1 and 1.0) and several other optional libraries
sessions	Directory where network and profiling sessions are maintained
wtklib	Contains the libraries for the Toolkit itself, including the properties of various device emulators

The *apps* folder is the directory where all the MIDlets that are created using the Toolkit are installed. Browse this folder, and you will notice several example MIDlets provided in their own folders. These have their own directory structure that allows clean separation of source code, libraries, and rest of the files associated with a MIDlet project.

The *bin* folder contains the executables for the Toolkit. The most important one is ktoolbar.exe (on Windows), which starts the main interface window for the Toolkit. This folder contains other executables as well, some of which we came across earlier (for example, preverify.exe andemulator.exe). Let us, however, concentrate on using the Toolkit now by running the ktoolbar.exe from the *bin* folder. The Toolkit will start and you will get the window shown in Figure 5.

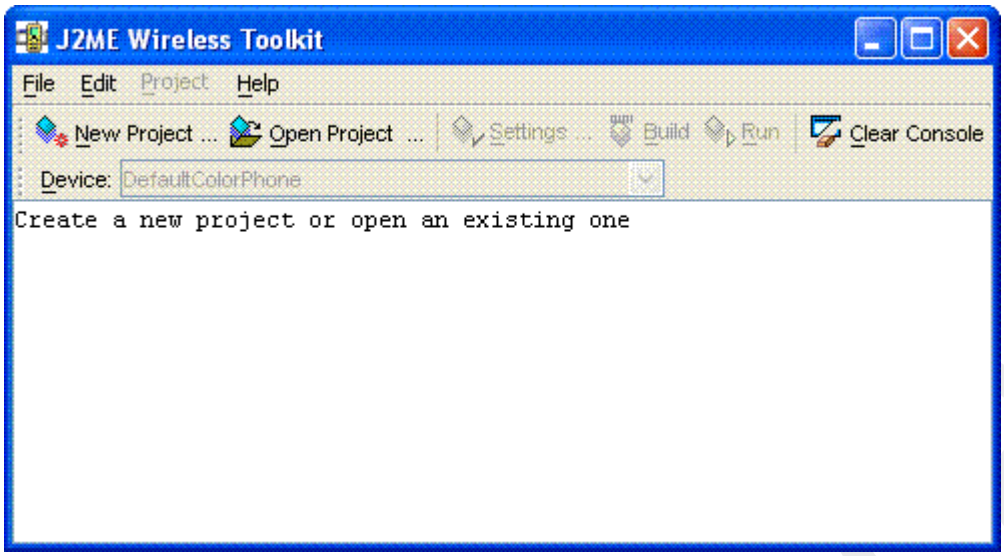


Figure 5. Main Toolkit window

As the message in the window says, from here, you can either create a new project or open an existing one. When you click on the Open Project menu button, you will be presented with a list of projects. As you may have guessed, this list of projects is the directory listing of the *apps* folder. Selecting a project from this list will open up the project and allow you to change its settings, build it (which includes compilation, preverification, and packaging) and run it. The steps of designing and coding are still to be done outside of this Toolkit.

Let's use the Toolkit to create the Date-Time MIDlet from the previous section. Click on New Project menu button, and enter the details in the window that comes up, as shown in Figure 6.

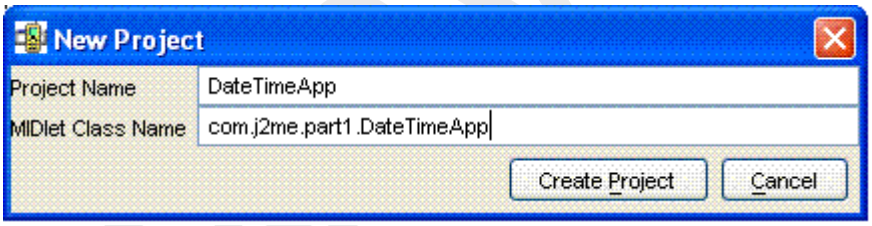


Figure 6. Creating a new project

The next window that comes up will allow you to change settings that control the target platform of your MIDlet. In this case, you want to target this MIDlet towards the MIDP 2.0 and CLDC 1.1 platforms and therefore, keep the Target Platform as JTWI, which preselects the MIDP 2.0 Profile. However, you will need to change the Configuration to CLDC 1.1 and uncheck the Optional Mobile Media API library, as shown in Figure 7.

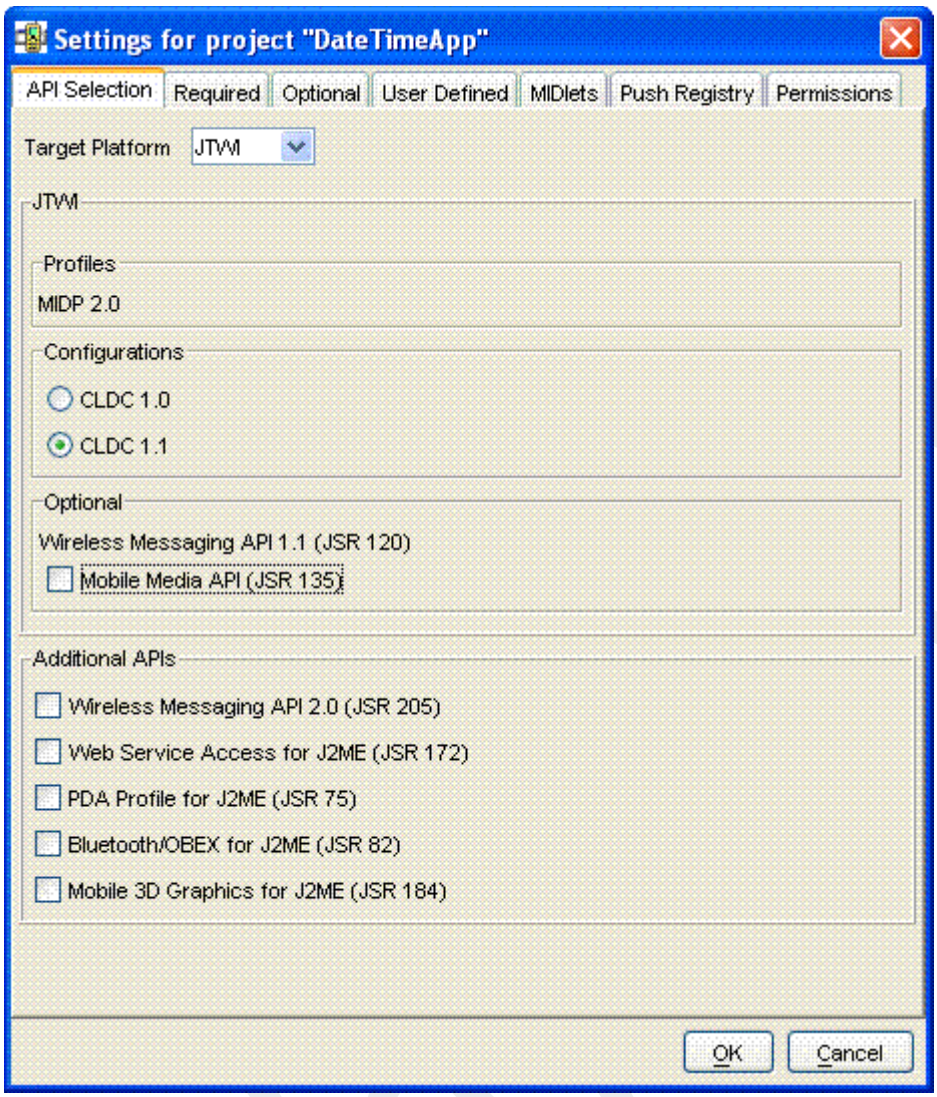


Figure 7. Changing project settings

You can review the rest of the settings by clicking on the tabs at the top, but for the moment, your project is ready to be created. Do so by clicking the OK button at the bottom. The project will be created with information about where to place the project files displayed on the screen, as shown in Figure 8. You can verify that the Toolkit has created a *DateTimeApp* folder under the *apps* folder by navigating to it.

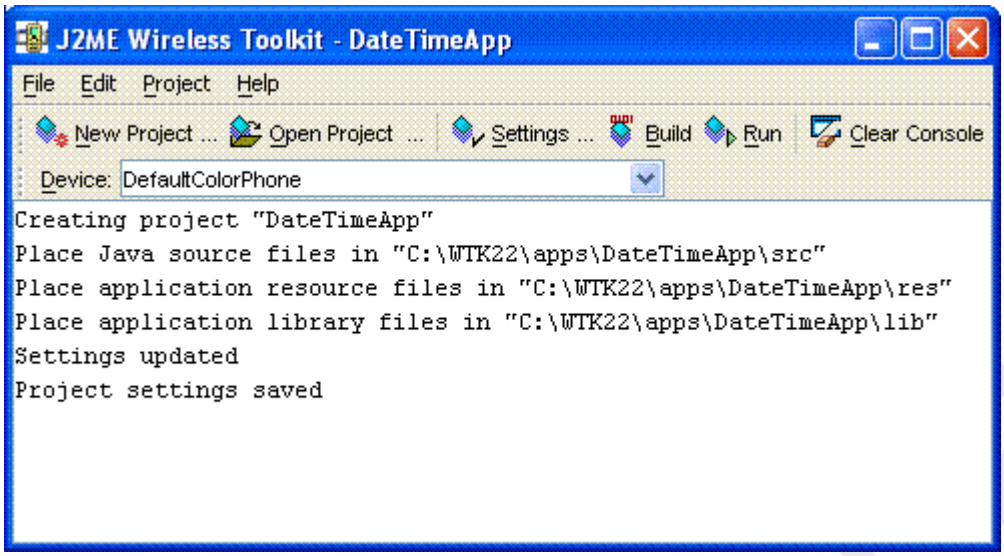


Figure 8. Project DateTimeApp created

You have already created the only required src file for this MIDlet in the previous section. Copy this file, *DateTimeApp.java*, from the folder*C:\WTK22\article\com\j2me\part1* to the fully qualified *src* folder (*C:\WTK22\apps\DateTimeApp\src\com\j2me\part1*). Note that the Toolkit created the fully qualified path based on the package name, so you don't have to. Once the copy is done, come back to the Toolkit, and hit the Run menu button. The Toolkit will compile, preverify, and package, and, provided everything goes OK, will run the DateTimeApp in the emulator. Seems simple enough, doesn't it? All you had to do was to create a new project, set the settings, write the code, drop it in the right directory, and hit the Run button. The Toolkit took care of the rest.

Before you leave this section, examine the rest of the folders under the DateTimeApp project. The *bin* folder under the *DateTimeApp* folder contains the JAD and the Manifest files, while the *classes* folder contains compiled classes. But where is the JAR file for this MIDlet? Well, the JAR file is not created by just running (or building) your application in the Toolkit. To create the JAR file, you will need to select the Project menu item, and then select one of the options under the Package submenu, as shown in Figure 9.

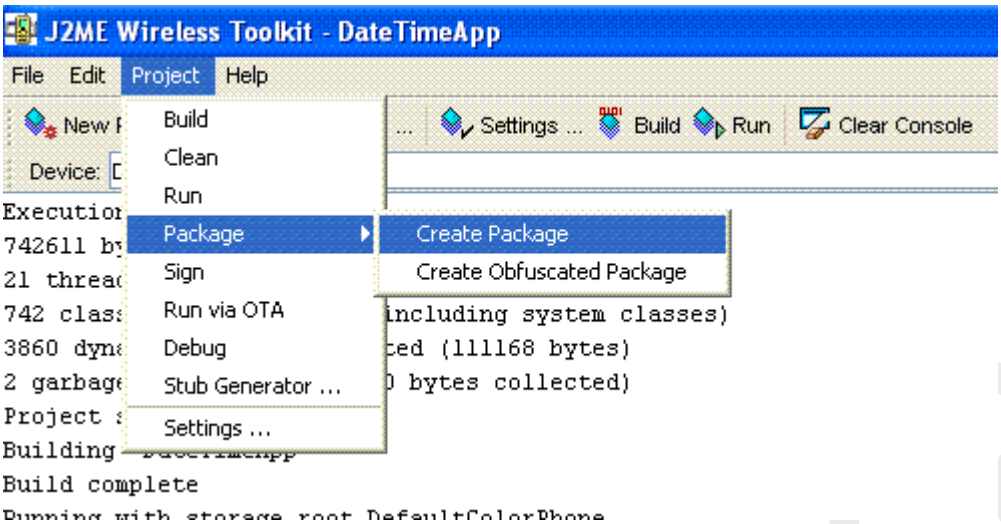


Figure 9. Creating the MIDlet's JAR file

By creating this package, the JAR file will be created, with correct Manifest information in the JAD file. You can even create the HTML file that we created in the deploy section previously, by clicking on the Run via OTA (Over The Air) menu. This will not only allow you to simulate your emulator running this MIDlet via the Internet, but also create the HTML file for you in the *bin* folder. Before you can use this HTML file to deploy on your own server, along with the JAD and the JAR files, you will need to change the hostname, which defaults to "localhost."

You now know how to create a simple MIDlet, both using the Toolkit, and without it. It's now time to look at the MIDlet lifecycle to understand what actually happens when your MIDlet is deployed and run.

The MIDlet Lifecycle

Mobile devices, whether emulators or real, interact with a MIDlet using their own software, which is called Application Management Software (AMS). The AMS is responsible for initializing, starting, pausing, resuming, and destroying a MIDlet. (Besides these services, AMS may be responsible for installing and removing a MIDlet, as well.) To facilitate this management, a MIDlet can be in one of three states which is controlled via the MIDletclass methods, that every MIDlet extends and overrides. These states are active, paused and destroyed.

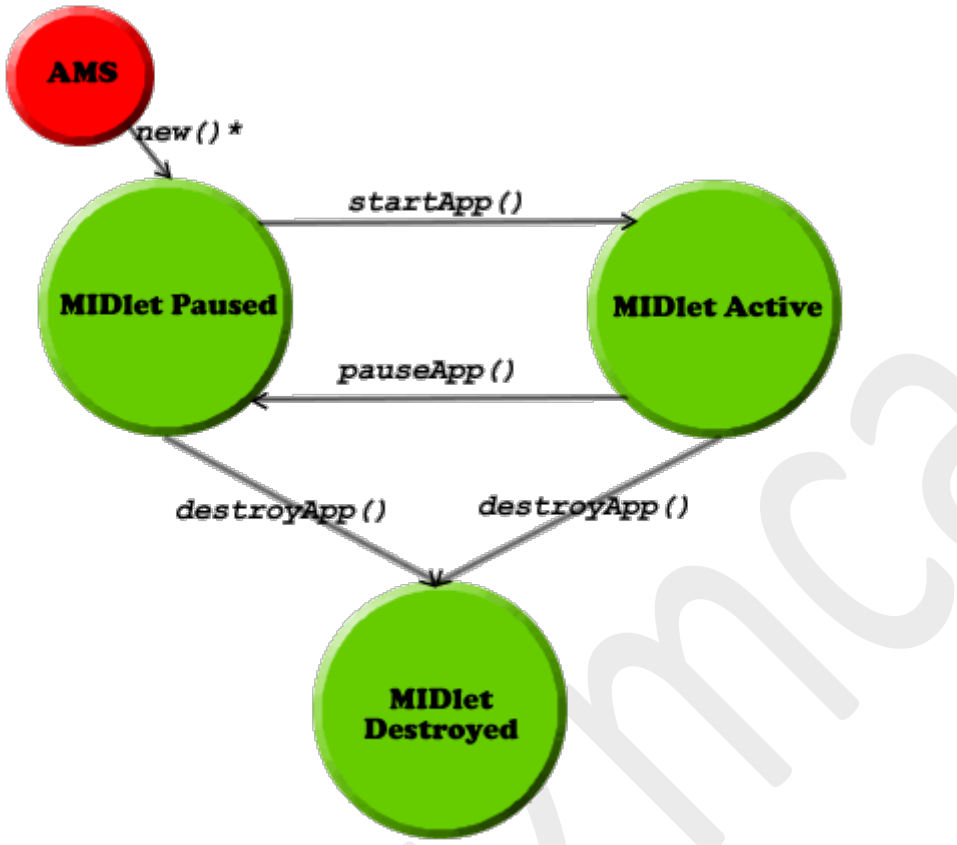


Figure 10. The possible states of a MIDlet and the transition between them

As you can see from Figure 11, an installed MIDlet is put into a paused state by the AMS creating an instance of it, by calling its no-args constructor. This is of course, not the only way that the MIDlet can be in a paused state. It can enter this state when the AMS calls the pauseApp() method on an active MIDlet (and the method returns successfully). It can also enter this state when the MIDlet pauses itself by calling the notifyPaused() method, as opposed to the pauseApp() method, which is called by the AMS. However, what exactly is happening with the MIDlet in the paused state?

In a paused state, the MIDlet is waiting for a chance to get into the active state. Theoretically, in this state, it should not be holding or using any of the device resources and should be passive in nature. Once the MIDlet is created, this is the state to be in before becoming active. Also, entering the paused state is necessary when the device requires it to consume fewer resources, because these resources may be required for handling other device functions, like handling an incoming call. This is when the device invokes the pauseApp() method through the AMS. If the MIDlet should inform the AMS that it has paused, it should invoke the notifyPaused() method, which tells the AMS that the MIDlet has indeed paused.

One final way in which a MIDlet can get into a paused state is when the MIDlet's startApp() method, which is called when the AMS invokes it to start the MIDlet (either the first time or from a paused state), throws a MIDletStateChangeException. Essentially, in case of an error, the MIDlet takes the safe road of staying in the paused state.

The active state is where every MIDlet wants to be! This is when the MIDlet can do its functions, hold the device resources and generally, do what it is supposed to do. As said previously, a MIDlet is in an active state when the AMS calls the startApp() method on a paused MIDlet (actually, the MIDlet enters the active state just *before* this method is called by the AMS). A paused MIDlet can request to go into the active state by calling the methodresumeRequest(), which informs the AMS that the MIDlet wishes to become active. The AMS may of course, choose to ignore this request or, alternatively, queue it if there are other MIDlets requesting the same.

The destroyed state is entered when a MIDlet's destroyApp(boolean unconditional) method is called and returns successfully, either from an active or paused state. This method is called by the AMS when it feels that there is no need for the MIDlet to keep running and is the place the MIDlet may perform cleanup and other last minute activities. The MIDlet can enter this state itself, by calling the notifyDestroyed() method, which informs the AMS that the MIDlet has cleaned up its resources and is eligible for destruction. Of course, since in this case, the destroyApp(boolean unconditional) method is not called by the AMS, any last-minute activities must be done before this method is invoked.

What happens if the AMS calls the destroyApp(boolean unconditional) method in the middle of an important step that the MIDlet may be doing, and may be loath to be destroyed? This is where the Boolean unconditional flag comes into the picture. If this flag is set to true, the MIDlet will be destroyed, irrespective of what the MIDlet is doing. However, if this flag is false, effectively, the AMS is telling the MIDlet that it wants the MIDlet to be destroyed, but if the MIDlet is doing something important, it can raise a MIDletStateChangeException, and the AMS will not destroy it just yet. However, note that even then, there are no guarantees that the MIDlet will not be destroyed, and it remains up to each device to decide how they should handle the request. If the device does honor the MIDlet's request, it may try and invoke the destroyApp(boolean unconditional) at a later stage.

Note that a destroyed state means that the MIDlet *instance* has been destroyed, but not uninstalled from the device. The MIDlet remains installed in the device, and a new instance of it may be created later.

Let me end this section, and this article, with a flow chart of a typical sequence of events while using the DateTimeApp MIDlet that we created in the previous sections, and the corresponding AMS actions and MIDlet states. This flow chart is shown in Figure 11.

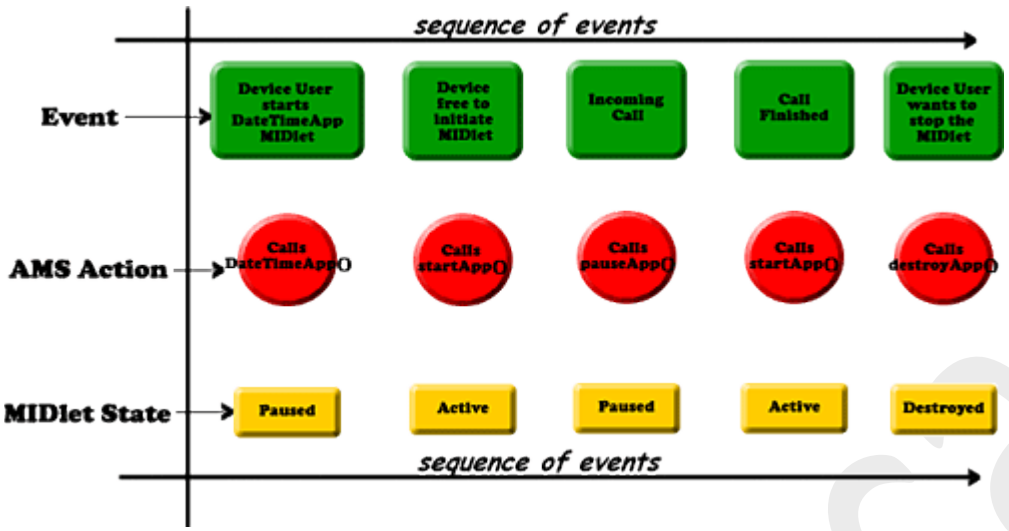


Figure 11. AMS actions and MIDlet states through a sequence of events

In the next part of this series, you will start creating useful MIDlets by understanding the User Interface API of MIDP 2.0. This will allow you to create powerful user interfaces, a key requirement for any MIDlet.

10.2 MIDP

Introduction

The Mobile Information Device Profile (MIDP) is a key element of the Java 2 Platform, Mobile Edition (J2ME). When combined with the Connected Limited Device Configuration (CLDC), MIDP provides a standard Java runtime environment for today's most popular mobile information devices, such as cell phones and mainstream personal digital assistants (PDAs). The MIDP specification was defined through the Java Community Process (JCP) by an expert group of more than 50 companies, including leading device manufacturers, wireless carriers, and vendors of mobile software. It defines a platform for dynamically and securely deploying optimized, graphical, networked applications.

CLDC and MIDP provide the core application functionality required by mobile applications, in the form of a standardized Java runtime environment and a rich set of Java APIs. Developers using MIDP can write applications once, then deploy them quickly to a wide variety of mobile information devices. MIDP has been widely adopted as the platform of choice for mobile applications. It is deployed globally on millions of phones and PDAs, and is supported by leading integrated development environments (IDEs). Companies around the world have already taken advantage of MIDP to write a broad range of consumer and enterprise mobile applications.

Specifications

- **MIDP 2.0** (JSR 118) is a revised version of the MIDP 1.0 specification. New features include an enhanced user interface, multimedia and game functionality, more extensive

connectivity, over-the-air provisioning (OTA), and end-to-end security. MIDP 2.0 is backward-compatible with MIDP 1.0, and continues to target mobile information devices like mobile phones and PDAs. More detailed information on these features can be found in What's New in MIDP 2.0.

- **MIDP 1.0** (JSR 37) is the original specification, which provides core application functionality required by mobile applications, including basic user interface and network security.

Reference Implementations

- MIDP RI 2.0 is based on the MIDP 2.0 specification and CLDC RI 1.0.4. The MIDP RI has been created for device manufacturers who want to port this J2ME profile to another platform.
- MIDP RI 1.0.3 is based on the MIDP 1.0 specification and CLDC RI 1.0.3. Its target audience is manufacturers who want to implement MIDP 1.0 on their devices.

Technology Compatibility Kit

The MIDP TCK can be licensed from Sun to certify a CLDC/MIDP implementation on a particular platform. For information about licensing terms, please contact Sun's Java Partner Engineering group.

Development Tools

- The Sun Java Wireless Toolkit provides complete development support for developing MIDP applications, and works in combination with today's leading IDEs.
- Sun ONE Studio, Mobile Edition is a Java-technology IDE for developing applications that can be deployed to Java technology-enabled mobile devices.

Optimized Implementation

MIDP for Palm OS is a J2ME implementation based on the CLDC 1.0 and MIDP 1.0 specifications, optimized for Palm OS 3.5.

Benefits

- **Rich User Interface Capabilities:** MIDP applications provide the foundation for highly graphical and intuitive applications. The graphical user interface is optimized for the small display size, varied input methods, and other native features of modern mobile devices. MIDP provides intuitive navigation and data entry by taking full advantage of phone keypads, extra buttons such as arrow keys, touch screens, and small QWERTY keyboards. MIDP applications are installed and run locally, can operate in both networked and unconnected modes, and can store and manage persistent local data securely.

- **Extensive Connectivity:** MIDP enables developers to exploit the native data network and messaging capabilities of mobile information devices. It supports leading connectivity standards, including HTTP, HTTPS, datagrams, sockets, server sockets, and serial port. MIDP also supports the Short Message Service and Cell Broadcast Service capabilities of GSM and CDMA networks, through the Wireless Messaging API (WMA) optional package.

- **Multimedia and Game Functionality:** MIDP is ideal for building portable games and multimedia applications. A low-level user-interface API complements the high-level UI API, giving developers greater control of graphics and input when they need it. A game API adds game-specific functionality, such as sprites and tiled layers, which take advantage of devices' native graphics capabilities. Built-in audio provides support for tones, tone sequences, and WAV files. In addition, developers can use the Mobile Media API (MMAPI) optional package to add video and other rich multimedia content to MIDP applications.

- **Over-the-Air-Provisioning:** A major benefit of MIDP is its capability to deploy and update applications dynamically and securely, over the air.

- **End-to-End Security:** MIDP provides a robust security model that complies with open standards and protects the network, applications, and mobile information devices. HTTPS support enables applications to use existing standards such as SSL and WTLS to send and receive encrypted data.