

Chapter1

Introduction

1.1 Broadband Wireless

Broadband wireless sits at the confluence of two of the most remarkable growth stories of the telecommunications industry in recent years. Both wireless and broadband have on their own enjoyed rapid mass-market adoption. Staggering growth of the Internet is the driving demand for higher-speed Internet-access services, leading to a parallel growth in broadband adoption. As a result, broadband subscription worldwide has grown from virtually zero to over 200 million in a decade.

Digital Subscriber Line (DSL) technology, which delivers broadband over twisted-pair telephone wires, and Cable Modem Technology which delivers over coaxial cable TV plant are the predominant mass-market broadband access technologies today. Both of these technologies typically provide up to a few megabits per second of data to each user, and continuing advances are making several tens of megabits per second possible. Since their initial deployment in the late 1990s, these services have enjoyed considerable growth. The United States has more than 50 million broadband subscribers, including more than half of home Internet users. Worldwide, this number is more than 200 million today and is projected to grow to more than 400 million by 2010. The availability of a wireless solution for broadband could potentially accelerate this growth. [1]

Broadband users worldwide are finding that it dramatically changes how we share information, conduct business, and seek entertainment. Broadband access not only provides faster Web surfing and quicker file downloads but also enables several multimedia applications, such as real-time audio and video streaming, multimedia conferencing, and interactive gaming. Broadband connections are also being used for voice telephony using Voice-over-Internet Protocol (VoIP) technology. More advanced broadband access systems, such as Fiber-to-The-Home (FTTH) and Very high Data rate

digital Subscriber Loop (VDSL) enable such applications as entertainment-quality video, including High-Definition TV (HDTV) and Video on Demand (VoD). As the broadband market continues to grow, several new applications are likely to emerge, and it is difficult to predict which ones will succeed in the future.

Broadband wireless can be defined as a wireless communication media which offer users certain unique benefits and convenience. There are two fundamentally different types of broadband wireless services. The first type attempts to provide a set of services similar to that of the traditional fixed-line broadband but using wireless as the medium of transmission. This type, called *fixed wireless broadband*, can be thought of as a competitive alternative to DSL or cable modem. The second type of broadband wireless, called *mobile broadband*, offers the additional functionality of portability, nomadicity and mobility. Mobile broadband attempts to bring broadband applications to new user experience scenarios and hence can offer the end user a very different value proposition. WiMAX (Worldwide Interoperability for Microwave Access) technology is designed to accommodate both fixed and mobile broadband applications.

The history of broadband wireless as it relates to WiMAX can be traced back to the desire to find a competitive alternative to traditional wireline-access technologies. During the past decade or so, a number of wireless access systems have been developed, mostly by start-up companies motivated by the disruptive potential of wireless. These systems varied widely in their performance capabilities, protocols, frequency spectrum used, applications supported, and a host of other parameters. Successful deployments have so far been limited to a few niche applications and markets. Clearly, broadband wireless has until now had a checkered record, in part because of the fragmentation of the industry due to the lack of a common standard. The emergence of WiMAX as an industry standard is expected to change this situation.

WiMAX is one of the most emerging technologies for BWA in metropolitan areas by providing an exciting addition to the current broadband techniques for the last-mile access. It is demonstrated that WiMAX is a viable alternative to the cable modem and DSL technologies due to its high resource utilization, easy implementation and low cost.

Furthermore, WiMAX not only enhances the existing features of the competitive cabled access networks, but provides high data rate applications with a variety of Quality of Service (QoS) requirements. WiMAX phones and PDAs are already being developed by companies like Motorola and Samsung and they are already in use in Korea, with WiMAX cousin technology, WiBro (Wireless Broadband) [2]. The bandwidth allocation algorithms have been designed for the efficient utilization of the scarce radio resources. In addition, to support multimedia traffics, the Medium Access Control (MAC) protocols will co-ordinate the transmission of traffic flows. The channel characteristics of users and traffic flow requirements are largely diverse, motivating us to design an efficient MAC layer protocols that can improve the system performance due to the channel and traffic dynamics.

1.2 Evolution of WiMAX Technology

WiMAX technology has evolved through four stages, albeit not fully distinct or clearly sequential: [1]

- narrowband wireless local-loop systems
- first-generation Line-of-Sight (LOS) broadband systems
- second-generation Non-Line-of-Sight (NLOS) broadband systems
- Standards-based broadband wireless systems.

1.2.1 Narrowband Wireless Local-Loop Systems

The first application for which a wireless alternative was developed and deployed was voice telephony. These systems, called Wireless Local-Loop (WLL), were quite successful in developing countries such as China, India, Indonesia, Brazil, and Russia, whose high demand for basic telephone services could not be served using existing infrastructure. In fact, WLL systems based on the Digital-Enhanced Cordless Telephony (DECT) and Code Division Multiple Access (CDMA) standards continue to be deployed in these markets.

1.2.2 First-Generation Broadband Systems

As DSL and cable modems began to be deployed, wireless systems had to evolve to support much higher speeds to be competitive. Systems began to be developed for higher frequencies, such as the 2.5GHz and 3.5GHz bands. In the late 1990s, one of the more important deployments of wireless broadband happened in the so-called Multichannel Multipoint Distribution Services (MMDS) band at 2.5GHz. The MMDS band was historically used to provide wireless cable broadcast video services, especially in rural areas where cable TV services were not available.

The first generation of these fixed broadband wireless solutions was deployed using the same towers that served wireless cable subscribers. These towers were typically several hundred feet tall and enabled LOS coverage to distances up to 35 miles, using high-power transmitters. First-generation MMDS systems required that subscribers install at their premises outdoor antennas high enough and pointed toward the tower for a clear LOS transmission path. Sprint and MCI launched two-way wireless broadband services using first-generation MMDS systems in a few markets in early 2000.

1.2.3 Second-Generation Broadband Systems

Second-generation broadband wireless systems were able to overcome the LOS issue and provide more capacity. This was done through the use of a cellular architecture and implementation of advanced-signal processing techniques to improve the link and system performance under multipath conditions. NLOS problem was solved by using such techniques as Orthogonal Frequency Division Multiplexing (OFDM), Code Division Multiple Access (CDMA), and multiantenna processing. A few megabits per second throughput over cell ranges of a few miles had become possible with second-generation fixed wireless broadband systems.

1.2.4 Emergence of Standards-Based Technology

In 1998, the Institute of Electrical and Electronics Engineers (IEEE) formed a group called 802.16 to develop a standard for what was called a wireless metropolitan area

network, or wireless MAN. Originally, this group focused on developing solutions for delivering high-speed connections to businesses that could not bear the fiber link. The IEEE 802.16 group produced a standard that was approved in December 2001. This standard, Wireless MAN-SC, specified a physical layer that used single-carrier modulation techniques and a Media Access Control (MAC) layer with a burst Time Division Multiplexing (TDM) structure that supported both Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD).

After completing this standard, the group started work on extending and modifying it to work in both licensed and license-exempt frequencies in the 2GHz to 11GHz range, which would enable NLOS deployments. This amendment, IEEE 802.16a, was completed in 2003, with OFDM schemes added as part of the physical layer for supporting deployment in multipath environments. By this time, OFDM had established itself as a method of choice for dealing with multipath for broadband and was already part of the revised IEEE 802.11 standards. Besides the OFDM physical layers, 802.16a also specified additional MAC-layer options, including support for Orthogonal Frequency Division Multiple Access (OFDMA).

1.3 Overview of WiMAX

After years of development and uncertainty, a standards-based interoperable solution is emerging for wireless broadband. A broad industry consortium, the WiMAX Forum [3] has begun certifying broadband wireless products for interoperability and compliance with a standard. WiMAX is based on Wireless Metropolitan Area Networking (WMAN) standards developed by the IEEE 802.16 group and adopted by both IEEE and the ETSI HIPERMAN group. In this section, we present a concise technical overview of the emerging WiMAX solution for broadband wireless.

As described above the IEEE 802.16 group was formed in 1998 to develop an air-interface standard for wireless broadband. 802.16a standard developed by this group covers NLOS applications in the 2GHz–11GHz band, using an Orthogonal Frequency Division Multiplexing (OFDM)-based physical layer. MAC layer improvements, such as

support for OFDMA, were also included. Further revisions resulted in a new standard in 2004, called IEEE 802.16-2004, which replaced all prior versions and formed the basis for the first WiMAX solution. These early WiMAX solutions based on IEEE 802.16-2004 targeted fixed applications, and we will refer to these as fixed WiMAX [4]. In December 2005, the IEEE group completed and approved IEEE 802.16e-2005, an amendment to the IEEE 802.16-2004 standard that added mobility support. The IEEE 802.16e-2005 forms the basis for the WiMAX solution for nomadic and mobile applications and is often referred to as mobile WiMAX [5]

The basic characteristics of the various IEEE 802.16 standards are summarized in Table 1.1 below.

 **Table 1.1 Basic Data on IEEE 802.16 Standards [6]**

Parameters	802.16	802.16-2004	802.16e-2005
Status	Completed December 2001	Completed June 2004	Completed December 2005
Frequency band	10GHz–66GHz	2GHz–11GHz	2GHz–11GHz for fixed; 2GHz–6GHz for mobile applications
Application	Fixed LOS	Fixed NLOS	Fixed and mobile NLOS
MAC architecture	Point-to-multipoint, mesh	Point-to-multipoint, mesh	Point-to-multipoint, mesh
Transmission scheme	Single carrier only	Single carrier, 256 OFDM or 2,048 OFDM	Single carrier, 256 OFDM or scalable OFDM with 128, 512, 1,024, or 2,048 subcarriers
Parameters	802.16	802.16-2004	802.16e-2005
Modulation	QPSK, 16 QAM, 64 QAM	QPSK, 16 QAM, 64 QAM	QPSK, 16 QAM, 64 QAM
Gross data rate	32Mbps–	1Mbps–75Mbps	1Mbps–75Mbps

	134.4Mbps		
Multiplexing	Burst TDM/TDMA	Burst TDM/TDMA/OFDMA	Burst TDM/TDMA/OFDMA
Duplexing	TDD and FDD	TDD and FDD	TDD and FDD
Channel bandwidths	20MHz, 25MHz, 28MHz	1.75MHz, 3.5MHz, 7MHz, 14MHz, 1.25MHz, 5MHz, 10MHz, 15MHz, 8.75MHz	1.75MHz, 3.5MHz, 7MHz, 14MHz, 1.25MHz, 5MHz, 10MHz, 15MHz, 8.75MHz
Air-interface designation	WirelessMAN-SC	WirelessMAN-SCa WirelessMAN-OFDM WirelessMAN-OFDMA WirelessHUMAN ^[a]	WirelessMAN-SCa WirelessMAN-OFDM WirelessMAN-OFDMA WirelessHUMAN ^[a]
WiMAX implementation	None	256 - OFDM as Fixed WiMAX	Scalable OFDMA as Mobile WiMAX

- a. WirelessHUMAN (Wireless high-speed unlicensed MAN) is similar to OFDM-PHY (physical layer) but mandates dynamic frequency selection for license-exempt bands.

For practical reasons of interoperability, the scope of the standard needs to be reduced, and a smaller set of design choices for implementation need to be defined. The WiMAX Forum does this by defining a limited number of system profiles and certification profiles. A *system profile* defines the subset of mandatory and optional physical- and MAC-layer features selected by the WiMAX Forum from the IEEE 802.16-2004 or IEEE 802.16e-2005 standard. It should be noted that the mandatory and optional status of a particular feature within a WiMAX system profile may be different from what it is in the original IEEE standard. Currently, the WiMAX Forum has two different system profiles [7]: one based on IEEE 802.16-2004, OFDM PHY, called the fixed system profile; the other one based on IEEE 802.16e-2005 scalable OFDMA PHY, called the mobility system profile. A *certification profile* is defined as a particular instantiation of a system

profile where the operating frequency, channel bandwidth, and duplexing mode are also specified. WiMAX equipments are certified for interoperability against a particular certification profile.

The WiMAX Forum has thus far defined five fixed certification profiles and fourteen mobility certification profiles. To date, there are two fixed WiMAX profiles against which equipment have been certified. These are 3.5GHz systems operating over a 3.5MHz channel, using the fixed system profile based on the IEEE 802.16-2004 OFDM physical layer with a point-to-multipoint MAC. One of the profiles uses FDD and the other uses TDD.

1.4 WiMAX Architecture Issues

Various architectural issues involved with WiMAX like-Point to Point (P2P) Vs Point to Multipoint (PMP), LOS Vs NLOS along with wireless radio are discussed in following text:

1.4.1 P2P Vs PMP

There are two scenarios for a wireless deployment: point-to-point and point-to-multipoint.

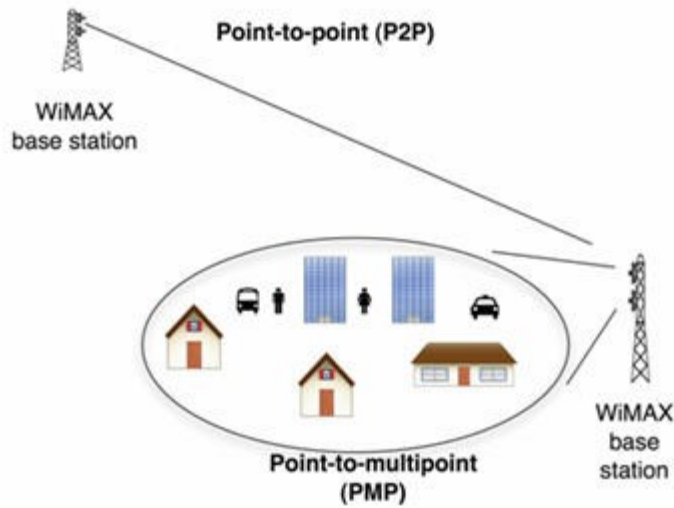


Figure 1.1: Point-to point and point-to-multipoint configurations [8]

Point-to-Point (P2P)

Point to point is used where there are two points of interest: one sender and one receiver. This is also a scenario for backhaul or the transport from the data source (data center, co-lo facility, fiber POP, Central Office, etc) to the subscriber or for a point for distribution using point to multipoint architecture. Backhaul radios comprise an industry of their own within the wireless industry. As the architecture calls for a highly focused beam between two points range and throughput of point-to point radios will be higher than that of point-to-multipoint products.

Point-to-Multipoint (PMP)

As seen in the fig. 1.1, point-to-multipoint is synonymous with distribution. One base station can service hundreds of dissimilar subscribers in terms of bandwidth and services offered.

1.4.2 LOS Vs NLOS

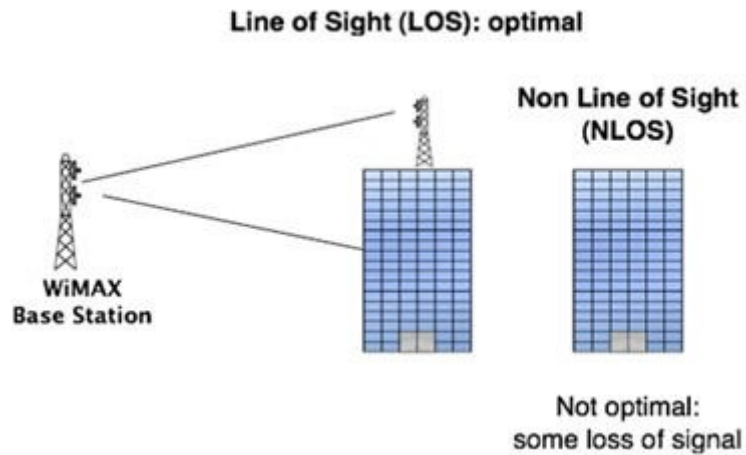


Figure 1.2: The difference between line of sight and non-line of sight [8]

Earlier wireless technologies e.g. LMDS (Local Multipoint Distribution Service), MMDS were unsuccessful in the mass market as they could not deliver services in NLOS scenarios. This limited the number of subscribers they could reach and, given the high cost of base stations and CPE, those business plans failed. WiMAX functions best in line of sight situations and, unlike those earlier technologies, offers acceptable range and throughput to subscribers who are not line of sight to the base station. Buildings between the base station and the subscriber diminish the range and throughput, but in an urban environment, the signal will still be strong enough to deliver adequate service. Given WiMAX's ability to deliver services non-line-of-sight, the WiMAX service provider can reach many customers in high-rise office buildings to achieve a low cost per subscriber because so many subscribers can be reached from one base station.

1.4.3 WiMAX Radios & Antennas

At the core of WiMAX is the WiMAX radio. A radio contains both a transmitter for sending and a receiver which receives. It generates electrical oscillations at a frequency known as the carrier frequency (in WiMAX that is usually between 2 and 11 GHz). A radio might be thought of as a networking device similar to a router or a bridge in that it

is managed by software and is composed of circuit boards containing very complex chip sets.

WiMAX architecture is built upon two components: radios and antennas. Most WiMAX products offer a base station radio separate from the antenna. Conversely, many CPE devices are also two piece solutions with an antenna on the outside of the building and subscriber station indoors as illustrated in the fig. 1.3.

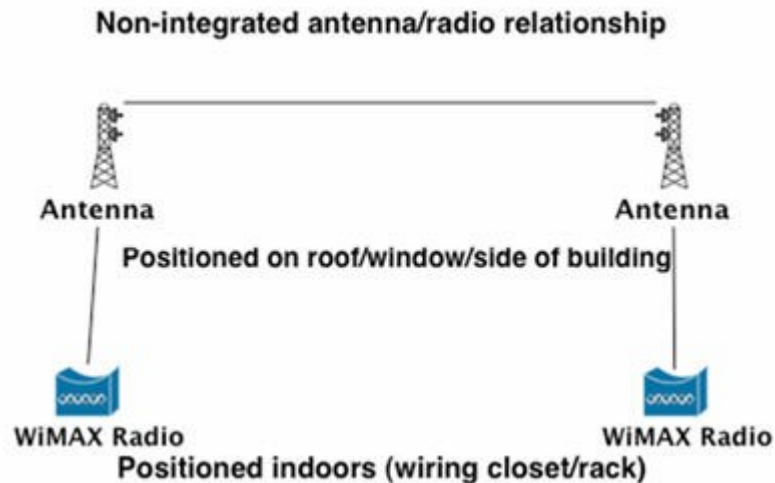


Figure 1.3: WiMAX Radio and Antennas [9]

The chief advantage of this is that the radio is protected from extremes of heat cold and humidity all of which detract from the radio's performance and durability. In addition, having the antenna outdoors optimizes the link budget (performance of the wireless connection) between transmitter and receiver especially in line of sight scenarios. The antenna is connected to WiMAX radio via a cable known as a "pigtail". One simple rule for wireless installations: keep the pigtail as short as possible so as to reduce the signal loss. The popular LMR-400 cable, for example will lose about 1 dB for every 10 feet of cable.

1.4.4 WiMAX Antennas

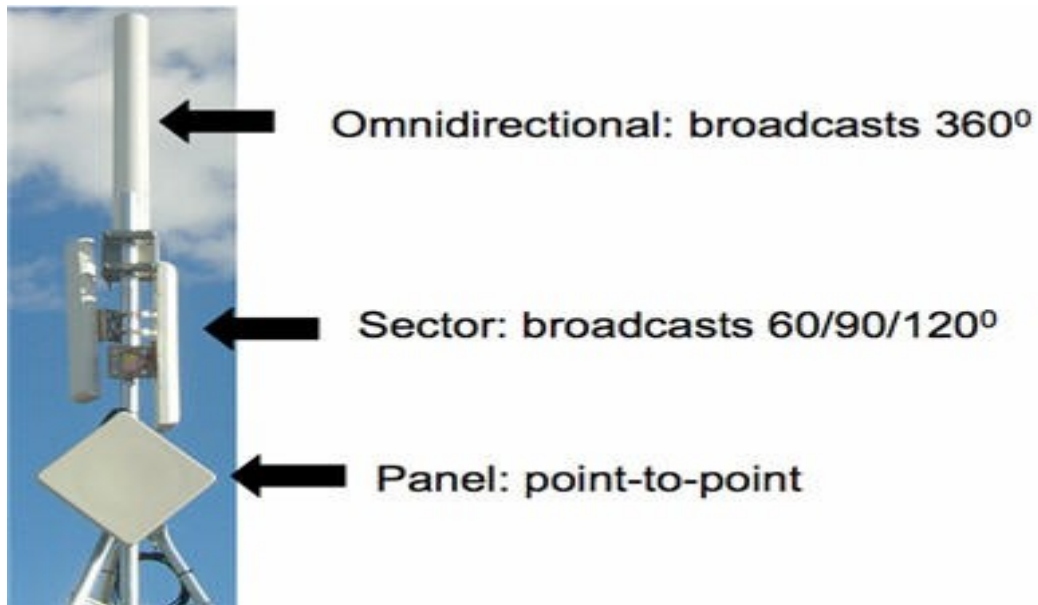


Figure 1.4: Different types of Antennas for different applications [10]

WiMAX antennas, just like the antennas for car radio, cell phone, FM radio, or TV, are designed to optimize performance for a given application. The fig. 1.4 illustrates the three main types of antennas used in WiMAX deployments. From top to bottom is an omnidirectional, sector and panel antenna each has a specific function.

1.4.5 Subscriber Stations

The technical term for Customer Premise Equipment (CPE) is subscriber station. The generally accepted marketing terms now focus on either "indoor CPE" or "outdoor CPE". There are advantages and disadvantages to both deployment schemes as described below.



Figure 1.5: An outdoor CPE device (Subscriber Station). Source: Airspan [9]

An outdoor CPE offers somewhat better performance over indoor CPE given that WiMAX reception is not impeded by walls of concrete or brick, RF blocking glass or steel in the building's walls .Indoor CPE can be installed by subscriber and service provider need not provide installation for it. Thus, it reduces waiting time needed for installation by service provider. Fig 1.6 shows typical Indoor CPE



Figure 1.6: Indoor WiMAX CPE (Airspar EasyST) - object on left) [9]

1.5 IP- Based WiMAX Network Architecture

The IEEE 802.16e-2005 standard provides the air interface for WiMAX but does not define the full end-to-end WiMAX network. The WiMAX Forum's Network Working Group (NWG) is responsible for developing the end-to-end network requirements, architecture, and protocols for WiMAX, using IEEE 802.16e-2005 as the air interface. This group has developed a network reference model to serve as an architecture framework for WiMAX deployments and to ensure interoperability among various WiMAX equipment and operators.

The architecture framework is defined such that the multiple players can be part of the WiMAX service value chain. More specifically, the architecture allows for three separate business entities: (1) network access provider (NAP), which owns and operates the ASN; (2) network services provider (NSP), which provides IP connectivity and WiMAX services to subscribers using the ASN infrastructure provided by one or more NAPs; and (3) application service provider (ASP), which can provide value-added services such as multimedia applications using IMS (IP multimedia subsystem) and corporate VPN (Virtual Private Networks) that run on top of IP. This separation between NAP, NSP, and

ASP is designed to enable a richer ecosystem for WiMAX service business, leading to more competition and hence better services.

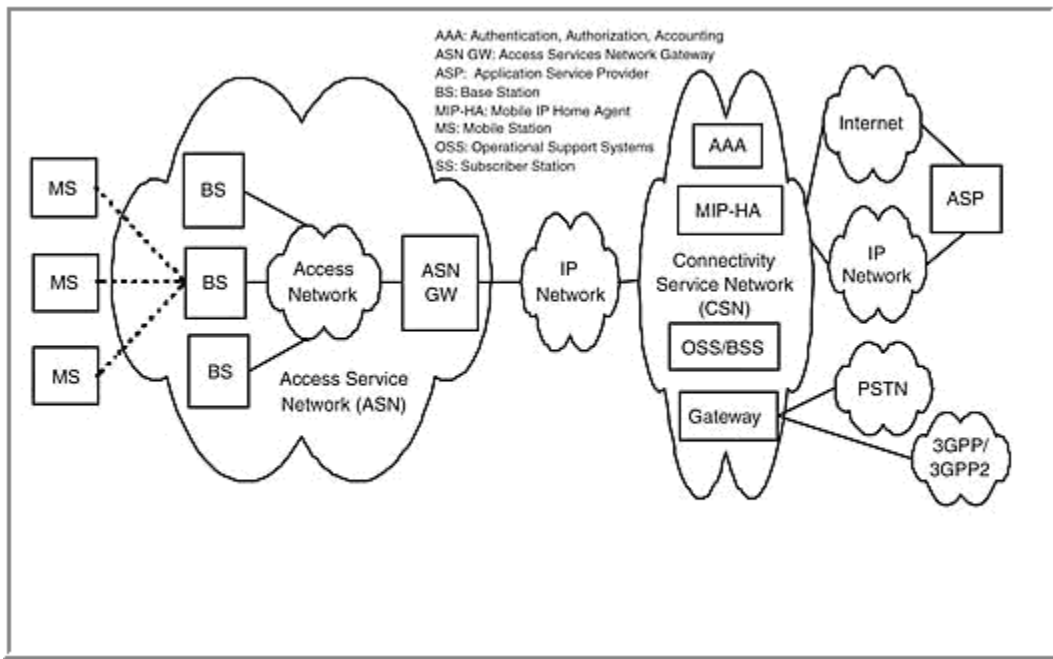


Figure 1.7: IP-Based WiMAX Network Architecture [6]

The network reference model developed by the WiMAX Forum NWG defines a number of functional entities and interfaces between those entities. (The interfaces are referred to as reference points.) Fig. 1.7 shows some of the more important functional entities.

Base station (BS): The BS is responsible for providing the air interface to the MS. Additional functions that may be part of the BS are micromobility management functions, such as handoff triggering and tunnel establishment, radio resource management, QoS policy enforcement, traffic classification, DHCP (Dynamic Host Control Protocol) proxy, key management, session management, and multicast group management.

Access service network gateway (ASN-GW): The ASN gateway typically acts as a layer 2 traffic aggregation point within an ASN. Additional functions that may be part of the ASN gateway include intra-ASN location management and paging, radio resource

management and admission control, caching of subscriber profiles and encryption keys, AAA client functionality, establishment and management of mobility tunnel with base stations, QoS and policy enforcement, and foreign agent functionality for mobile IP, and routing to the selected CSN.

Connectivity service network (CSN): The CSN provides connectivity to the Internet, ASP, other public networks, and corporate networks. The CSN is owned by the NSP and includes AAA servers that support authentication for the devices, users, and specific services. The CSN also provides per user policy management of QoS and security. The CSN is also responsible for IP address management, support for roaming between different NSPs, location management between ASNs, and mobility and roaming between ASNs. Further, CSN can also provide gateways and interworking with other networks, such as PSTN (Public Switched Telephone Network), 3GPP, and 3GPP2.

1.6 Applications of WiMAX

WiMAX has the potential to replace a number of existing telecommunications infrastructures. In a fixed wireless configuration it can replace the telephone company's copper wire networks, the cable TV's coaxial cable infrastructure while offering Internet Service Provider (ISP) services. In its mobile variant, WiMAX has the potential to replace cellular networks. Fig. 1.8 represents various potential applications in which WiMAX can be used as an alternative to existing options.

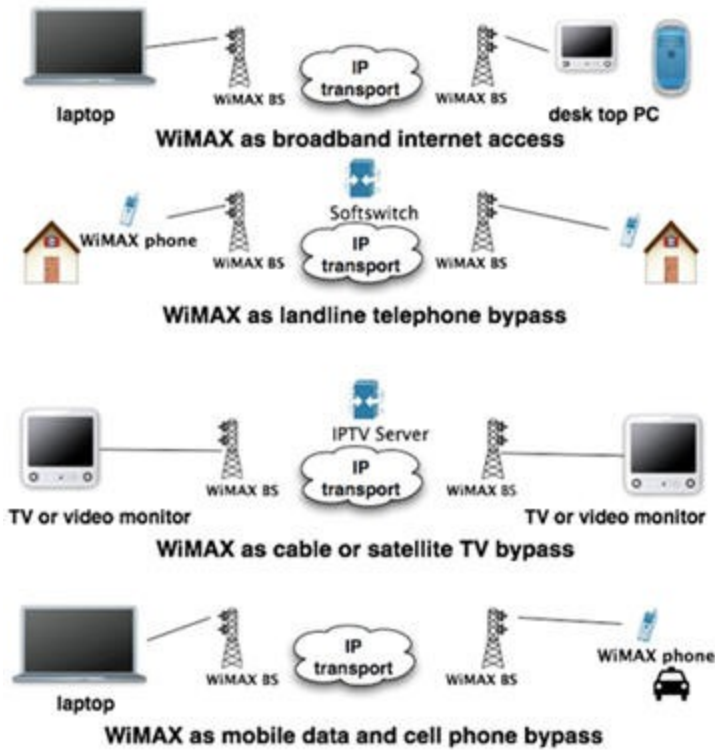


Figure 1.8: WiMAX Applications [9]

Applications using a fixed wireless solution can be classified as point-to-point or point-to-multipoint. Point-to-point applications include interbuilding connectivity within a campus and microwave backhaul. Point-to-multipoint applications include (1) broadband for residential, Small Office/Home Office (SOHO), and Small- to Medium-Enterprise (SME) markets, (2) T1 or fractional T1-like services to businesses, and (3) wireless backhaul for Wi-Fi hotspots. While in its mobile variant, WiMAX can be competent alternate to cellular networks and also can be used for emerging applications like Mobile TV, Streaming audio/video when subscriber is mobile.

Chapter 2

Literature Survey

Physical and MAC layer issues in context of WiMAX are described below first and then various scheduling algorithm for ensuring QoS of various service flows are described.

2.1 Physical Layer in 802.16

The WiMAX physical layer is based on orthogonal frequency division multiplexing. OFDM is the transmission scheme of choice to enable high-speed data, video, and multimedia communications and is used by a variety of commercial broadband systems, including DSL, Wi-Fi, Digital Video Broadcast-Handheld (DVB-H), and MediaFLO, besides WiMAX. OFDM is an elegant and efficient scheme for high data rate transmission in a non-line-of-sight or multipath radio environment. In this section, we cover the basics of OFDM and other physical layer issues.

2.1.1 Multicarrier Modulation (OFDM)

OFDM belongs to a family of transmission schemes called multicarrier modulation, which is based on the idea of dividing a given high-bit-rate data stream into several parallel lower bit-rate streams and modulating each stream on separate carriers—often called subcarriers, or tones. Multicarrier modulation schemes eliminate or minimize Inter Symbol Interference (ISI) by making the symbol time large enough so that the channel-induced delays—delay spread being a good measure of this in wireless channels[11]—are an insignificant (typically, <10 percent) fraction of the symbol duration. Therefore, in high-data-rate systems in which the symbol duration is small, being inversely proportional to the data rate, splitting the data stream into many parallel streams increases the symbol duration of each stream such that the delay spread is only a small fraction of the symbol duration.

OFDM is a spectrally efficient version of multicarrier modulation, where the subcarriers are selected such that they are all orthogonal to one another over the symbol duration, thereby avoiding the need to have nonoverlapping subcarrier channels to eliminate intercarrier interference. Choosing the first subcarrier to have a frequency such that it has an integer number of cycles in a symbol period, and setting the spacing between adjacent subcarriers (subcarrier bandwidth) to be $BSC = B/L$, where B is the nominal bandwidth (equal to data rate), and L is the number of subcarriers, ensures that all tones are orthogonal to one another over the symbol period. It can be shown that the OFDM signal is equivalent to the Inverse Discrete Fourier Transform (IDFT) of the data sequence block taken L at a time. This makes it extremely easy to implement OFDM transmitters and receivers in discrete time using IFFT (Inverse Fast Fourier) and FFT, respectively. [12]

2.1.2 Subchannelization (OFDMA)

The available subcarriers may be divided into several groups of subcarriers called subchannels. Fixed WiMAX based on OFDM-PHY allows a limited form of subchannelization in the uplink only. The standard defines 16 subchannels, where 1, 2, 4, 8, or all sets can be assigned to a Subscriber Station (SS) in the uplink. Uplink subchannelization in fixed WiMAX allows subscriber stations to transmit using only a fraction (as low as 1/16) of the bandwidth allocated to it by the base station, which provides link budget improvements that can be used to enhance range performance and/or improve battery life of subscriber stations. A 1/16 subchannelization factor provides a 12 dB link budget enhancement.

Mobile WiMAX based on OFDMA-PHY, however, allows subchannelization in both the uplink and the downlink, and here, subchannels form the minimum frequency resource-unit allocated by the base station. Therefore, different subchannels may be allocated to different users as a multiple-access mechanism. This type of multiaccess scheme is called orthogonal frequency division multiple access (OFDMA), which gives the mobile WiMAX PHY its name.

2.1.3 Adaptive Modulation and Coding

WiMAX supports a variety of modulation and coding schemes and allows for the scheme to change on a burst-by-burst basis per link, depending on channel conditions. Using the channel-quality feedback indicator, the mobile can provide the base station with feedback on the downlink channel quality. For the uplink, the base station can estimate the channel quality, based on the received signal quality. The base station scheduler can take into account the channel quality of each user's uplink and downlink and assign a modulation and coding scheme that maximizes the throughput for the available signal-to-noise ratio. Adaptive modulation and coding significantly increases the overall system capacity, as it allows real-time trade-off between throughput and robustness on each link.

Table 2.1 lists the various modulation and coding schemes supported by WiMAX. In the downlink, QPSK, 16 QAM, and 64 QAM are mandatory for both fixed and mobile WiMAX; 64 QAM is optional in the uplink. FEC coding using convolutional codes is mandatory. Convolutional codes are combined with an outer Reed-Solomon code in the downlink for OFDM-PHY. The standard optionally supports turbo codes and Low-Density Parity Check (LDPC) codes at a variety of code rates as well. A total of 52 combinations of modulation and coding schemes are defined in WiMAX as burst profiles.

Table 2.1 Modulation and Coding Supported in WiMAX [6]

	Downlink	Uplink
Modulation	BPSK, QPSK, 16 QAM, 64 QAM; BPSK optional for OFDMA-PHY	BPSK, QPSK, 16 QAM; 64 QAM optional
Coding	Mandatory: convolutional codes at rate 1/2, 2/3, 3/4, 5/6 Optional: convolutional turbo codes at rate 1/2, 2/3, 3/4, 5/6; repetition codes at rate 1/2, 1/3, 1/6, LDPC, RS-Codes for OFDM-PHY	Mandatory: convolutional codes at rate 1/2, 2/3, 3/4, 5/6 Optional: convolutional turbo codes at rate 1/2, 2/3, 3/4, 5/6; repetition codes at rate 1/2, 1/3, 1/6, LDPC

2.2 MAC Layer

The primary task of the WiMAX MAC layer is to provide an interface between the higher transport layers and the physical layer. The MAC layer takes packets from the upper layer—these packets are called MAC Service Data Units (MSDUs)—and organize them into MAC Protocol Data Units (MPDUs) for transmission over the air. For received transmissions, the MAC layer does the reverse. The IEEE 802.16-2004 and IEEE 802.16e-2005 MAC design includes a convergence sublayer that can interface with a variety of higher-layer protocols, such as ATM, TDM Voice, Ethernet, IP, and any unknown future protocol. Given the predominance of IP and Ethernet in the industry, the WiMAX Forum has decided to support only IP and Ethernet at this time. Besides providing a mapping to and from the higher layers, the convergence sublayer supports MSDU header suppression to reduce the higher layer overheads on each packet. Various sublayers of MAC layer and corresponding issues are clear from following fig. 2.1. of WiMAX reference model.

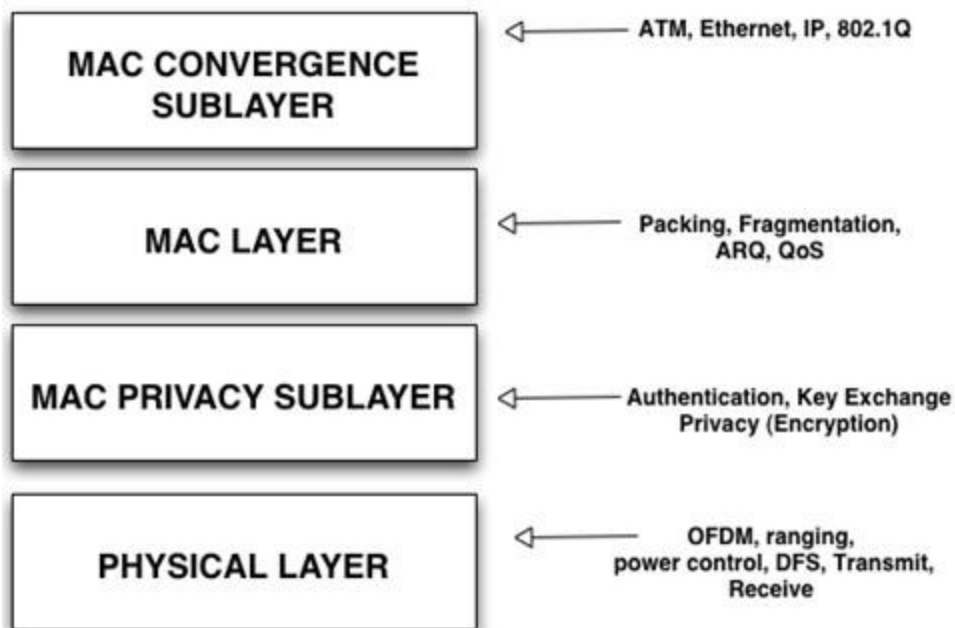


Figure 2.1: WiMAX Reference Model [9]

The WiMAX MAC is designed from the ground up to support very high peak bit rates while delivering quality of service similar to that of ATM and DOCSIS. The WiMAX MAC uses a variable-length MPDU and offers a lot of flexibility to allow for their efficient transmission. For example, multiple MPDUs of same or different lengths may be aggregated into a single burst to save PHY overhead. Similarly, multiple MSDUs from the same higher-layer service may be concatenated into a single MPDU to save MAC header overhead. Conversely, large MSDUs may be fragmented into smaller MPDUs and sent across multiple frames. Various issues related to MAC layers are as follows:

2.2.1 Channel-Access Mechanisms

In WiMAX, the MAC layer at the base station is fully responsible for allocating bandwidth to all users, in both the uplink and the downlink. The only time the MS has some control over bandwidth allocation is when it has multiple sessions or connections with the BS. In that case, the BS allocates bandwidth to the MS in the aggregate, and it is up to the MS to apportion it among the multiple connections. All other scheduling on the downlink and uplink is done by the BS. For the downlink, the BS can allocate bandwidth to each MS, based on the needs of the incoming traffic, without involving the MS. For the uplink, allocations have to be based on requests from the MS.

The WiMAX standard supports several mechanisms by which an MS can request and obtain uplink bandwidth. Depending on the particular QoS and traffic parameters associated with a service, one or more of these mechanisms may be used by the MS. The BS allocates dedicated or shared resources periodically to each MS, which it can use to request bandwidth. This process is called polling. Polling may be done either individually (unicast) or in groups (multicast). Multicast polling is done when there is insufficient bandwidth to poll each MS individually. When polling is done in multicast, the allocated slot for making bandwidth requests is a shared slot, which every polled MS attempts to use. WiMAX defines a contention access and resolution mechanism for the case when more than one MS attempts to use the shared slot. If it already has an allocation for sending traffic, the MS is not polled. Instead, it is allowed to request more bandwidth by (1) transmitting a stand-alone bandwidth request MPDU, (2) sending a bandwidth request

using the ranging channel, or (3) piggybacking a bandwidth request on generic MAC packets.

2.2.2 Quality of Service

Support for QoS is a fundamental part of the WiMAX MAC-layer design. Strong QoS control is achieved by using a connection-oriented MAC architecture, where all downlink and uplink connections are controlled by the serving BS. Before any data transmission happens, the BS and the MS establish a unidirectional logical link, called a connection, between the two MAC-layer peers. Each connection is identified by a Connection Identifier (CID), which serves as a temporary address for data transmissions over the particular link.

WiMAX also defines a concept of a service flow. A service flow is a unidirectional flow of packets with a particular set of QoS parameters and is identified by a Service Flow Identifier (SFID). The QoS parameters could include traffic priority, maximum sustained traffic rate, maximum burst rate, minimum tolerable rate, scheduling type, ARQ type, maximum delay, tolerated jitter, service data unit type and size, bandwidth request mechanism to be used, transmission PDU formation rules, and so on. Service flows may be provisioned through a network management system or created dynamically through defined signaling mechanisms in the standard. The base station is responsible for issuing the SFID and mapping it to unique CIDs.

To support a wide variety of applications, WiMAX defines five scheduling services (Table 2.2) that should be supported by the base station MAC scheduler for data transport over a connection. It should be noted that design of an effective scheduler for prioritizing is critical to the overall capacity and performance of a WiMAX system.

Table 2.2: Service classes supported in WiMAX for prioritizing traffic [9]

Service Class	Applications	QoS Specifications
Unsolicited Grant Service (UGS)	VoIP	Jitter Tolerance Maximum Latency Tolerance Maximum Sustained Rate
Real-time Polling Service (rtPS)	Streaming audio or video	Traffic Priority Maximum Latency Tolerance Maximum Reserved Rate Maximum Sustained Rate
Extended real time Polling Services (ErtPS)	VoIP (Voice with Activity Detection)	Traffic Priority Maximum Latency Tolerance Jitter Tolerance Maximum Sustained Rate Maximum Reserved Rate
Non-real time polling services (nrtPS)	FTP	Traffic Priority Maximum Reserved Rate Maximum Sustained Rate
Best- Effort (BE)	Data Transfer, Web Browsing etc.	Traffic Priority Maximum Sustained Rate

2.2.3 Mobility Support

In addition to fixed broadband access, WiMAX envisions four mobility-related usage scenarios:

1. Nomadic. The user is allowed to take a fixed subscriber station and reconnect from a different point of attachment.

2. Portable. Nomadic access is provided to a portable device, such as a PC card, with expectation of a best-effort handover.

3. Simple mobility. The subscriber may move at speeds up to 60 kmph with brief interruptions (less than 1 sec) during handoff.

4. Full mobility: Up to 120 kmph mobility and seamless handoff (less than 50 ms latency and <1% packet loss) is supported.

WiMAX networks will initially be deployed for fixed and nomadic applications and then evolves to support portability to full mobility over time. The IEEE 802.16e-2005 standard defines a framework for supporting mobility management. In particular, the standard defines signaling mechanisms for tracking subscriber stations as they move from the coverage range of one base station to another when active or as they move from one paging group to another when idle. The standard also has protocols to enable a seamless handover of ongoing connections from one base station to another. The WiMAX Forum has used the framework defined in IEEE 802.16e-2005 to further develop mobility management within an end-to-end network architecture framework. The architecture also supports IP-layer mobility using mobile IP.

2.2.4 Security Functions

Unlike Wi-Fi, WiMAX systems were designed at the outset with robust security in mind. The standard includes state-of-the-art methods for ensuring user data privacy and preventing unauthorized access, with additional protocol optimization for mobility. Security is handled by a privacy sublayer within the WiMAX MAC. The key aspects of WiMAX security for fixed and wireless are as shown in following fig. 2.2.

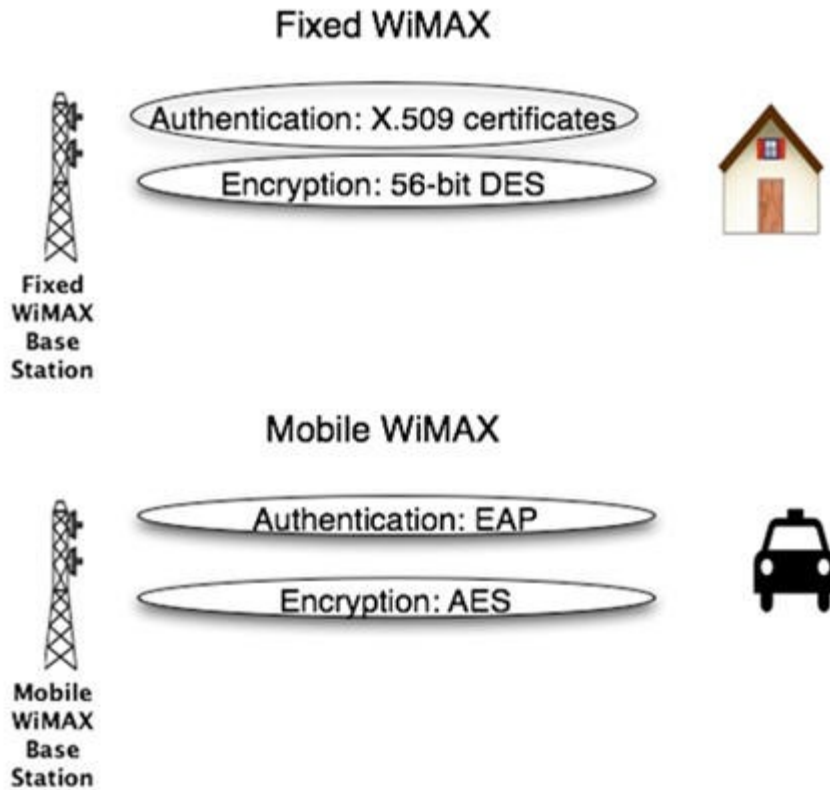


Figure 2.2: WiMAX Security through Authentication and Encryption [9]

2.3 Scheduling Architecture

Typical IEEE 802.16 architecture includes one Base Station and multiple Subscriber Stations. Communication occurs in two directions: from BS to SS is called Downlink and From SS to BS is called Uplink. Thus each frame is divided in two subframe i.e. downlink subframe and uplink subframe as shown in fig. 2.3. During downlink BS broadcasts data to all subscribers and subscribers selects packets destined for it. While Uplink channel is shared by multiple SSs and to ensure this sharing this channel is slotted and these slots are allocated by BS to various SSs in one uplink frame by using TDD or FDD. This slot allocation information is broadcasted by the BS through the Uplink Map Message (UL-MAP) at the beginning of each frame. UL-MAP contains Information Element (IE) which includes the transmission opportunities, i.e., and the time slots in

which the SS can transmit during the uplink subframe. After receiving the UL-MAP, each SS will transmit data in the predefined time slots as indicated in IE.

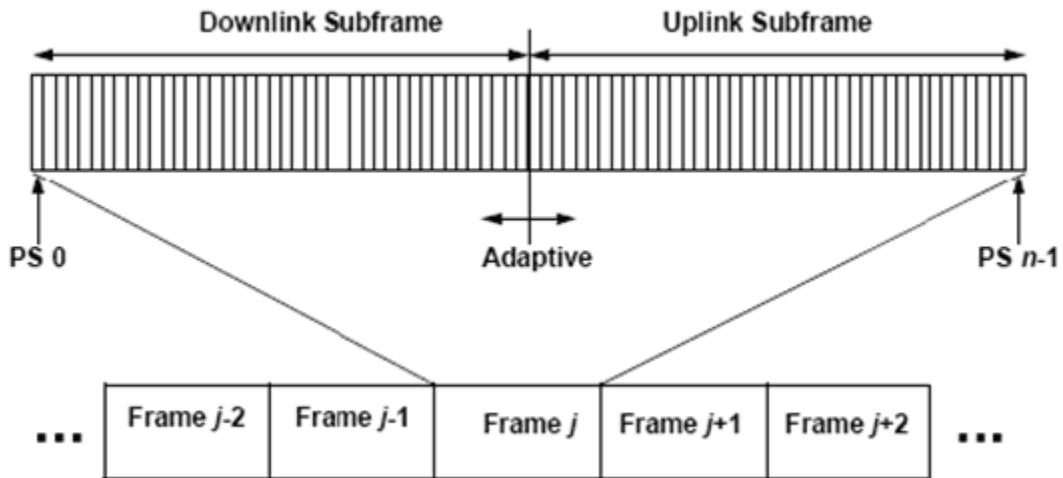


Figure 2.3 TDD frame structure in IEEE 802.16 [4]

2.3.1 QoS Architecture for IEEE 802.16 MAC Protocol

IEEE 802.16 can support multiple applications (data, voice, and video) with different QoS requirements.

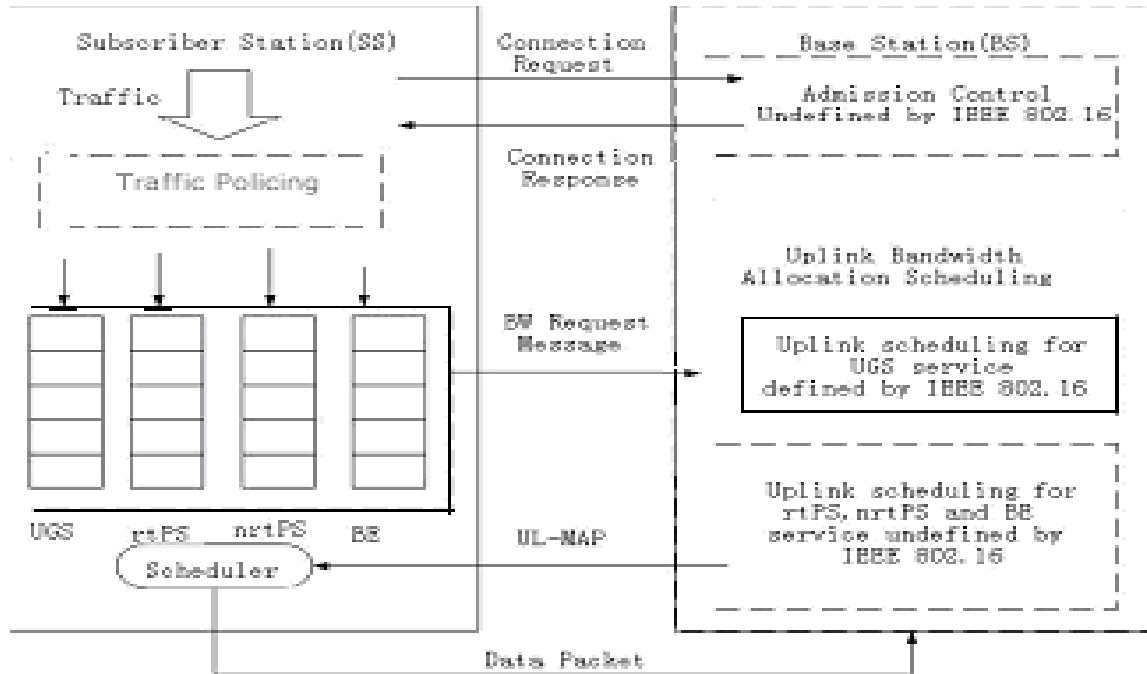


Figure 2.4 QoS architecture defined in 802.16

Fig. 2.4 shows the existing QoS architecture of IEEE 802.16. Uplink Bandwidth Allocation scheduling resides in the BS to control all the uplink packet transmissions. Since IEEE 802.16 MAC protocol is connection oriented, the application first establishes the connection with the BS as well as the associated service flow (UGS, rtPS, nrtPS or BE). BS will assign the connection with a unique connection ID (CID)

Each connection requests for desired bandwidth to the BS. Bandwidth is granted per connection or per subscriber in which per connection allocation is done by SS. IEEE 802.16 defines:

- 1) The signaling mechanism for information exchange between BS and SS such as the connection setup, BW-Request and UL-MAP.
- 2) The Uplink Scheduling for UGS service flow.

And IEEE 802.16 does not define:

- 1) The Uplink Scheduling for rtPS, nrtPS and BE service flow.
- 2) The Admission Control and Traffic Policing process.

Out of these undefined parts in next section we discuss various approaches to provide Uplink scheduling for rtPS, nrtPs and BE service flow. ErtPS (Extended Real-time polling Service) is defined only in IEEE 802.16e-2005 and not in IEEE 802.16-2004. So, we are not considering it as separate service flow from rtPS

2.4 Scheduling Algorithms for QoS Support in WiMAX

2.4.1 Scheduling Strategies

As Scheduling for rtPS, nrtPS and BE services is not defined in 802.16 standard, there are various approaches described to address this issue. Some of them consider modifying scheduling algorithms defined for other types of networks i.e. wired networks. But this approach does not give accurate results for wireless networks because wireless channel is going to fade over time and so channel quality does not remain same. So, scheduling algorithm should consider this factor in scheduling. Besides this, different algorithms works on a specific approach e.g. some algorithm are just priority based which schedules all connection by just one centralized scheduling. Some algorithm works by selecting different algorithm for each different class. To be more specific in this approach some algorithms are also developed which just address one class. Here we are going to discuss three algorithms which represent these three different approaches.

2.4.2 Overview of various Scheduling Algorithms

In this section we describe various available scheduling algorithms and analyze how good they are as far as QoS requirements of various applications are considered.

2.4.2.1 Round Robin (RR)

RR is one of the simplest scheduling algorithms designed especially for a time sharing system, where the scheduler assigns time slots to each queue in equal portions without priority. Once a queue is served, it is not visited again until all the other SSs in the system

have been served. RR can provide a fair resource access to each SS, and every queue is allocated with the same portion of system resources regardless of the channel condition. However, the RR scheduler has the same bandwidth efficiency as a random scheduler. Also, it cannot guarantee different QoS requirements for each queue.

2.4.2.2 Proportionate Fair Scheduling [13]

PF was proposed by Qualcomm Company, which was realized in the IS-856 standard for the downlink traffic scheduling (also known as High Data Rate (HDR)) [13]. It is devised to avoid conflict between full use (by selecting user with highest data rate) and fairness. It does this by maintaining trade-off between system throughput and starvation of low priority users. This algorithm is based on one priority function:

$U_i(t) = r_i(t) / R_i(t)$ where $r_i(t)$ is the current data rate and $R_i(t)$ is an exponentially smoothing average of the service rate received by SS i up to slot t . Queue having highest value of $U_i(t)$ is served at time slot t . For updating average throughput of the queue following function is used:

$R_i(t+1) = (1-1/T_c) * R_i(t) + (1/T_c)r_i(t)$ if connection i is served at time-slot t

$R_i(t+1) = (1-1/T_c) * R_i(t)$ if connection i is not served at time-slot t

Here T_c is time-constant for finding out moving average which is generally taken 1000 slots in CDMA-HDR system [13]. By adjusting this T_c parameter we can make perceived throughput less sensitive to short-term starvation on the queue. So scheduler waits for long time for a particular connection for improvement of its channel quality. Also, when large number of users co-exists in system we can leverage from feature called multi-user diversity gain in which we can schedule connection having highest data rate at particular time slot. Thus PF algorithm is simple and efficient but it fails to deliver QoS parameter requirements for a particular connection as mentioned for various service classes e.g. delay, throughput etc.

2.4.2.3 Novel Scheduling Algorithm [14]

In this Scheduling Algorithm, Author propose scheme in which they are using different scheduling algorithm for different classes of service flow. They have selected WFQ for

(Wireless Fair Queuing) for rtPS service flow. In this, priority of connection is calculated according to its weight which is equal to average connection rate/total average traffic rate. WFQ is selected because it is suitable for service flow which requires delay bound. For nrtPS service they have selected DRR (Dynamic Round Robin) as it is easy to implement and so more suitable for middle priority service like nrtPS. For BE service they have used FIFO algorithm because there is no QoS needed for BE service. Thus they have taken into consideration requirement of real-time service, complexity of scheduling algorithm and fairness. But they have not shown how they have integrated all these different algorithms for providing centralized control from one BS. Also, detail about DRR strategy for nrtPS is not mentioned.

2.4.2.4 Cross- Layer Scheduling Algorithm [15]

To manage resource allocation and grants an appropriate QoS per connection, other scheduling schemes are proposed. These scheduling schemes rely on different priority functions to handle different classes of services for matching their QoS requirements. To have a comprehensive introduction, a representative cross-layer scheduling algorithm with QoS support by [15] is briefed as follows.

In this algorithm various metrics of different class are considered to calculate priority of each connection. These metrics are briefly described as below:

1. UGS: - PER (Packet Error Rate) and Service Rate. Because it requires guarantee on throughput, latency jitter upto some tolerance of packet loss i.e. voip
2. RTPS: - PER and Max. Delay after which packet is useless i.e. streaming video.
3. NRTPS: - PER and min. Reserved Rate i.e. FTP
4. BE: - No any guarantee needed but PER should be maintained i.e. http .e -mail

So, based on these parameters and channel condition priority of a particular connection is calculated. But this priority calculated for only three of service flow mentioned above. For UGS connection, there is no such any priority is calculated but certain fix number of slots are allocated to it. After that remaining slots are allocated to connection having highest priority. For calculating priority various parameters like channel condition, delay requirement (for rtPS) and average data rate (for nrtPS) are considered and priority is

calculated. In order to differentiate the priority of the four types of services such that $rtPS > nrtPS > BE$, the class coefficients are assigned to the queues of each service type.

This algorithm is further described in detail in chapter 4.

2.4.2.5 TCP-Aware Uplink Scheduling Algorithm [16]

This algorithm works with only one class of 4 classes defined for QoS. It deals with BE class. As this class has not any specific QoS requirement it is not advantageous to use bandwidth request mechanism for this class and to waste that bandwidth. Also, it is not advisable to equally allocate remaining bandwidth to all remaining BE connections because all connections can't utilize all bandwidth allocate to them and some may have more requirements than allocated. So, this algorithm works by calculating bandwidth for a particular connection according to sending rate of that connection. Also as sending rate is going to change dynamically, it is not proper to allocate fix amount of bandwidth to a particular connection. So, to properly allocate bandwidth this algorithm works as follows:

Step1: Compute the sending rate.

Step2: If sending rate < allocated bandwidth

Then demand = sending rate

Step3: if sending rate = allocated bandwidth

Then demand = Increment allocated bandwidth in fix proportion

Step4: If Sending rate > allocated bandwidth

Then increase bandwidth until sending rate becomes stable

The main strategy of this algorithm is to allocate bandwidth somewhat higher than actual sending rate of connection so that we can safely estimate the sending rate at any given time. To detect changes in sending rate its max & min. values are maintained over a period of time and are changed according to rate change. Whenever these values are to be changed above algorithm is used for demand estimation. After estimating demand for each connection, it uses max-min fair scheduling for allocating total bandwidth among all connections.

Chapter 3

Objective

3.1 Research Challenges

As discussed in chapter 1 Broadband wireless access industry has seen very significant growth in present time due to its unique advantages compared to wired technology like rapid deployment and easy configuration, high scalability, lower maintenance cost, lower investment and also end-user preference. But as this industry has grown, different alternatives also emerged to capture available market. In this scenario, it is very important to provide outstanding service to end-user by ensuring that diversifying requirements of different applications are satisfied in the best way possible. Due to this Quality of Service (QoS) has become very important issue in present era to differentiate oneself from other competing technology. WiMAX (IEEE 802.16) is one such standard that specifies general QoS architecture and the most important part of this architecture is Scheduling Algorithm and this part is not defined and left open for vendors to implement as per their needs.

Scheduling algorithms serve as an important component in any communication network to satisfy the QoS requirements. The design is especially challenged by the limited capacity and dynamic channel status that are inherent in wireless communication systems. We have described various scheduling algorithm in chapter 2 which are proposed for ensuring QoS requirements of various service flows. In general, algorithm should satisfy following features and criteria for optimizing system performance. [17-19]

1) Bandwidth utilization

Efficient bandwidth utilization is the most important in the algorithm design. The algorithm must utilize the channel efficiently. This implies that the scheduler

should not assign a transmission slot to a connection with a currently bad link. Also slots allocated to particular connection are to be used fully.

2) QoS requirements

The proposed algorithm should support different applications to exploit better QoS. To support delay-sensitive applications, the algorithm provides the delay bound provisioning. The long-term throughput should be guaranteed for all connections when the sufficient bandwidth is provided.

3) Fairness

The algorithm should assign available resource fairly across connections. The fairness should be provided for both short term and long term.

4) Implementation complexity

In a high-speed network, the scheduling decision making process must be completed very rapidly, and the reconfiguration process in response to any network state variation. Therefore, the amount of time available to the scheduler is limited. A low-complexity algorithm is desirable.

5) Scalability

The algorithm should operate efficiently as the number of connections or users sharing the channel increases.

An ideal algorithm should satisfy all of above criteria for efficient scheduling. Below we provide a comparative study of three representative algorithms for approaches mentioned in chapter 2 which describes advantages and disadvantages in terms of above mentioned criteria.

Table 3.1 Comparison of different QoS algorithms in WiMAX (IEEE 802.16)

Algorithm	Advantage	Disadvantage
PF Algorithm	Fairness In Scheduling Priority Based Simple Implementation Multi-User Diversity Gain	No QoS Guarantee
Cross-Layer Scheduling Algorithm	QoS Guarantee Channel Quality is also considered in Scheduling	Complex Implementation All slots per frame are allocated to highest priority connection
TCP-aware uplink Scheduling Algorithm	Efficient utilization of resources among BE connections	Treats only one class Complex implementation

Apart from above three there are many other algorithms like Round Robin , WFQ(Wireless Fair Queue), EDF(Earliest Deadline First) etc. but all of them are going to relate to one of the strategies mentioned in section 3.

3.2 Problem Statement

Out of the three strategies mentioned above, first strategy which involves modification to algorithm developed for other type of networks is not suitable because fading of channel quality is very important issue to consider. Also, to satisfy QoS to a particular connection is another important issue. These two issues are not handled in this approach and so it doesn't give expected performance. Third approach suggests to develop algorithm which is suitable for a particular class of defined service flow in WiMAX has also some drawbacks:

- 1) As it treats only one class in isolation it is not applicable in situation when we want to integrate effect of fading channel quality to all connections available for scheduling.
- 2) As each connection has its own requirements it is not efficient to combine all connections of one service flow into group and to allocate slots to each group. QoS

requirements are mentioned for each connection and not for each service flow. Service flow only indicates type of QoS requirement.

Second approach mentioned in section 2.4.2.3 which makes allocation of available slots on per connection basis has not abovementioned drawback. Representative Cross-Layer Scheduling algorithm works well as it is going to handle QoS requirements for each connection as well as it considers channel quality per frame in determining priority of a particular connection. It is complex compare to first approach but this complexity has to be handled if we want to satisfy QoS requirements of each connection more accurately. Increased complexity can be taken care by using more efficient hardware. But this Algorithm has also one drawback. It allocates all available slots to a particular connection for a given frame and only one connection can transmit during each frame. This may result in underutilization of resources i.e. slots when all available slots are not needed for a particular connection and so they can be allotted to other connections.

Efficient utilization of resources is very important factor because it is going to affect QoS satisfaction of each connection under given resources. The Cross-Layer Scheduling algorithm can be modified so that better utilization of resources can be done.

Chapter 4

Scheduler Design & Implementation

The Scheduler Design with a typical network configuration of fixed PMP system and various parameters of QoS are described below. Modification to Scheduling Algorithm is also described after that.

4.1 Network Configuration

Typical network topology is as shown in fig. 4.1 in which there is one BS (Base Station) and multiple SSs (Subscriber Station). Each SS can support multiple connections (sessions, flows). This is a typical star topology which is applicable to wide variety of networks like cellular networks, mobile ad hoc networks and wireless sensor networks.



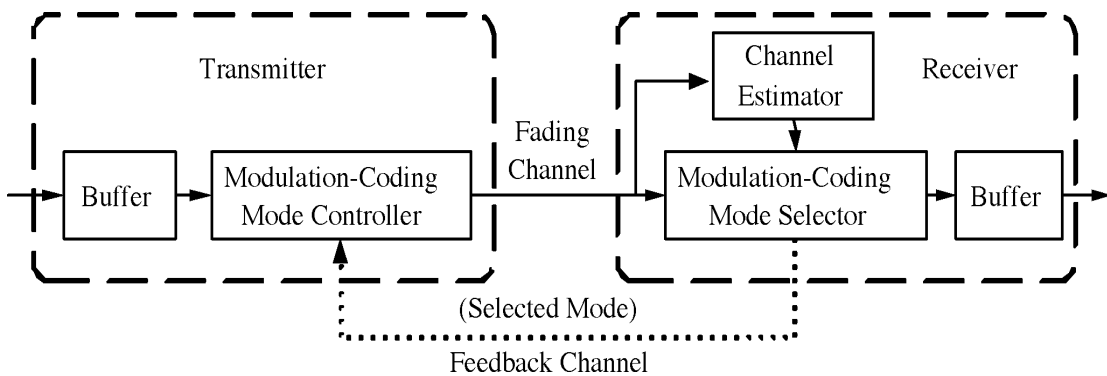
Figure 4.1 Network Topology

As there is only one BS to serve multiple SSs, inter BS interference can be neglected. The downlink channel is shared by all SSs in a time division multiplexing manner, where a downlink scheduler is deployed at the BS to schedule the transmissions corresponding to

each queue. The transmission occurs within a fixed-sized time frame. Only the selected queues can be served within the frame. The air interface specification is OFDM which employs a Fast Fourier Transform (FFT) of size 256. All carriers are assigned to one queue for the data transmission in a time slot. All connections in the network communicate with BS using Time–Division Multiplexing (TDM)/Time- Division Multiple Access (TDMA).

The signaling and procedure for the service setup and maintenance of each connection are defined as in the IEEE 802.16 standard [4]. However, the standard does not define the scheduling mechanism or the admission control and traffic policing processes. The signaling overhead is not included in our design and analysis. Typical Wireless link for each connection from BS to each SS is as shown in following Fig. 4.2.

Figure 4.2 Wireless link from BS to SS [15]



BS maintains buffer for each connection which operates as First-In-First-Out (FIFO) mode. This buffer is followed by AMC controller which receives its input from AMC selector implemented at SS (Receiver). At the PHY, multiple transmission modes are available to each user, with each mode representing a pair of a specific modulation format and a Forward Error Control (FEC) code, as in IEEE 802.11/15/16, 3GPP, and 3GPP2 standards. Based on channel estimates obtained at the receiver, the AMC selector determines the modulation-coding pair (mode or burst profile), whose index is sent back to the transmitter through a feedback channel, for the AMC controller to update the

transmission mode. Finally according to mode received from transmitter, bits are decoded and these decoded bit streams are mapped to packets and these packets are handled by MAC layer. As various groups of transmission modes are defined in the IEEE 802.16 standard [4], we consider following particular group of transmission modes:

Transmission Modes (TM): Various parameters defined for this mode are modulations, RS code and CC code rate etc. The modulations are Mn -ary rectangular/square Quadrature Amplitude Modulators (QAMs), and the FEC codes are Reed–Solomon (RS) concatenated with convolutional codes (CC) as shown in Table 4.1. Although we focus on this TM, other transmission modes can be similarly constructed [4], [20-22].

Table 4.1 Transmission Modes in IEEE 802.16 Standard [4]

Mode n	1	2	3	4	5	6
Modulation	QPSK	QPSK	16QAM	16QAM	64QAM	64QAM
RS Code	(32,24,4)	(40,36,2)	(64,48,8)	(80,72,4)	(108,96,6)	(102,108,6)
CC Code Rate	2/3	5/6	2/3	5/6	3/4	5/6
Coding Rate R_c	1/2	3/4	1/2	3/4	2/3	3/4
R_n (bits/symbol)	1.0	1.5	2.0	3.0	4.0	4.5
a_n (dB)	232.9242	140.7922	264.0330	208.5741	216.8218	220.7515
g_n	22.7925	8.2425	6.5750	2.7885	1.0675	0.8125
γ_{pn} (dB)	3.7164	5.9474	9.6598	12.3610	16.6996	17.9629

At the PHY, the processing unit is a frame consisting of multiple transmitted symbols. At the MAC, the processing unit is a packet comprising multiple information bits. Fig. 4.3 describes the packet and frame structures in detail.

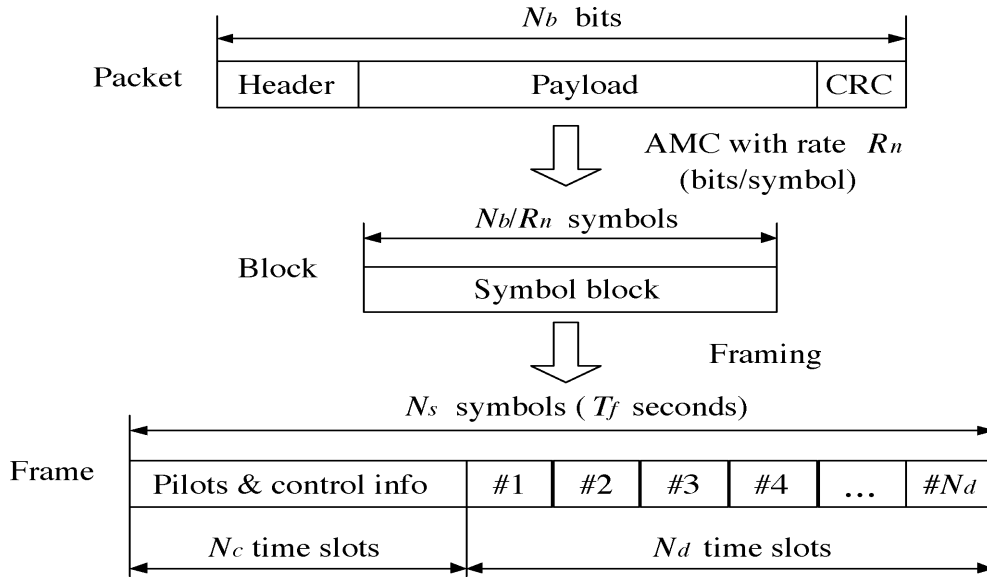


Figure 4.3 Processing Unit at MAC and PHY layer [15]

1. At the MAC, each packet contains a fixed number of bits N_b , which include packet header, payload, and Cyclic Redundancy Check (CRC) bits. After modulation and coding with mode n of rate R_n as in Table 4.1, each packet is mapped to a symbol block containing N_b/R_n symbols.
2. At physical layer the data are transmitted frame by frame basis over wireless channel where each frame contains a fixed number of symbol N_s . Given a fixed symbol rate, the frame duration T_f (in seconds) is constant and represents the time unit. With TDM, each frame is divided into $N_c + N_d$ time slots, where for convenience we let each time slot contain a fixed number of $2N_b/R_1$ symbols. As a result, each time slot can transmit exactly $2R_n/R_1$ packets with transmission mode n . For the TM in particular, one time slot can accommodate $2R_1/R_1 = 2$ packets with mode $n = 1$, $2R_2/R_1 = 3$ packets with mode $n = 2$, and so on. The N_c time slots contain control information and pilots. The N_d time slots convey data, which are scheduled to different connections dynamically. Each connection is allocated a certain number of time slots during each frame. How these slots are

allocated to a particular connection is decided by scheduler and will be described next.

4.2 Parameters indicating QoS requirement

Scheduling Algorithm in WiMAX should satisfy diverse requirements of various service flows mentioned in 802.16 QoS architecture. It should also consider channel quality as one parameter during scheduling. In Cross-Layer Scheduling algorithm, various metrics of different class are considered to calculate priority of each connection. These metrics are briefly described as below [15]:

- UGS: - PER (Packet Error Rate) and Service Rate. Service rate defines required guarantee on throughput, latency and jitter to necessary levels.
- rtPS: - PER and Max. Delay after which packet is useless i.e. streaming video.
- nrtPS: - PER and min. Reserved Rate i.e. FTP
- BE: -Not any guarantee needed but PER should be maintained i.e. http .e –mail

4.3 AMC Design at the PHY

Efficient bandwidth utilization for a prescribed PER performance at the PHY can be accomplished with AMC schemes, which match transmission parameters to the time-varying wireless channel conditions adaptively and have been adopted by many standard wireless networks, such as IEEE 802.11/15/16 and 3GPP/3GPP2 [4],[20-22].

Each connection with rtPS, nrtPS, and BE services relies on AMC at the PHY. The objective of AMC is to maximize the data rate by adjusting transmission modes to channel variations while maintaining a prescribed PER P_o , and the design procedure is similar to that proposed in [23] and [24]. Let N denote the total number of transmission modes available ($N = 6$ for TM). As in [23], we assume constant power transmission and partition the entire Signal-to-Noise Ratio (SNR) range in $N + 1$ nonoverlapping consecutive intervals, with boundary points denoted as $\{\gamma_n\}_{n=0}^{N+1}$. In this case mode n is chosen when $\gamma \in [\gamma_n, \gamma_{n+1})$, for $n = 1, \dots, N$.

To avoid deep-channel fades, no data are sent when $\gamma_0 \leq \gamma < \gamma_1$, which corresponds to the mode $n = 0$ with rate $R_0 = 0$ bit/symbol. The design objective of AMC is to determine the boundary points $\{\gamma_n\}_{n=0}^{N+1}$.

According to abovementioned PER, various thresholds for different modes as defined in IEEE 802.16 standard [4] are calculated based on following equation [25].

$$\gamma_0 = 0$$

$$\gamma_n = \frac{1}{g_n} \ln \left(\frac{a_n}{P_0} \right), n=1, 2, \dots, N$$

$$\gamma_{N+1} = \infty$$

Values of a_n and g_n are defined for various modes in IEEE 802.16 standard as shown in table 4.1. $\{\gamma_n\}_{n=0}^N$ represents minimum SNR required to guarantee prescribed PER (P_0) for particular connection. Thus it is a switching threshold based on which mode is selected for a particular connection having a specific SNR which depends on channel condition of a connection at that time. This mode determines current data rate of connection.

4.4 Scheduler Design

After mode is selected for a particular connection to determine scheduling priority for each connection is calculated as following:

First of all fixed number of timeslots are separately allocated for UGS services. Also transmission mode for UGS connection is fixed which is determined initially via training to meet prescribed PER and service requirement and same during whole service time. After that remaining slots are allocated to particular service among rtPS, nrtPS and BE services according to priority. The queues for rtPS are managed with an Earliest Deadline First (EDF) algorithm [26], which is sensitive to delay latency and reliable for real-time services. An opportunistic scheme which is similar to the PF algorithm [13] is deployed for the queues supporting nrtPS, while the queues for BE traffic are managed based on a Best-Rate discipline. In order to differentiate the priority of the three types of services such that $rtPS > nrtPS > BE$, the class coefficients are assigned to the queues of each service type in decreasing order respectively.

The scheduling algorithm is implemented according to the following Formulas:

Suppose one frame has N_d time slots available. Out of these N_d time-slots, fixed no. of time-slots say N_{ugs} are allocated for UGS connection. Remaining $N_r = N_d - N_{ugs}$ time slots are allocated to connection having highest priority.

Priority function is defined as follows:

$$\varnothing_i(t) = \left\{ \begin{array}{l} \beta_{class} \frac{R_i(t)}{R_N} \frac{1}{F_i(t)}; F_i(t) \geq 1, R_i(t) \neq 0 \\ \beta_{class}; F_i(t) < 1, R_i(t) \neq 0 \\ 0; R_i(t) = 0 \end{array} \right\}$$

We can set β_{class} variable according to priority of a particular class. Generally coefficients

are selected such that $\beta_{rtPS} > \beta_{nrtPS} > \beta_{BE}$. Parameter $\frac{R_i(t)}{R_N}$ indicates normalized channel

quality of a particular connection. Because R_N indicates maximum possible datarate and $R_i(t)$ indicates current data rate. So, this factor considers multiuser diversity gain advantage. $F_i(t)$ is a satisfaction parameter which is defined as follows:

$F_i(t) = T_i - \Delta T_i - W_i(t) + 1$ for real-time connection, where T_i is delay requirement specified for connection and $W_i(t)$ is maximum current delay requirement and ΔT_i is guard time region ahead of the deadline T_i .

$F_i(t) = \eta_i(t) / n_i$ for non real-time connection, where n_i is data rate specified for connection and $\eta_i(t)$ is average data rate which is calculated as follows:

$\eta_i(t+1) = \eta_i(t) (1-1/Tc) + (1/Tc) R_i(t)$, $R_i(t)$ is current data rate and Tc is time constant for calculating average

$F_i(t)$ for BE connection is not defined and so it can be taken as 1.

Thus for real-time connection satisfaction parameter calculates difference between current maximum delay requirement and which is described for a particular connection and if it is less than 1 means that packet should be immediately transmitted. This is ensured by assigning highest priority to such connection i.e. β_{rtps} . Similarly, for non-realtime connection it calculates ratio of average data rate to data rate described for a particular connection. So, if it is less than 1 means data rate requirements are not satisfied properly and so highest possible priority should be assigned to such connection which is β_{nrtps} . For Best Effort no such function is defined as there is no specific QoS requirement and so priority is calculated only based on normalized channel quality.

For any connection i , possible transmission rate (capacity) i.e. the number of packets that can be transmitted at time t can be expressed as

$$C_i(t) = N_r R_i(t), \text{ Where } N_r \text{ is remaining time slots for a given frame.}$$

4.5 Proposed Modification in Slot Allocation Strategy

As mentioned above, in cross-layer scheduling algorithm all remaining slots of a frame are allocated to highest priority connection. Due to this only one connection can transmit during each frame. This may result in underutilization of resources when all available slots are not needed for a particular connection and so they can be allotted to other connections. So, we propose modification to this algorithm so that available time-slots are distributed among half higher priority connections. Slots are distributed evenly among all these connections. Desirable feature of this modification are as follows:

- As slots are distributed among half of the connections, it will result in good average performance compared to cross-layer scheduling algorithm over time because at any frame every connection has more chance to be allocated some slots.
- There is no wastage of slots which may happen in Cross-layer scheduling algorithm when all slots allocated to particular connection are not used.

- When two connections have same priorities, Cross-layer algorithm selects randomly any one connection. This may result in very low satisfaction for other connection which is not allocated any slots but having same priority.

The simulation and analysis for original algorithm and above mentioned modification are presented in the next chapter.

Chapter 5

Simulation and Analysis

We have analyzed performance of our proposed modification compared to original cross-layer modification via simulation. Simulation tool is developed in C++. Coefficients for rtPS, nrtPS and BE are selected 1.0, 0.8 and 0.5 respectively. We have performed simulation and analysis in various scenarios:

We have compared performance on following output parameters:

- Real-Time Connection : Delay Outage Probability
- Non-Real Time & Best-Effort Connection: Average DataRate.

5.1 Simulation for 6 different connections

We have first performed simulation for 6 connections with following parameters:

- rtPS1: PER 0.01, Delay 30 ms, Guard Time 1 ms
- rtPS2: PER 0.02, Delay 50 ms, Guard Time 2 ms
- nrtPS1: PER 0.001, DataRate 10.0
- nrtPS2: PER 0.002, DataRate 8.0
- BE1: PER 0.001
- BE2: PER 0.002

Performances of above 6 connections are evaluated for 30 frames where each frame has 5 slots to be allocated. Along with above parameters, we have provided SNR (Signal to Noise Ratio) as input for each connection for each frame which indicates varying channel condition.

According to Cross-Layer Scheduling Algorithm, all 5 slots are allocated to highest priority connection while in proposed modification these 5 slots are distributed among top 3 higher

priority connections. Top 2 connections are allocated 2 slots while 3rd connection is allocated 1 slot.

Output of Simulation is as following:

1) For Original Algorithm:

DELAY OUTAGE PROBABILITY OF REAL-TIME CONNECTIONS

1st Realtime:

0.01	0.01	0.01	0.02	0.02	0.02	0	0	0.02	0.02
0	0.02	0.01	0.02	0.02	0.05	0	0.02	0.02	0
0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0

2nd Realtime:

0.05	0	0.01	0.01	0.01	0.2	0	0.01	0.01	
0	0	0	0.01	0.01	0	0.01	0.01	0.01	0
0.14	0.01	0.05	0	0	0.01	0.01	0.01	0.2	0

DATARATE PERFORMANCE OF NON-REALTIME CONNECTIONS

1st Non-Realtime:

0	0	15	11.25	9	7.5	6.43	8.48	7.54	11.23
10.21	9.36	8.64	11.48	10.72	10.05	9.46	10.7	10.13	9.63
9.17	8.75	8.37	8.02	8.95	10.41	10.02	9.66	9.33	9.71

2nd Non-Realtime:

0	0	0	10	8	6.67	5.71	5	4.44	4
3.64	3.33	3.08	2.86	4.1	3.84	3.61	3.41	3.23	3.07
2.93	2.79	2.67	2.56	2.46	2.36	3.04	3.49	3.37	3.26

DATARATE PERFORMANCE OF BEST-EFFORT CONNECTIONS

1st Best-Effort:

0	0	0	0	0	7.5	6.43	5.63	5	4.5
4.09	3.75	3.46	3.21	3	2.81	2.65	2.5	4.74	4.5
4.29	4.09	3.91	3.75	3.6	3.46	3.33	3.21	3.1	3

2nd Best-Effort:

0	0	0	0	9	7.5	6.43	5.63	5	4.5
4.09	3.75	3.46	3.21	3	2.81	2.65	2.5	2.37	2.25
2.14	2.05	1.96	1.88	1.8	1.73	1.67	1.61	1.55	1.5

2) For Modified Algorithm:

DELAY OUTAGE PROBABILITY OF REAL-TIME CONNECTIONS

1st Realtime:

0.01	0.03	0.01	0.02	0.02	0.01	0	0	0.03	0.02
0	0.04	0.01	0.02	0.02	0.01	0	0.02	0.01	0
0.03	0.02	0.01	0.03	0.01	0.02	0.02	0.02	0	0

2nd Realtime:

0.1	0	0.01	0.01	0.01	0.04	0	0.01	0.01	
0	0	0.02	0.01	0.01	0	0.01	0.01	0.01	0.02
0.03	0.02	0.1	0	0	0.01	0.01	0.01	0.04	0

DATARATE PERFORMANCE OF NON-REALTIME CONNECTIONS

1st Non-Realtime:

12	15	13	15.75	14.1	11.75	10.07	8.81	8.33	9.28
9.03	9.92	10.15	10.81	10.95	10.27	10.04	10.19	9.65	9.59
9.23	9.57	9.7	10.08	10.18	10.51	10.58	10.35	9.99	9.94

2nd Non-Realtime:

6	21	14	10.5	8.4	7	7	7.27	6.96	8.26
7.51	7.71	7.86	8.22	7.68	7.4	7.71	7.75	7.34	7.19
6.85	6.96	6.93	7.04	7.5	7.7	7.72	7.66	7.61	7.64

DATARATE PERFORMANCE OF BEST-EFFORT CONNECTIONS

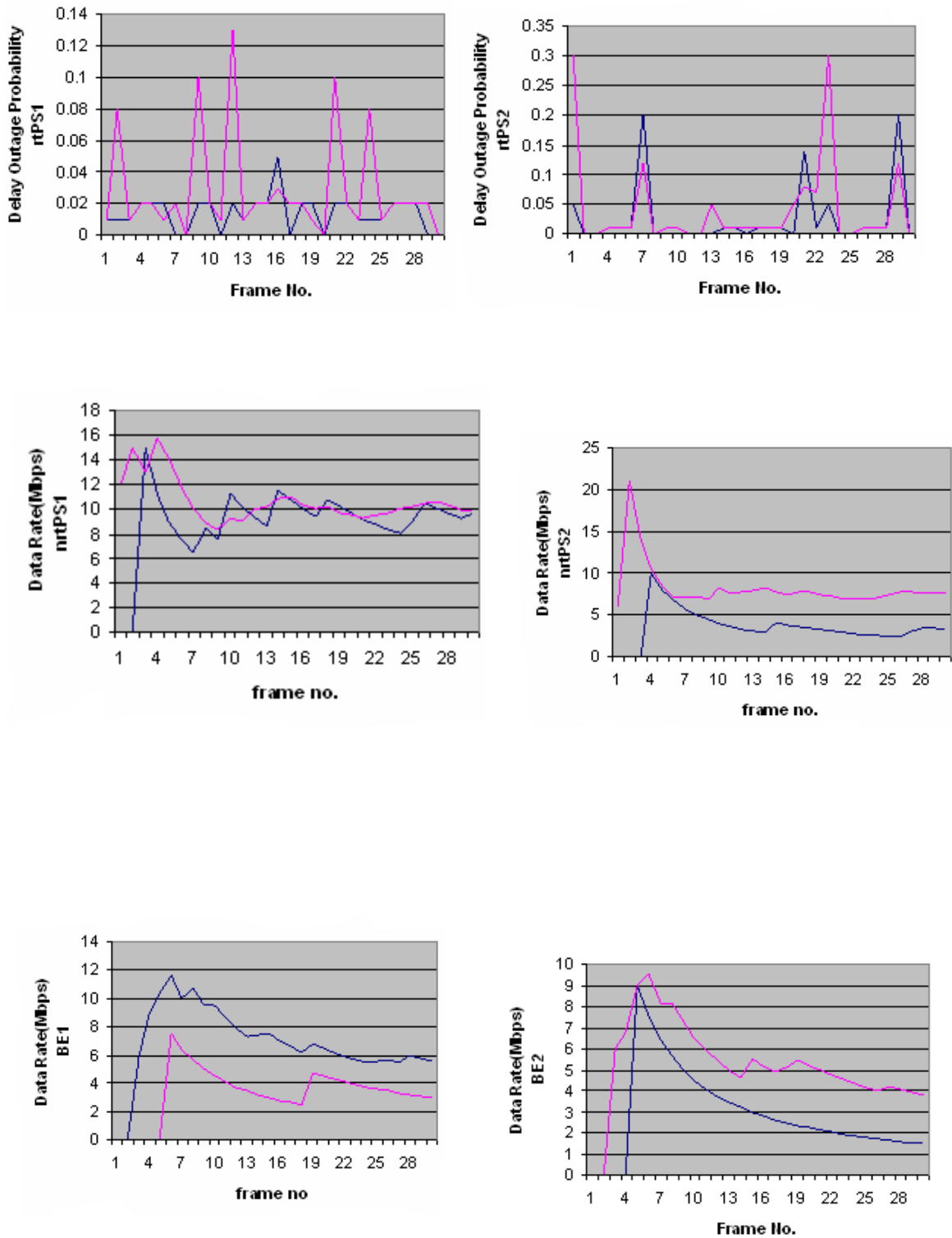
1st Best-Effort:

0	0	6	9	10.4	11.67	10	10.75	9.56	9.5
8.64	7.92	7.31	7.43	7.47	7	6.59	6.22	6.84	6.5
6.19	5.91	5.65	5.42	5.56	5.69	5.48	5.93	5.72	5.53

2nd Best-Effort:

0	0	6	6.75	9	9.5	8.14	8.13	7.22	6.5
5.91	5.42	5	4.64	5.53	5.19	4.88	5.11	5.47	5.2
4.95	4.73	4.52	4.33	4.16	4	4.19	4.04	3.9	3.77

Result of Comparison between these two algorithms for each connection is as shown in following graphs:



**Figure 5.1: Performance Comparison of two approaches for 6 connections
And 5 slots per Frame**

5.2 Simulation for 9 different connections

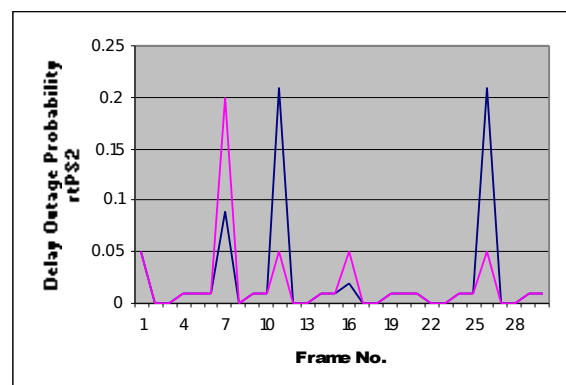
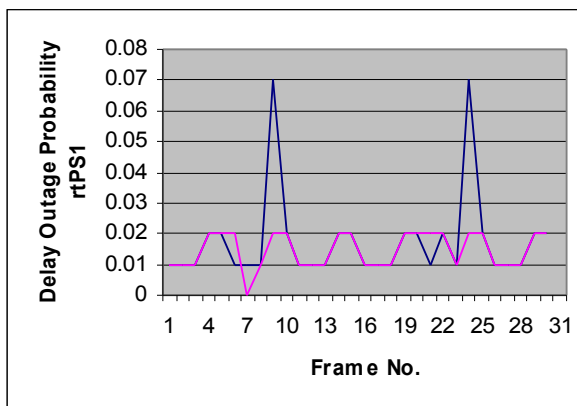
Then we have performed simulation for 9 connections with following parameters to check performance of algorithms for scalability.

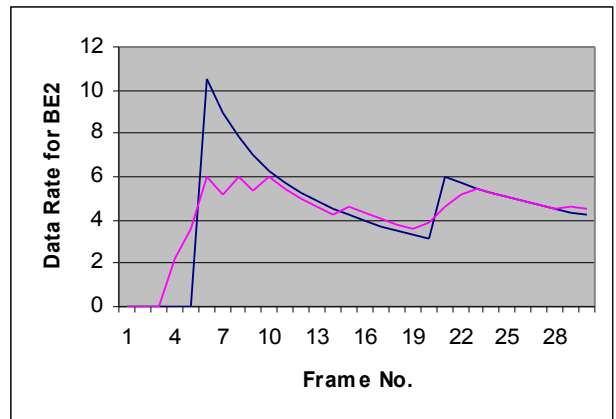
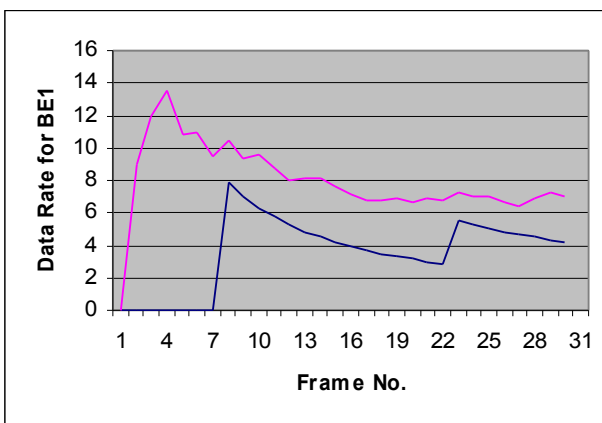
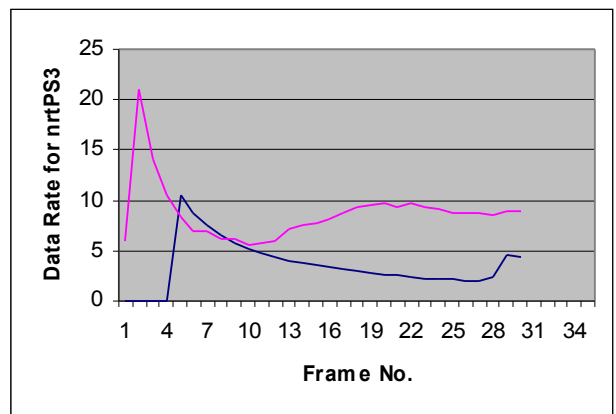
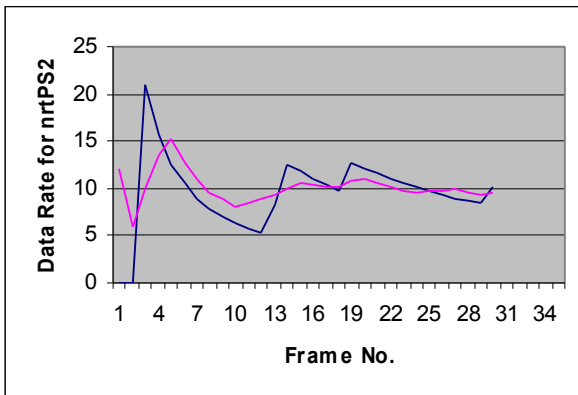
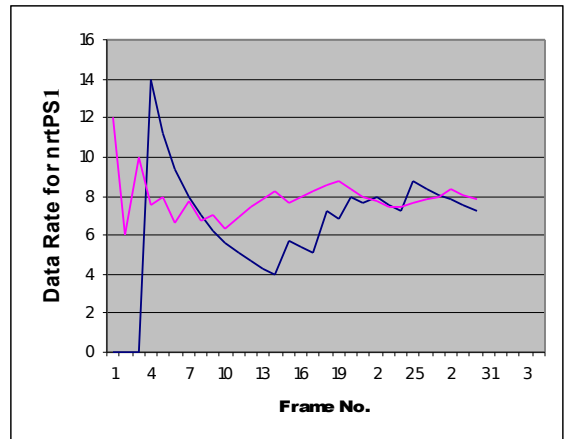
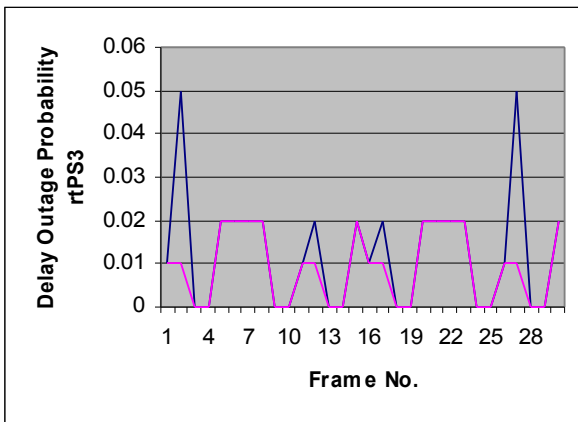
- rtPS1: PER 0.01, Delay 30 ms, Guard Time 1 ms
- rtPS2: PER 0.02, Delay 50 ms, Guard Time 2 ms
- rtPS3: PER 0.03, Delay 40 ms, Guard Time 1.5 ms
- nrtPS1: PER 0.001, DataRate 10.0
- nrtPS2: PER 0.002, DataRate 8.0
- nrtPS3: PER 0.003, DataRate 9.0
- BE1: PER 0.001
- BE2: PER 0.002
- BE3: PER 0.003

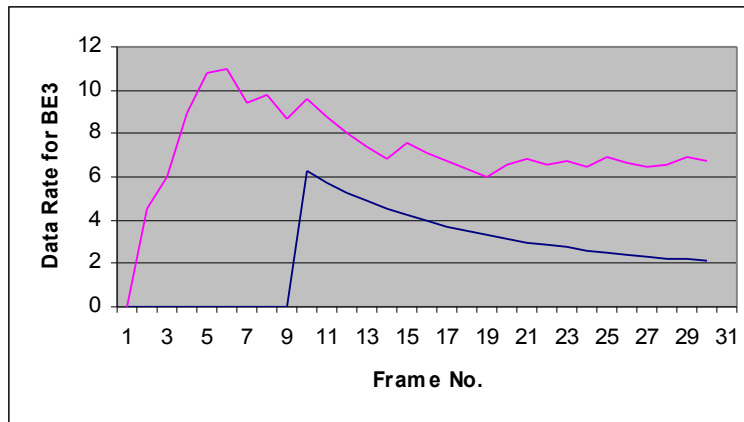
Performances of above 9 connections are evaluated for 30 frames where each frame has 7 slots to be allocated. Along with above parameters, we have provided SNR (Signal to Noise Ratio) as input for each connection for each frame.

According to Cross-Layer Scheduling Algorithm, all 7 slots are allocated to highest priority connection while in proposed modification these 7 slots are distributed among top 4 higher priority connections. Top 3 connections are allocated 2 slots while 4th connection is allocated 1 slot.

Result of Comparison between these two approaches for each connection is as shown in following graphs:



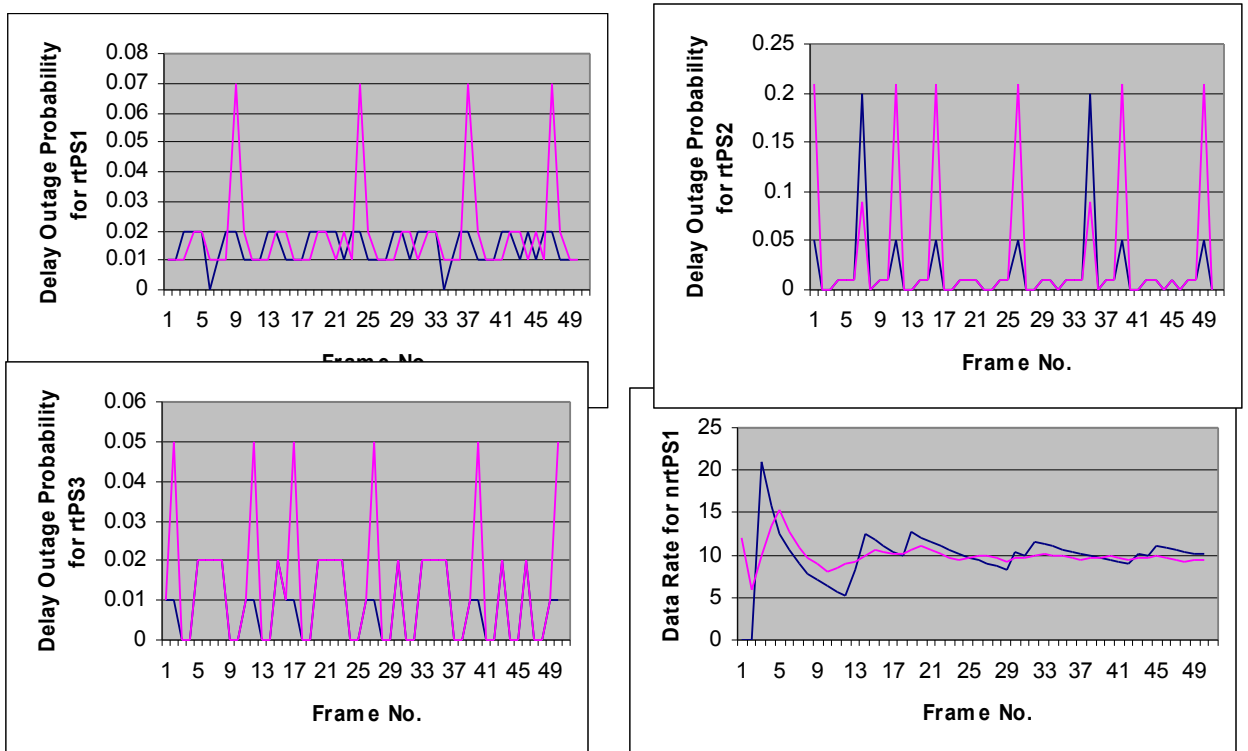




**Figure 5.2: Performance Comparison of two approaches for 9 connections
And 7 slots per Frame**

5.3 Simulation Result when numbers of frames are increased

Then we have performed simulation for same 9 connections as mentioned in section 5.2 when number of frames is increased to 50. Result of comparison is as shown in following graphs:



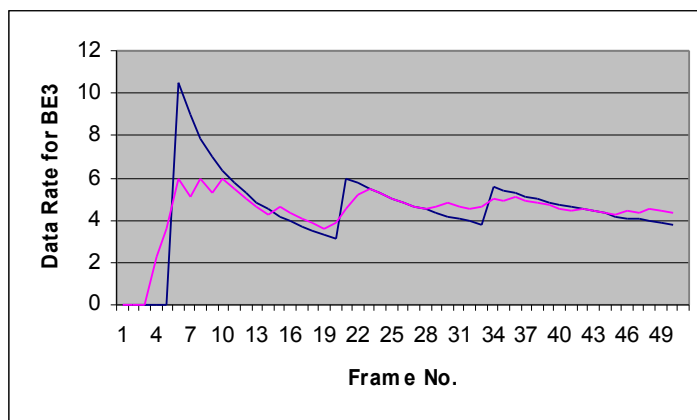
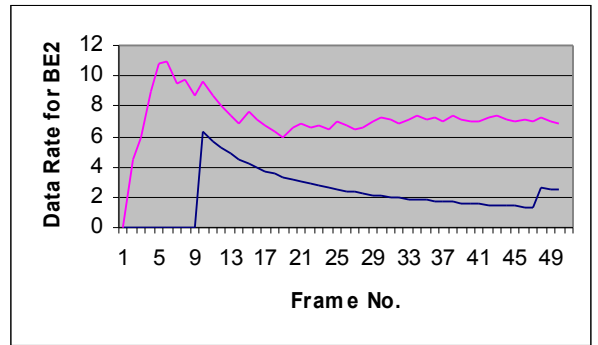
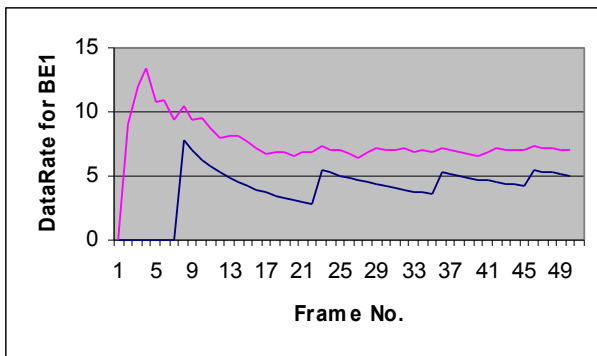
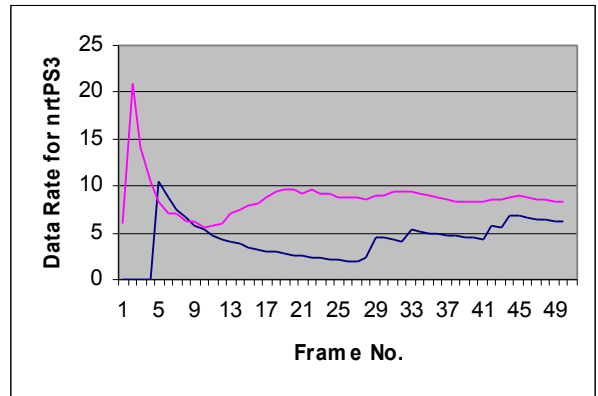
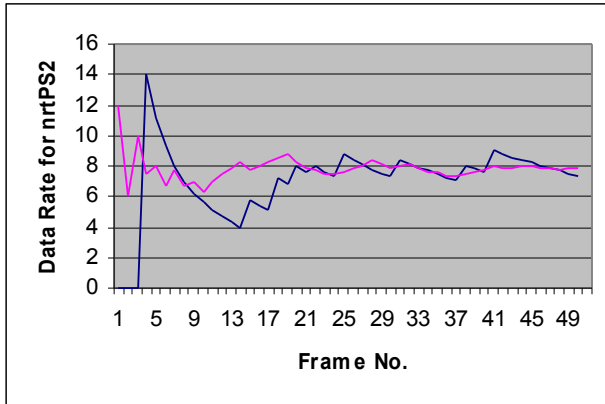


Figure 5.3 Performance Comparison of two approaches for 50 frames

From above graphs, we can see following results:

- There is considerable improvement in Data Rates for Non-Realtime and Best-Effort Connections.
- For Real-time connection delay outage probability has increased slightly but still it is in the limit of 1% for most of cases which is quite good.
- The main advantage of this approach is that there is a consistent performance of data-rate for Non-Realtime and Best-Effort connections; as a result the possibility of jitter is reduced quite significantly.
- We also observe that as number of frames is increased this consistent performance is maintained.

Thus proposed modification delivers good overall results compared to original algorithm and satisfies all desirable features mentioned in section 4. Also the algorithm is scalable as it continues to deliver good results as number of connections, number of slots and number of frames is increased.

Chapter 6

Conclusions and Future Directions

6.1 Conclusions

In this thesis we have made comparative study of various scheduling algorithms for ensuring QoS requirements in IEEE 802.16 standard. On the basis of this study, we conclude that there are three main approaches for scheduling. Out of these three approaches, approach which considers each connection separately for determining allocation of slots is found best as far as QoS requirements are to be satisfied. Also, this approach does centralized scheduling of all connection by defining different priority function for each service flow. We have made extensive study of representative Cross-Layer Scheduling algorithm which follows this approach. Then we have proposed modification to the Cross-Layer Scheduling algorithm so that available slots can be used more efficiently. Further to validate our modification we have developed simulation in C++ for both these algorithms. We have performed simulation for different situations which involves different number of connections and different number of slots to allocate. Simulation studies shows that proposed modification delivers better performance compared to original algorithm.

6.2 Extensibility

We have proposed modification for Cross-Layer algorithm which can be considered as a viable alternative for scheduling in wide variety of wireless networks like IEEE 802.11, IEEE 802.15 and cellular networks. In fact it can be considered for any network where QoS requirements have to be satisfied of different applications which are mandatory feature for emerging applications.

6.3 Future Work

Scheduling for classes of service flows in WiMAX is quite a challenging research topic. There are many issues which can be further investigated.

- In proposed modification we have distributed available slots among top half priority connections equally. Optimum allocation of available slots to various connection are worthy of further research so that we can have most efficient utilization of resources along with QoS satisfaction.
- In IEEE 802.16e-2005 new service class i.e. ErtPS is defined. To integrate this service class for slot allocation among existing four service classes and effect of performance can be further investigated.
- Performance of this algorithm in various situations like multiple BS, mobile SS are also issues to further investigate.
- Performance of proposed modification to actual network environment is also worth further study.

References

- [1] http://www.wimax.com/commentary/wimax_weekly accessed on 25/5/2008
- [2] Mark C. Wood, “An Analysis of the Design and Implementation of QoS over IEEE 802.16”
- [3] www.wimaxforum.org/technology accessed on 25/5/2008
- [4] IEEE. Standard 802.16-2004. Part16: Air interface for fixed broadband wireless access systems. October 2004.
- [5] IEEE. Standard 802.16e-2005. Part16: Air interface for fixed and mobile broadband wireless access systems—Amendment for physical and medium access control layers for combined fixed and mobile operation in licensed band. December 2005.
- [6] Jeffrey G. Andrews, [Arunabha Ghosh](#), [Rias Muhamed](#), “Fundamentals of WiMAX: Understanding Broadband Wireless Networking”, Prentice Hall
- [7] WiMAX Forum. WiMAX Forum Mobile System Profile. 2006–07.
- [8] http://www.wimax.com/education/wimax/wireless_architectures accessed on 25/5/2008
- [9] Frank Ohrtman, “WiMAX in 50 Pages: A Simple Explanation To A Complex Subject”
- [10] <http://www.wimax.com/education/wimax/antennas> accessed on 25/5/2008
- [11] WiMAX Forum. Mobile WiMAX—Part II: A comparative analysis. White Paper. April 2006
- [12] WiMAX Forum. Mobile WiMAX—Part I: A technical overview and performance evaluation. White Paper. March 2006

- [13] A.Jalali, R. Padovani, and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," Proc. IEEE VTC pp. 1854-1858, 2000, Tokyo, Japan.
- [14] Chang-an Feng, Fangmin Xu, Zheng Zhou, "A Novel Resource Allocation and Reservation Algorithm for IEEE 802.16 network"
- [15] Q.Liu, S.Zhou, and G.B.Giannakis, "A Cross-layer scheduling Algorithm with QoS support in wireless networks," IEEE Transactions on vehicular Technology, vol.55, No.3, May.2006.
- [16] Seungwoon Kim and Ikjun Yeom, "TCP-aware Uplink scheduling for IEEE 802.16". IEEE Communication Letter, Feb 2007.
- [17] H.Fattah and C.Leung, "An overview of scheduling algorithms in wireless Multimedia networks," IEEE Wireless Commun., vol.9, no.5, pp.76-83, Oct. 2002
- [18] H.Zhang, "Service Disciplines for Guaranteed Performance Service in Packet switching Networks," Proceedings of the IEEE, vol. 83, pp. 1374-96, Oct. 1995,
- [19] R.Guerin and V.Peris, "Quality-of-service in Packet Networks: Basic Mechanisms and Directions," Computer Networks, vol.31, pp.169-89, Feb. 1999.
- [20] 3GPP TR 25.848 V4.0.0, *Physical Layer Aspects of UTRA High Speed Downlink Packet Access (Release 4)*, 2001.
- [21] 3GPP2 C.S0002-0 Version 1.0, *Physical Layer Standard for cdma2000 Spread Spectrum Systems*, Jul. 1999.
- [22] J. Karaoguz, "High-rate wireless personal area networks," *IEEE Commun. Mag.*, vol. 39, no. 12, pp. 96–102, Dec. 2001.
- [23] M. S. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Kluwer J. Wireless Commun.*, vol. 13, no. 1/2, pp. 119–143, May 2000.
- [24] Qingwen Liu; Shengli Zhou; Giannakis, G.B. "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.
- [25] E. Hossain and V. K. Bhargava, "Link-level traffic scheduling for providing predictive QoS in wireless multimedia networks," *IEEE Trans. Multimedia*, vol. 6, no. 1, pp. 199–217, Feb. 2004.

- [26] C.L.Liu, and J.W. Layland, "Scheduling algorithms for multiprogramming in a hard real time environment", Journal of the Association for Computing Machinery, Vol.20,no.1, pp. 44-61, Jan. 1973

List of Papers Published/Communicated

- [1] **Mohit Saini, Ashish Jain and Anil K. Verma, "Comparison of Wi-Fi and Wi-MAX Wireless Technologies", in 11th Punjab Science Congress February 7-9, 2008 (Status: Accepted)**
- [2] **Ashish Jain and Anil K. Verma, "Comparison of different scheduling algorithm for WiMAX (802.16)", in National Conference on Mobile and Pervasive computing (CoMPC -2008) August 7-8, 2008 (Status: Submitted)**
- [3] **Ashish Jain and Anil K. Verma, "Efficient Scheduling Algorithm for QoS Support in WiMAX", in Fifth International Conference on Computer Networks And Security (ICCNS 2008) September 27-28, 2008 (Status: Submitted)**