Resource Management for Cross Layered Star and Mesh Networks

Ayda Basyouni

A Thesis
In the Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements For the Degree of Doctor of Philosophy at
Concordia University
Montreal, Québec, Canada

September, 2008

Your file Votre référence
ISBN: 978-0-494-45649-1
Our file Notre référence
ISBN: 978-0-494-45649-1

ABSTRACT

Resource Management for Cross Layered Star and Mesh Networks
Ayda Basyouni, Ph.D.
Concordia University, 2008

The use of wireless services is rapidly spreading around the world and many of the world population no longer know how to cope without their cell phones; the feel of always being connected offers a great sense of flexibility and security. So far, voice has been the primary wireless application. However, with the Internet continuing to influence our daily lives, the demand for wireless data is extensively increasing. Already, in the countries that have cellular-data services readily available, the number of cellular subscribers taking advantage of data services has reached significant proportions.

In this thesis, we investigate resource management techniques for cross layered star and mesh wireless data networks. In particular, we investigate several aspects related to resource management techniques over the reverse packet data channel in cdma2000 1xEV star networks. We provide an upper bound for the reverse packet data channel throughput as a function of the number of mobile stations that are allowed to transmit instantaneously on each time slot. We also provide a lower bound for the average sector throughput based on the number of users per sector and propose several autonomous rate assignment, and scheduling techniques that provide a significant throughput improvement relative to other published techniques.

We also develop analytical models for lowest-rate-first, highest-rate-first priority scheduling techniques, and two round-robin fair scheduling techniques over

the reverse data channel in cdma2000 1xEV star networks. For these four scheduling techniques, the distribution of the mobile stations among the possible data rates is modelled as a Markov process. An analytical expression for the steady state system throughput is derived from the steady state distribution of the above Markov process. The above model is extended to evaluate the performance of the cross layered design between the hybrid ARQ, rate assignment, and time slot scheduling over the reverse packet data channel in cdma2000 1xEV. Expressions for the steady state system throughput and file transmission delay are derived from the steady state distribution of this model.

Fountain codes are a class of erasure codes with the property that a potentially limitless sequence of encoding symbols can be recovered from any subset of size equal to or slightly larger than the number of source symbols. In this thesis we relate the parameters of the higher layer Fountain codes to those of the physical layer codes to present a cross layered coding technique. Based on the bit error rate of physical layer codes, the upper layer Fountain code's parameters are designed to obtain a prespecified performance. The performance of the proposed cross layered coding technique is found to be comparable to that of physical layer based Hybrid ARQ technique. As an application we studied the performance of WiMAX backhaul mesh networks where cashing is allowed at the service stations and proposed cross layered coding technique is used as its main coding scheme. In particular, the performance of the above network is modelled as a Markov process and analytical expressions for the steady state system performance are derived from its associated steady state distribution.

All the above analytical models are validated through simulations.

To my Father, who has been dreaming about this day since I was a child.

To my Mother, may Allah bless her.

Acknowledgement

I would like to express my sincere gratitude and thanks to my supervisors, Dr. Ahmed Elhakeem, and Dr. Anjali Agarwal for their valuable support, enlightening guidance, and sincere encouragement throughout the course of this work.

Special thanks to my husband, Amr, for his love, understanding, and support, and to my children Reem and Maryam, whose smiles make all the difference to my world.

Words fall short of expressing my love, appreciation, and gratitude for my parents in-law, family and friends for being always there for us, and giving us the hand whenever we needed it. Without their love, encouragement and support this work would have never been possible.

I also wish to acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), and by Concordia University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Acronames

1xEV-DV	Evolution-Data Voice
1xEV-DO	Evolution-Data Optimization
1xRTT	Single Carrier Radio Transmission Technology
3G	Third Generation
4G	Forth Generation
3GPP2	Third Generation Partnership Project 2
8PSK	8 Phase Shift Keying
ACK	Acknowledgement
BW	Bandwidth
BS	Base Station
BPSK	Binary Phase Shift Keying
CDF	Cumulative Density Function
CDMA	Code Division Multiple Access
CRC	Cyclic Redundancy Code
DL	Down Link
DRC	Data Rate Control Channel
DS/CDMA	Direct Sequence CDMA
FL	Forward Link
F-PDCH	Forward Packet Data Channel
F-SCH	Forward Supplemental Channel
F-GCH	Forward Grant Channel
F-ACKCH	Forward Acknowledgement Channel
F-RCCH	Forward Rate Control Channel
H-ARQ	Hybrid Automatic Repeat Request
HSDPA	High Speed Down-Link Packet Access
IS-95	Interim Standard 95
Kbps	kilo bit per Second
Mbps	Mega bit per Second
MHz	Mega Hertz
MS	Mobile Station
MAC	Medium Access Control
MSIB	Mobile Status Indicator Bit
NAK	Not Acknowledged

OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
QAM	Quadrature Amplitude Modulation
RL	Reverse Link
R-PDCH	Reverse Packet Data Channel
R-SCH	Reverse Supplemental Channel
R-PDCCH	Reverse Packet Data Control Channel
R-REQCH	Reverse Request Channel
RRI	Reverse Rate Indicator
RoT	Rise Over Thermal
SINR	Signal to Noise and Interference Ratio
TD	Time Division
TDM	Time Division Multiplexing
TDD	Time Division Duplex
TCP	Transmission Control Protocol
UMTS	Universal Mobile Telecommunications System
UDP	User Datagram Protocol
UL	Up Link
WiMAX	Worldwide Interoperability for Microwave Access
WCDMA	Wideband CDMA

Chapter 1

Introduction

Ever since people wanted to talk to other people who were out of earshot, there have been attempts at communication. These attempts have advanced from carrier pigeons, signalling towers, smoke signals through telegraph, telephone, radio and satellite. Technology has advanced to allow TV, Internet, high speed data transmission. During the past years, there has been a quickly rising interest in radio access technologies for providing mobile as well as fixed services for voice, video and data. The 3G/4G technologies such as cdma2000 1xEV, WCDMA, Wi-Fi and WiMAX enable network operators to offer users a wider range of more advanced services while achieving greater network capacity through improved resource management. The semi-annual wireless industry survey published by CITA at the end of December 2007 [1] shows the growth in the wireless industry. At the end of 2007, the number of wireless subscribers had reached almost 256 million with an increase of 22 million subscribers since the end of 2006. The total twelve-month revenues reached more than 138 billion dollars in 2007.

1.1 Motivations and Objectives

Mobile systems have recently evolved from being simple devices used for voice communications to systems that support new features such as data transmissions, video streaming, and Internet access. The capacity demands to handle these new services have led to a new era of wireless technologies, which resulted in a huge cellular market all over the world. The providers of this technology have to offer attractive services that are reasonably priced. Efficient resource management and standards that provide high data rates are the tools used by those providers to ensure quality of service (QoS) with affordable prices.

QoS could be satisfied by different resource management means such as, call admission control [2]-[12], power/rate control [13]-[31], and scheduling [32]-[45]. With the new era of wireless technologies, the need for dynamic resource allocation and optimized scheduling techniques arises in order to efficiently use the available bandwidth (BW). Packet delays, user and sector throughputs and fairness between users, are also important issues to investigate. The performance of the backhaul networks is also considered to be an important factor in the overall network performance. Recently the cross layer design [67]-[74] was adopted to fully optimize wireless networks. Rate, power and coding at the physical layer can be adapted to meet the requirements of the applications given the current channel and network conditions. In cross layer design knowledge is shared between layers to obtain the best performance.

In wireless networks a star topology consists of one central base station (BS), relay, or wireless modem and several end nodes such as mobile stations (MSs), and computers (see Fig. 1.1). All the end nodes are connected directly to the central hub. In mesh networks each node could be connected to other nodes directly or through multiple hubs (see Fig. 1.2). Fully connected mesh network is a network where all the nodes are connected to each other. In mesh networks a routing algorithm is applied to determine the best route from any node to any other node.

In the last few years wide range of wireless technologies were standardized. Throughout this thesis we consider CDMA-based standards as a star network. In particular, we studied cdma2000 1xEV-DV (evolution-data voice) revisions C and D [47] [48], and single carrier evolution data-optimization (1xEV-DO) [49] star networks. We also consider the IEEE 802.16e worldwide interoperability for microwave access (WiMAX) [50] as a backhual mesh network. It is important to mention that WiMAX standard supports both star and mesh topologies, however in this thesis we only consider the mesh topology.

The objectives of this thesis is to improve the resource management techniques associated with the cdma2000 1x reverse link data channel. We also study the improvements in the performance of WiMAX backhaul network. In particular, we provide an upper bound for the reverse packet data channel throughput as a function of the number of mobile stations that are allowed to transmit instantaneously on each time slot. We propose several rate assignment, and scheduling

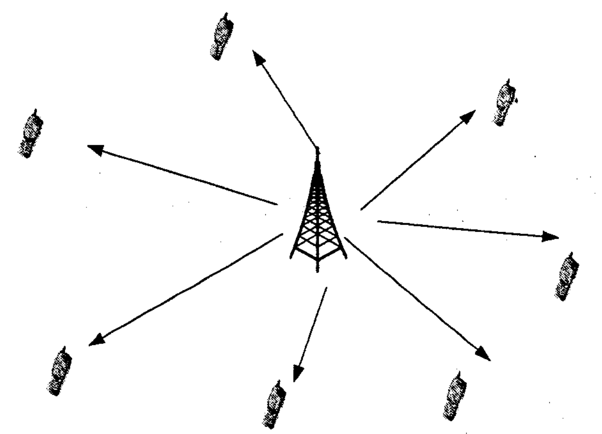


Figure 1.1: An example of star network

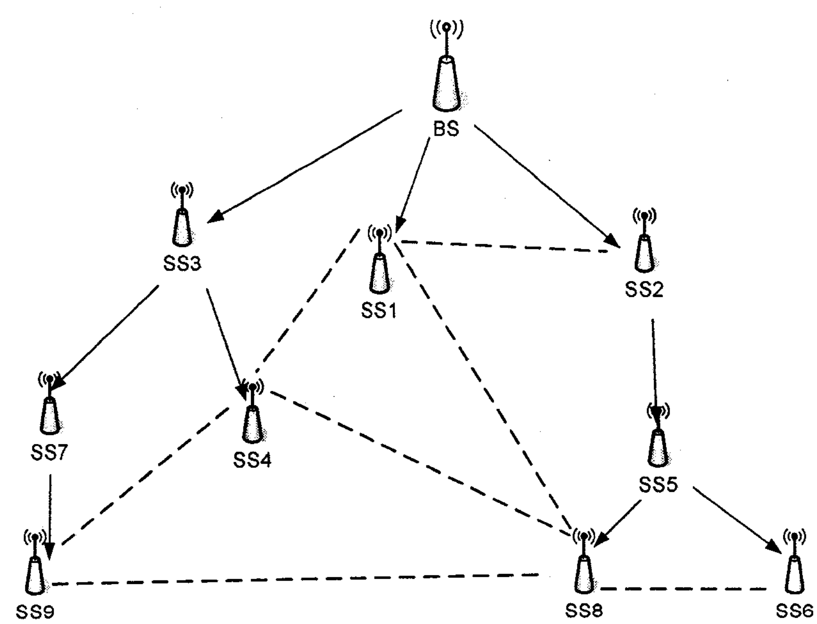


Figure 1.2: An example of mesh network

techniques that provide a significant throughput improvement relative to other published techniques in cdma2000 1x reverse link data channel. We also develop analytical models for lowest-rate-first, highest-rate-first priority scheduling techniques, and two round-robin fair scheduling techniques over the reverse data channel in cdma2000 1xEV star networks. For these four scheduling techniques, the distribution of the mobile stations among the possible data rates is modelled as a Markov process. An analytical expression for the steady state system throughput is derived from the steady state distribution of the above Markov process. The above model is extended to evaluate the performance of the cross layered design between the hybrid ARQ, rate assignment, and time slot scheduling over the reverse packet data channel in cdma2000 1xEV. We also investigate some techniques that could be used to improve the performance of backhaul mesh networks, specifically,we proposed cross layer coding technique in which we relate the parameters of the higher layer Fountain codes to those of the physical layer codes to present a cross layered coding technique. Based on the bit error rate of physical layer codes, the upper layer Fountain code's parameters are designed to obtain a prespecified performance. As an application we studied the performance of WiMAX backhaul mesh networks where cashing is allowed at the service stations and proposed cross layered coding technique is used as its main coding scheme.

1.2 Outline

The rest of the thesis is organized as follows. In the rest of this chapter we briefly describe the standards considered in this thesis and review the resource management techniques in the literature.

In Chapter 2 we consider a cdma2000 1xEV star network, for which we derive upper and lower bounds for the reverse packet data channel throughput as a function of the number of mobile stations that are allowed to transmit instantaneously on each time slot. We propose several autonomous rate assignment, and scheduling techniques that provide a significant throughput improvement relative to the other published schemes.

In Chapter 3 and based on the results of Chapter 2 we develop an analytical model for several scheduling techniques over the reverse data channel in cdma2000 1xEV. The models are extended to include the effect of the cross layer design between H-ARQ, rate assignment and scheduling. For these scheduling techniques, the distribution of the mobile stations among the possible data rates is modelled as a Markov process. An analytical expression for the steady state system throughput is derived from the steady state distribution of the above Markov process. The developed models are validated through simulations.

In Chapter 4 we proposed a cross layered coding technique. The performance of the proposed coding technique is found to be comparable to that of physical layer based Hybrid ARQ technique. As an application of the multi layer coding technique proposed, Markov process is used to model and analyze a WiMAX backhaul mesh network where caching is allowed at service stations. Throughout our analysis the proposed cross layered coding technique is used as the WiMAX coding technique. An analytical expressions for the steady state system performance are derived from the steady state distribution of the above Markov process. Results of the developed model is validated through simulations.

Finally, in Chapter 5 we discuss the conclusions and future work.

Some of the work presented in this thesis has been published in [101]-[108].

1.3 Standards

1.3.1 The cdma2000 1xEV

The IS-95A was the first CDMA- based wireless system deployed in the nineties [51]. However, the cdma2000 1xRTT (single carrier radio transmission technology) [53] is considered to be the first phase in CDMA2000 evolution. The cdma2000 1xEV-DO (evolution-data optimization) [54], which classified as a third generation (3G) system, is the third phase in CDMA2000 evolution. In this version of the standards, the data rates reached up to 2Mbps. The cdma2000 1xEV-DV (evolution-data voice) revisions C and D [47] [48] were real evolution on the forward and reverse links. By creating dedicated physical channels to carry high speed data from the base station (BS) to the mobile station (MS) and vise versa, the data rates on the reverse and forwarded links reached the third generation set of goals. In what follows we briefly describe both cdma2000 1xEV-DO and 1xEV-DV standards.

1.3.1.1 The cdma2000 1xEV-DO

In the last few years, the third generation partnership project two (3GPP2) approved the single carrier evolution data-optimization (1xEV-DO) [49] standards to satisfy the demands for high data rate wireless networks.

In 1xEV-DO, each MS transmits on the reverse traffic channel which consists of a data channel, reverse rate indicator (RRI), pilot channel, data rate control channel (DRC), and an acknowledgement channel (ACK). Each channel in the reverse traffic channel is spread by an appropriate orthogonal Walsh function. A slot is a basic transmission unit, and is 1.666 ms long (2048 chips). A group of 16 slots is referred to as a frame.

The data channel supports five data rates (see Table 1.1) ranging from 9.6 kbps to 153.6 kbps with 26.66 ms time frame packets. The forward link supports data rates up to 2.4 Mbps. The RRI channel indicates the data rate of the associated data channel. The pilot channel is used for channel estimation and coherent detection. The DRC channel informs the access network of the best serving cell and the supportable data rate on the forward traffic channel. A DRC message is repeated over DRCLength (1, 2, 4, or 8 slots). The ACK channel informs the access network whether a packet transmitted on the forward traffic channel has been

R	Data Rates in Kbps	$TPRD_k$ in dB
R_1	9.6	3.75
R_2	19.2	6.75
R_3	38.4	9.75
R_4	76.8	13.25
R_5	153.6	18.5

Table 1.1: Traffic to pilot ratio for 1xEV-DO data channel, channel BW = 1.25 MHz

received successfully.

1.3.1.2 The cdma2000 1xEV-DV

The cdma2000 1xEV-DV (evolution-data voice) revisions C and D [47] [48] were a real evolution on the forward and reverse links. By creating the forward packet data channel (F-PDCH) and the reverse packet data channel (R-PDCH), dedicated physical channels to carry high speed data from the BS to the MS and vise versa, the data rates on the reverse and forwarded links reached the third generation set of goals.

The BW of cdma2000 1xEV-DV is 1.25 MHz. On the R-PDCH, each MS can transmit using one of the eleven allowed rates (see Table 1.2). Two power levels are associated with each rate (normal or boosted). Throughout our thesis, we assume that only normal transmission power gain is used.

Fig. 1.3 shows the transmission process on R-PDCH. Cyclic redundancy check (CRC) bits and encoder tail bits are added to the information bits to form the en-

R	Data Rates in Kbps	$TPRD_k$ in dB
R_1	19.2	0.75
R_2	40.8	3.75
R_3	79.2	6.75
R_4	156.0	9.625
R_5	309.6	11.875
R_6	463.2	13.625
R_7	616.8	14.875
R_8	924.0	16.625
R_9	1231.2	18
R_{10}	1538.4	19.125
R_{11}	1845.6	21.25

Table 1.2: Traffic to pilot ratio for 1xEV-DV R-PDCH, channel BW = 1.25 MHz

coder packet. The encoder packet is fed into a turbo encoder with rate 1/5 and then interleaved. Each encoded packet is divided into three sub-packets. The sub-packet symbol selector selects the symbols to be transmitted in each sub-packet such that the effective code rate becomes lower as more sub-packets are transmitted [55]. The sub-packet is modulated using BPSK, QPSK or 8-PSK depending on the user's data rate. The hybrid automatic repeat request (H-ARQ) protocol is applied to sub-packets transmissions (Fig. 1.4). Each user is assigned four ARQ stop and wait channels, each of 10 ms. Each data packet will use one of these ARQ channels to transmit its sub-packets. All the ARQ channels could have the same data rate or different data rates. The ARQ channel keeps the same data rate as long as it is transmitting the same packet. Incremental redundancy technique is used, where a subset of the turbo encoded packet is sent (sub-packet). This first sub-packet will have an encoding rate higher than the original encoder. As more redundancy (sub-packets) are transmitted, the received data packets reach the original encoding rate.

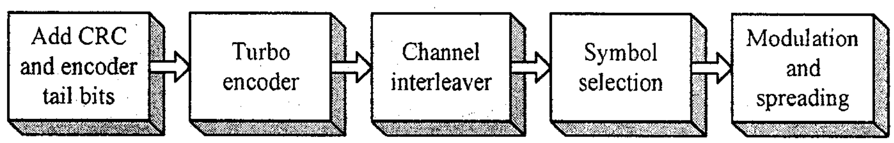


Figure 1.3: Transmission process on R-PDCH

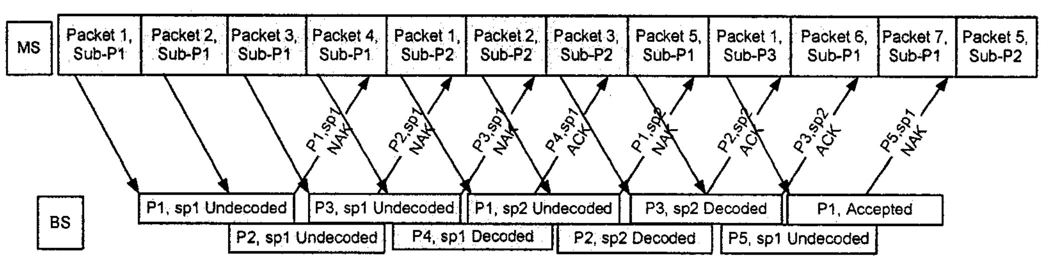


Figure 1.4: H-ARQ on the R-PDCH

If the BS decodes the information correctly, an ACK will be transmitted and the MS will not send the rest of this packet. A NAK will be transmitted if the information is not decoded correctly and the MS sends the successive sub packets. The H-ARQ protocol has a 40 ms cycle (4 slots) from the time the MS starts transmitting the sub-packet until it receives the ACK/ NAK and processes it. The 1xEV turbo encoder has a final coding rate of $R_c = 1/5$. Each encoded packet is divided into three sub-packets. The first sub-packet has encoding rates of $R_c = 1/2$, the first and second sub-packets combined will have an effective lower encoding rate of $R_c = 1/3$, the three sub-packets combined will reach the final encoding rate of $R_c = 1/5$.

The R-PDCH is supported by the reverse packet data control channel (R-PDCCH) which is used to transmit control information that correspond to the current transmission on the R-PDCH. Amongst these, the size of the data packet, which sub packet is currently transmitted, whether or not the transmission is done using elevated power levels, and an indicator if the MS has enough power and data to transmit above the current rate. The reverse request channel (R-REQCH) transmits the MS's current application (FTP, video streaming, interactive gaming, etc.), the amount of available data, and its current power head room. The forward rate control channel (F-RCCH) allows the MS to change its current rate by one step (higher or lower). Forward grant channel (F-GCH) allows the MS to change its current rate to the granted rate. The granted rate could be any one of the allowed eleven rates. The forward acknowledgement channel (F-ACKCH) provides the MS with an ACK or NAK for the transmitted sub packet.

The rate assignment mechanisms supported by the 1xEV reverse data channel can be categorized as autonomous, differential, or absolute rate schemes [55].

In autonomous rate assignment schemes, the BS grants each MS a rate from the set of allowable rates. The MS is then free to transmit using this rate or any other rate below it. This technique is considered the best option for delay sensitive applications and it has the least computational overhead.

Differential rate assignment allows the MS to change its rate by one step up or down. The BS can send one of three commands (UP, Down, or Hold). The problem with this technique is ramping up delays; it takes relatively long time to reach high data rate transmissions.

In absolute rate assignment schemes, the BS authorizes the MS to send on a certain rate. The grant message is initiated when the BS receives a change rate request from the MS.

The 3GPP2 [56] requires that all resource scheduling techniques, for 1xEV standards, to satisfy the fairness criterion. Fairness is defined as providing all mobile stations with a minimal level of throughput and is evaluated by determining the cumulative density function (CDF) of the normalized throughput, with respect to the average user throughput, for all users. In order to satisfy the fairness criterion, the standard specifies that the CDF shall lie to the right of the curve given by the three points in Table 1.3. The first point in the table requires that the normalized user throughput of at least 90% of the users exceeds 0.1. Similarly, the second and third points require that the normalized user throughput of at least 80% and 50% of the users exceeds 0.2 and 0.5 respectively.

Good surveys on the cdma2000 1xEV-DV system can be found in [57] [58] [13] [59] [55] [60]. In [57], Derryberry and Pi describe the reverse link (RL) physical layer enhancements introduced to support 1xEV-DV. The forward link (FL)

Normalized throughput	Cumulative density function (CDF)
0.1	0.1
0.2	0.2
0.5	0.5

Table 1.3: Fairness criterion

physical layer enhancements can be found in [58]. Kwon *et al.* [13] discuss the system performance evaluation for both FL and RL, describe how the evaluation was conducted and analyze the results. The hybrid automatic repeat request (H-ARQ) mechanism used in release D is described in [59]. This modified technique allows for variable retransmission power levels resulting in four modes of operation: normal, reduction, boost and boost reduction modes. In [55], Comstock *et al.* describe the design enhancements of the upper layers of the cdma2000 revision D, including basic channel operation, the multiplexing of traffic, H-ARQ and mechanisms for managing RL radio resources. Kim and Ti [60] look at the new system from practical prospective; they discuss the needs, services and market drivers and provide the market and revenue statistics for voice and data services.

1.3.2 WiMAX

The rapid increase in user demands for faster connections to the Web and VoIP services has led to faster developments in wireless communications technology.

WiMAX [50] is rapidly proving itself as a key player in the area of long-range, high-speed fixed and mobile wireless technology. It also provides fast wireless solution for backhaul systems.

The physical layer of WiMAX is based on orthogonal frequency division multiple access (OFDMA) [62]. The OFDMA can be thought of as a multi-user version of the orthogonal frequency division multiplexing (OFDM). In OFDM, the BW is subdivided into multiple frequency sub-carriers. Also, the input data stream is divided into several parallel sub-streams of reduced data rate (thus increased symbol duration) and each sub-stream is modulated and transmitted on a separate orthogonal sub-carrier. The increased symbol duration improves the robustness of OFDM to delay spread [64]. Multiple access in OFDMA is achieved by assigning subsets of sub-carriers to individual users, which allows simultaneous transmission from different users.

The WiMAX OFDMA structure consists of three types of sub-carriers: data sub-carriers for data transmission, pilot sub-carriers for estimation and synchronization purposes, and null sub-carriers as guard bands [64]. Data and pilot sub-carriers are grouped into subsets of sub-carriers called sub-channels, the sub channelization is supported in both down link and up link [50].

Each sub-carrier can be modulated using one of the following modulation and coding techniques shown in Table 1.4 where $SINR_k$ denotes the signal to interference and noise ratio required by each scheme to achieve a BER of 10^{-6} [50]. The parameters of WiMAX up-link for 5 and 10 MHz BW are shown in Table 1.5.

	Modulation Technique	Coding rate C_k	Data bits per sub-carrier d_k	$SINR_k$ in dB
M_1	QAM4	1/2	1	5
M_2	QAM4	3/4	1.5	8
M_3	QAM16	1/2	2	10.5
M_4	QAM16	3/4	3	14
M_5	QAM64	2/3	4	18
M_6	QAM64	3/4	4.5	20

Table 1.4: Allowable modulation techniques

Both convolutional code and convolutional turbo code with variable code rate are supported. Hybrid Automatic Repeat Request (H-ARQ) is also supported by WiMAX to enhance the system performance.

The frame structure for a WiMAX time division duplex (TDD) implementation is divided into down link (DL) and up link (UL) sub-frames separated by transition gaps to prevent DL and UL transmission collisions. Each frame has its control information to ensure optimal system operation. As shown in Fig. 1.5, each UL and DL sub-frame is further divided horizontally to symbol durations and vertically to sub-channels and each frame is transmitted over T_f time [64].

1.4 Literature Review

As we will cover more than one topic in this thesis, our literature review is divided into two parts. In the first part, we discuss the currently available rate

	5 MHz BW	10 MHz BW
Total number of sub-carriers (N_{FFT})	512	1024
Number of Null sub-carriers	104	184
Number of Pilot sub-carriers	136	280
Number of Data sub-carriers	272	560
Number of used sub-carriers (N_{used})	408	840
Number of sub-channels N	17	35
Number of sub-carriers N_c per sub-channel	16	16
Symbol duration (T_s) micro sec	102.9	102.9
Frame duration (T_f) msec.	5	5
Number of data symbols/frame (N_{sym})	44	44

Table 1.5: Wimax parameters [64]

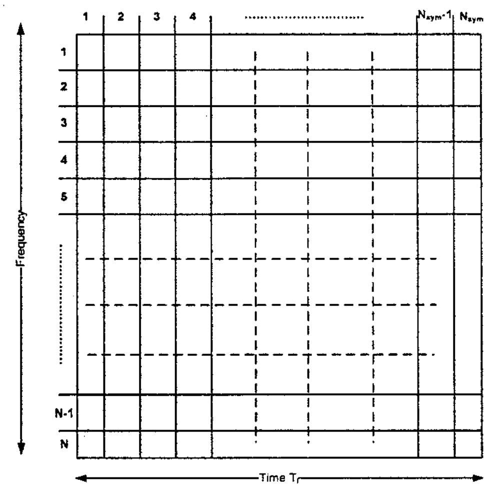


Figure 1.5: Data time frame of T_f msec

assignment techniques. In the second part, we review the relevant scheduling schemes.

1.4.1 Rate Assignment

One should mention that the methodology by which the BS determines the appropriate data rate for each MS is not specified in the standards. In the rest of this section, we give a review of the current rate assignment methodologies.

Kwon *et al.* [13] [14] proposed a hybrid rate technique for the 1xEV-DV system. It is called rate control with quick start (RCQS) in which high data rates are assigned to the users at the beginning in order to avoid the delay of ramping up. However, they didn't provide any specific details for the rate assignment scheme.

Shu and Niu [15] proposed a dynamic rate assignment based on computing the optimum number of simultaneous transmissions on one time slot with respect to multiple access interference. In this scheme, the authors assume that all the users have the same packet size, each user is allowed to send more than one packet per slot, and the number of packets are computed as a function of the optimum number of transmissions and the buffer length of each user.

Lee, Yeo, and Cho [16] [18] [20] proposed rate assignment techniques for the 1xEV-DO system. The techniques are based on modelling the rate control as discrete Markov process and adopting the RoT as the traffic load measure. These techniques were proposed for a set of five data rates (1xEV-DO allowable data rates).

Vannithamby and Sousa [21] provided an analytical model for data rates on the forward link in WCDMA. Two cases are considered: only one user is transmitting per slot, and more than one user are transmitting at the same time. They compared more than one technique. Their techniques were mainly based on the signal to interference ratio required to provide QoS, transmitted power allocated for mobiles, and the normalized interference to each mobile.

Rodriguez and Goodman [22] addressed power and data rate allocations for each terminal, such that the network weighted throughput is maximized. The weights admit various interpretations, including levels of importance, "utility", and price. They stated that at least one terminal should operate at the highest available data rate. Lowering the highest available data rate increases the number of terminals which should operate at maximum data rate. They utilized a model which can accommodate many physical layer configurations of practical interest.

Rodriguez, Goodman, and Marantz [23] investigated power and data rate allocations that maximize the network weighted throughput. Each terminal has one of two possible weights, which admits various practical interpretations. In their work, they introduced a general procedure to seek a global optimizer allocation. Their analysis was based on classical optimization theory, and accommodate a wide variety of physical layer configurations.

Bjorklund *et al.* [24] studied the problem of forward link (FL) bandwidth allocation in WCDMA networks. They proposed an algorithm for assigning transmission rates to the users with the objective of optimizing some performance criterion such as the total throughput. They formulated the problem using one multiple-choice knapsack model per cell.

Price and Javidi [26] considered a WCDMA network with arbitrary but known layout and variable rate assignments, which is connected to a traditional wired network. They showed that by using an optimization framework it is possible to construct a distributed rate assignment algorithm which addresses issues like interference and congestion control. They formulated this as a maximization problem subject to interference and congestion constraints, and developed a distributed algorithms to solve this maximization problem.

Price and Javidi [27] addressed the QoS requirements and continually changing resource demands. In that paper, they examined such optimal resource allocation through non-uniform rate assignment in a CDMA system. In particular, they focused on reverse link rate assignment at the MAC layer, and showed that a pricing structure can be used to de-centrally regulate each mobile's bandwidth consumption (transmission rate) based on the air-link's interference. Using an optimization formulation, they showed that unequal rate assignments perform better than equal-rate assignments, particularly as the density of mobiles at the boundary of two cells increases.

Fattah and Leung [29] proposed a load-based transmission rate (LTR) assignment scheme for non-real time data services in an integrated voice/data direct sequence code division multiple access (DS/CDMA) system. The LTR scheme optimally determines the transmission rate for each session according to its individual load to minimize the overall average packet transfer delay.

Ci and Guizani [30] [31] proposed an optimal rate assignment technique for the 1xEV-DV forward link. The scheme was proposed to support delay bounded multimedia services. In their analysis, they consider one user with mixed traffic per time slot. The scheme's objective is to maximize the one slot capacity subject to some constraints such as the number of available Walsh codes, service type, packet length, and maximum allowable data rate.

Kulkarni *et al.* [82] addressed the problem of assigning sub-carriers and bits to point-to-point wireless links in the presence of co-channel interference and Rayleigh fading in WiMAX networks. The objective is to minimize the total transmitted power over the entire network while satisfying the data rate requirement of each link. Simulation results show that the approach results in an efficient assignment of sub-carriers and transmitter power levels in terms of the energy required for transmitting each bit of information.

In [74] the authors proposed a cross-layered resource allocation scheme over wireless relay networks with the objective of maximizing the relay network throughput subject to a given delay QoS constraint.

1.4.2 Scheduling Techniques

Efficient Scheduling algorithms are essential to guaranteeing QoS such as delay and throughput. The design of scheduling schemes for the wireless systems is challenging due to the high error rates in the wireless links. With the increasing demand for wireless data services, the demand for optimum scheduling techniques is also increasing. The perfect wireless scheduling technique should satisfy many aspects. In particular, it should maximize throughput, minimize the delay, and tolerate the high error rates on the wireless links. During the last few years, many scheduling techniques were proposed. Some of these proposals were inherited from the wired links such as the first come first serve scheduler, fair queuing technique, weighted fair queuing [35] [41]. In our review, we focused only on the ones designed specially for wireless systems. In the following we review some of the available wireless scheduling schemes.

Jalali, Padovani, and Pankaj [32] introduced the idea of proportional fair scheduling to improve the system throughput; the users at a certain time are given priorities based on their current rates. The priority metric for any user at time t

is a function of its current rate and its average throughput till time t. The proportional fair algorithm does not have delay control mechanism, which makes it unsuitable for delay-constrained applications.

Shakkottai and Stolyar [33] proposed an exponential scheduling algorithm that guarantees some delay constraints. Their proposal is based on the idea of proportional fairness but with different metric. The priority metric for any user at time t is a function of its maximum allowable packet drop probability, its initial life time of each packet, and its head of line delay.

Shin and Lee [34] proposed a modification of the exponential fair algorithm where they identified a new fairness term which reflects the overall buffer status of the i^{th} user at time t. The new priority metric for the modified exponential algorithm is a function of the weighted sum of the number of packets in the i^{th} user's buffer at time t.

Almajano and Romero [36] presented scheduling algorithms that focus on giving QoS guarantee in terms of bit rate in a packet switched CDMA by taking into account the service classes that are defined in the UMTS system, i.e., conversational, streaming, interactive and background. This scheduler was based on the leaky bucket algorithm; they introduced a new parameter called "Service Credit" (SCr) which measures the difference between the bit rate requested by a user and the bit rate that the system has offered.

Lopez *et al.* [38] used energy transmitted per correct received bits as a cost function for evaluating the radio resource consumption of each user within a

UMTS network, and used a modified round robin scheme as their scheduler.

Malik and Zeghlache [39] presented a resource scheduling scheme to improve throughput and fairness for non real time data traffic on the FL WCDMA network. The proposed Fair Resource Scheduling (FRS) scheme takes into account link conditions to enhance throughput and provides fair resource sharing among traffic flows.

Al-Manhari *et al.* [43] proposed a priority packet scheduler algorithm for high speed down-link packet access (HSDPA). Users priorities are set based on their instantaneous channel conditions and their average throughputs.

Long *et al.* [44] found, through mathematical analysis, that TDM is more efficient than CDM for FL in high speed CDMA networks. Based on this finding, they proposed a new scheme for packet scheduling in high speed CDMA networks, namely Channel States Dependent Fair Service for CDMA (CSDFSC), which try to maximize channel throughput under fairness and transmission power constraints. This scheme assigns different queue weights to various traffic classes.

Wang *et al.* [45] proposed the channel adaptive fair queuing (CAFQ) algorithm for multicode TD-CDMA system. In each time slot, the CAFQ algorithm keeps on allocating codes in a fair manner until the whole bandwidth resource is used up (subject to the interference budget). The CAFQ algorithm guaranteed short-term fairness.

Viswanathan and Sayandev [79] developed a linear programming framework for determining an optimum routing and scheduling of flows that maximizes throughput in a wireless WiMAX mesh network. They also discussed the application of mesh networking for load balancing of wired backhaul traffic under unequal access traffic conditions. Numerical results show a significant benefit for mesh networking under unbalanced loading.

Cao *et al.* [80] introduced a new fairness criteria based on the actual traffic demands and formulated a scheduling technique with the objective of maximizing the system throughput under the proposed fairness model in WiMAX networks. The scheduling algorithm is evaluated through simulations.

Cao *et al.* [81] developed a stochastic model for the distributed scheduler of the IEEE 802.16 mesh mode and analyzed the scheduler performance under various conditions.

A cross-layer packet scheduling scheme for video transmission over wireless downlink is presented in [73], a gradient based scheduling scheme is used in which user data rates are dynamically adjusted based on channel quality as well as gradients of a utility function. In addition user utilities are designed as a function of the distortion of the received video.

Chapter 2

Resource Management Techniques for cdma2000 1xEV Star Networks

In order to improve the BW efficiency, the 3GPP2 utilizes efficient mechanisms such as H-ARQ, adaptive modulation, and coding as mentioned in Section 1.3.1. While current standards fully specify several parameters, such as modulation schemes, encoder packet size, and coding techniques, resource management techniques remain unexplored.

In this chapter, and based on the developed rise over thermal (RoT) model, we derive an upper bound for the reverse packet data channel throughput as a function of the number of mobile stations that are allowed to transmit instantaneously on each time slot. We also provide a lower bound for the average sector throughput based on the number of users per sector. We propose several autonomous rate assignment, and scheduling techniques that provide a significant throughput improvement relative to the other published schemes.

The rest of the chapter is organized as follows. The rise over thermal model used throughout the thesis is developed in Section 2.1. In Section 2.2, the upper and lower bounds for the maximum instantaneous throughput over the R-PDCH is developed as a function of the number of active mobile stations per time slot.

The proposed rate assignment techniques are described in Section 2.3. The proposed scheduling techniques are described in Section 2.4. The analysis and simulation results are presented in Section 2.5. A summary of the chapter is given in Section 2.6.

2.1 RoT Constraint

The RoT [16] is defined as the ratio between the total power received to the thermal noise power at the base station. To ensure QoS, 3GPP2 [56] specifies that the RoT per slot should not exceed 7 dB more than 1% of the time.

In what follows we explain the model for the RoT to be used in our work. Let P_T, and P_N be the total received power and the noise thermal power at the BS respectively. Then the RoT is given by $RoT = P_T/P_N$. Assume M active mobiles in the sector under consideration. Therefore, $P_T = \sum_{i=1}^{M} P_i + P_N$, where P_i is the received power of the i^{th} mobile at the BS. Thus $RoT = P_T/(P_T - \sum_{i=1}^{M} P_i)$, which can be expressed as

$$RoT = (1 - \sum_{i=1}^{M} \frac{P_i}{P_T})^{-1}. \tag{2.1}$$

Let $\overline{P_i}$ be the pilot channel power for the i^{th} mobile. Define the traffic to pilot ratio for the i^{th} mobile as $TPR_i = P_i/\overline{P_i}$. Denote the targeted signal to noise and interference ratio for the pilot channel as $\tau_i = \overline{P_i}/(P_T - P_i)$, where $P_T - P_i$ is the total interference and the thermal power for the i^{th} mobile. By substituting into equation 2.1, we get

$$RoT = (1 - \sum_{i=1}^{M} \frac{TPR_i}{TPR_i + 1/\tau_i})^{-1}, \tag{2.2}$$

which can be rewritten as,

$$RoT = (1 - \sum_{k=1}^{r} N_k \frac{TPR_k}{TPR_k + 1/\tau})^{-1}, \tag{2.3}$$

where r denotes the number of allowable rates, N_k denotes the number of active mobiles transmitting with rate R_k at the current time slot, and $\tau = \tau_i = 10^{0.1SINR}$, where $SINR$ is the targeted signal to noise plus interference ratio for the pilot channel (constant for all users). The total traffic to pilot ratio corresponding to rate R_k, TPR_k, is determined by

$$TPR_k = \begin{cases} 1 + 10^{0.1TPRD_k} + 10^{0.1TPDRC}, & \text{1xEV-DO} \\ 1 + 10^{0.1TPRD_k} + 10^{0.1TPRC} + 10^{0.1TPRQ}, & \text{1xEV-DV} \end{cases} \tag{2.4}$$

where $TPRD_k$ denotes the traffic to pilot ratio (in dB) on the reverse data channel associated with rate R_k (Tables 1.1 and 1.2). $TPDRC, TPRC$, and $TPRQ$ denote the traffic to pilot ratio (in dB) associated with the data rate control (DRC) channel (1xEV-DO), reverse control channel and reverse request channel (1xEV-DV) respectively. As specified by the standard, the values of $TPDRC, TPRC, TPRQ$, and $SINR$ are shown in Table 2.1.

To satisfy the RoT constraint, RoT must not be greater than 7 dB. Substituting

Parameter	Value
$TPDRC$	-5 dB
$TPRC$	-5 dB
$TPRQ$	-1 dB
$SINR$	-22 dB

Table 2.1: Values used for the simulation parameters

in equation 2.3,

$$(1 - \sum_{k=1}^{r} N_k \frac{TPR_k}{TPR_k + 1/\tau})^{-1} \leq 10^{0.7} \tag{2.5}$$

Let

$$c_k = \frac{TPR_k}{TPR_k + 1/\tau} \tag{2.6}$$

be a constant for each rate R_k. Therefore Eq. 2.5 can be re-written as

$$\begin{aligned} &\tfrac{1}{(1-\sum_{k=1}^{r} c_k N_k)} \leq 10^{0.7} \\ &\Rightarrow 1 \leq 10^{0.7}(1 - \textstyle\sum_{k=1}^{r} c_k N_k) \\ &\Rightarrow 10^{0.7} \textstyle\sum_{k=1}^{r} c_k N_k \leq 10^{0.7} - 1 \\ &\Rightarrow \textstyle\sum_{k=1}^{r} c_k N_k \leq \tfrac{10^{0.7}-1}{10^{0.7}} = 0.8005 \end{aligned} \tag{2.7}$$

2.2 Theoretical Throughput Bounds

In this section, we develop bounds for the instantaneous throughput on the reverse packet data channel in cdma2000 1xEV as a function of the number of active MSs per time slot and subject to the rise over thermal constraint [56].

2.2.1 Upper Bound

On the reverse packet data channel, each encoded packet is divided into three sub-packets using incremental redundancy. Only one sub-packet is transmitted at a time. The BS uses a soft combining technique to decode the received packet (Section 1.3.1.2). In developing our upper bound, we assume that all the transmitted data are accepted after the first transmission.

In the following analysis we consider the two possible cases:

- Case 1: All mobile stations must be assigned a non zero rate from the set of data rates **R** (Table 1.2) on each time slot.
- Case 2: Some mobile stations are allowed to be idle (i.e., not transmitting) on the given time slot.

Our objective is to maximize the instantaneous link throughput while satisfying the RoT constraint. Hence, finding an upper bound on the instantaneous throughput is equivalent to solving the following integer optimization problem:

Maximize $Throughput = \sum_{k=1}^{r} N_k R_k$

subject to

$C1.\ \sum_{k=1}^{r} N_k = M$ (for case 1)

$\grave{C}1.\ \sum_{k=1}^{r} N_k \leq M$ (for case 2)

$C2.\ \sum_{k=1}^{r} c_k N_k \leq 0.8005$

where the condition $C2$ is obtained from Eq. 2.7. The above optimization problem is solved using the branch and bound optimization technique.

2.2.2 Lower Bound

In bounding the minimum achievable throughput we consider the worst case in which, each mobile station is assigned the lowest rate R_1 and each packet will be accepted after the third sub-packet transmission. The trivial case of all the users being idle is not considered. The minimum bound could be computed as follows,

$$Throughput = M\frac{R_1}{3}, \tag{2.8}$$

and it should satisfy,

$$c_1 M \leq 0.8005, \tag{2.9}$$

where $c_1 = \frac{TPR_1}{TPR_1+1/\tau}$ is constant for rate R_1, M the number of active users, and $\frac{R_1}{3}$ is the effective rate after the retransmissions.

We apply the above optimization problem on cdma2000 1xEV-DV standards where $r = 11$ and the values for $TPRD_k$ are given in Table 1.2. Values of $TPRC$, $TPRQ$, and $SINR$ are shown in Table 2.1.

Fig. 2.1 shows the maximum theoretical link throughput for case 1 and case 2 respectively. Also the lower link throughput that could be achieved is demonstrated in the same figure. The developed lower and upper bounds are used to check our proposed resource management techniques.

From Fig. 2.1, it is clear that the maximum throughput in the first case is reached when three MSs are allowed to transmit. As the number of active users gets higher, we get a lower throughput. The second case confirms the same conclusion, i.e., if the system is allowed to choose the number of active users to transmit on each slot, only three users are chosen to transmit on each time slot.

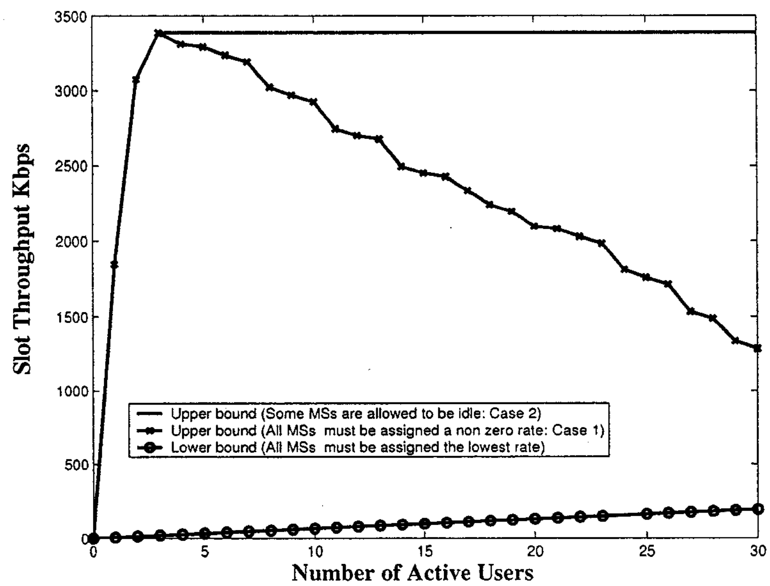


Figure 2.1: Upper and lower throughput bounds for cdma2000 1xEV-DV

Let $(N_1, N_2, \cdots, N_{10}, N_{11})$ denote the rate distribution that achieves the maximum instantaneous throughput. Table 2.2 shows this distribution as a function of the number of active mobile stations, M, for case 1. For case 2, the optimum rate distribution is given by

$$(0,0,0,0,0,0,0,0,0,0,1), \quad M = 1$$

$$(0,0,0,0,0,0,0,0,0,2,0), \quad M = 2$$

$$(0,0,0,0,1,0,0,0,0,2,0), \quad M > 2.$$

Thus, for both cases, the maximum throughput is reached with 2 users assigned the rate R_{10} and one user assigned the rate R_5. One should note that these values of the throughput may not be achieved in practice since we have some timing constraints on the H-ARQ channel. Also, rates should only be assigned if there is enough data in the buffer to form at least one packet with such rate, and packets may not be accepted from the first transmission.

From the above results, it is clear that we can achieve a better throughput if we allow only a subset of active users to transmit on each slot. One can also argue that the rate R_{11} should be used only if we have one active MS, otherwise it is better to use a combination of the other rates.

2.3 Proposed Rate Assignment Techniques

In this section, we present our proposed autonomous rate assignment schemes. For autonomous rate assignment, each MS is free to transmit using the autonomous rate R_k assigned to it by the base station or any other rate below it (Sec. 1.3.1). In order to simplify our analysis, we assume that if the MS is assigned a rate R_k, then it will always transmit using R_k and not a lower rate.

M	$(N_1, N_2, N_3, N_4, N_5, N_6, N_7, N_8, N_9, N_{10}, N_{11})$
1	$(0,0,0,0,0,0,0,0,0,0,1)$
2	$(0,0,0,0,0,0,0,0,0,2,0)$
3	$(0,0,0,0,1,0,0,0,0,2,0)$
4	$(0,0,1,1,0,0,0,0,0,2,0)$
5	$(1,1,0,1,0,0,0,0,0,2,0)$
6	$(0,4,0,0,0,0,0,0,0,2,0)$
7	$(4,1,0,0,0,0,0,0,0,2,0)$
8	$(5,0,0,1,0,0,0,0,1,1,0)$
9	$(4,3,0,0,0,0,0,0,1,1,0)$
10	$(8,0,0,0,0,0,0,0,1,1,0)$
11	$(4,5,0,0,0,0,0,1,0,1,0)$
12	$(8,2,0,0,0,0,0,1,0,1,0)$
13	$(11,0,0,0,0,0,0,1,0,1,0)$
14	$(7,5,0,0,0,0,1,0,0,1,0)$
15	$(11,2,0,0,0,0,1,0,0,1,0)$
16	$(14,0,0,0,0,0,1,0,0,1,0)$
17	$(13,2,0,0,0,1,0,0,0,1,0)$
18	$(12,4,0,0,1,0,0,0,0,1,0)$
19	$(16,1,0,0,1,0,0,0,0,1,0)$
20	$(10,9,0,0,0,0,0,0,0,1,0)$
21	$(18,1,0,1,0,0,0,0,0,1,0)$
22	$(17,4,0,0,0,0,0,0,0,1,0)$
23	$(21,1,0,0,0,0,0,0,0,1,0)$
24	$(22,0,0,1,0,0,0,0,1,0,0)$
25	$(21,3,0,0,0,0,0,0,1,0,0)$
26	$(25,0,0,0,0,0,0,0,1,0,0)$
27	$(21,5,0,0,0,0,0,1,0,0,0)$
28	$(25,2,0,0,0,0,0,1,0,0,0)$
29	$(25,2,0,1,0,0,1,0,0,0,0)$
30	$(24,5,0,0,0,0,1,0,0,0,0)$

Table 2.2: Rate distribution for case 1

For the trivial case of one active MS, the MS' assigned rate is R_{11}. In what follows, we assume that we have more than one active MS.

For our schemes, the rate assigned to the MS is determined based on the user's loaded file size.

Denote the loaded file size of the i^{th} user by F_i. Let T_{k-1} and T_k be the minimum and maximum thresholds for rate R_k, then the i^{th} user will be assigned rate R_k if

$$T_{k-1} \leq F_i < T_k, \tag{2.10}$$

where $T_k \in \mathbf{T}$ and $\mathbf{T}$ denote the set of predetermined thresholds. Based on the results of the previous section, we investigate the following two cases:

- Case A: $\mathbf{R} = \{R_k, 1 \leq k \leq 11\}$
- Case B: $\mathbf{R} = \{R_k, 1 \leq k \leq 10\}$

In the rest of this section, we propose two rate assignment schemes to determine the set of thresholds $\mathbf{T}$. Motivated by fairness amongst users, the first technique adapts to the file length, i.e file proportional rate assignment (FPR). In this technique fairness between users is provided in the sense of service time. The second technique is motivated by minimizing the fluctuation of the system load over time. In this technique each allowable rate is assigned equal number of times over a long interval of time, i.e equiprobable rates assignment (EPR).

2.3.1 FPR Rate Assignment Technique

In this scheme, the set $\mathbf{T}$ is determined such that each MS is assigned a rate proportional to its file length and thus the differences between the number of packets in each MS's buffer is minimized. Hence, the number of time slots required to transmit any data file (not including the effect of time scheduling) is the almost the same for all MSs. Based on the traffic model, a minimum (F_{min}) and maximum (F_{max}) file size can be assumed. For case A, we have

$$T_k = \begin{cases} F_{min}, & k = 0 \\ F_{max}, & k = 11 \\ (F_{max} - F_{min}) \times \frac{R_k}{R_{11}}, & 1 \leq k \leq 10. \end{cases}$$

Similarly, for case B, we have

$$T_k = \begin{cases} F_{min}, & k = 0 \\ F_{max}, & k = 10 \\ (F_{max} - F_{min}) \times \frac{R_k}{R_{10}}, & 1 \leq k \leq 9. \end{cases}$$

2.3.2 EPR Rate Assignment Technique

In this scheme, the MSs are assigned rates in such a way that, over a long interval of time, each allowable rate is assigned equal number of times. This can be achieved by analyzing the pdf distribution of the file size curve. The set $\mathbf{T}$ is determined by dividing the area under the pdf curve into (eleven for case A and ten for case B) equal parts. Let $f(x)$ denote the pdf of the file size, then the set $\mathbf{T}$ is determined such that we have

$$\int_{T_{k-1}}^{T_k} f(x)dx = \int_{T_k}^{T_{k+1}} f(x)dx.$$

The above two schemes were applied to the FTP traffic, the FTP file length distribution is modelled as a truncated lognormal distribution with file size between $0.5K$ bytes and $500K$ bytes, which is obtained from the following lognormal distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp \frac{-(\ln x - \mu)^2}{2\sigma^2}, x \geq 0 \tag{2.11}$$

where $\sigma = 2.0899, \mu = 0.9385$ [56].

Table 2.3 shows the proposed set of file size thresholds $\mathbf{T}$ for FPR and EPR rate assignment schemes (case A).

Similarly, Table 2.4 show the proposed set of file size thresholds $\mathbf{T}$ for for FPR and EPR rate assignment schemes (case B).

Since each data packet is transmitted in 10 msec slot (Sec. 1.3.1), the number of bits required to construct a packet for rate R_k is equal to $R_k \times (10 \times 10^{-3})$ bits. Thus in both the schemes above, to ensure that there is enough data bits to construct at least one packet with data rate R_k, the threshold T_{k-1} should also satisfy

$$T_{k-1} \geq R_k/100. \tag{2.12}$$

It is easy to verify that the threshold sets for both FPR and EPR techniques satisfy the above inequality.

	FPR Scheme	EPR Scheme
T_0	0.5	0.5
T_1	5.19	0.78
T_2	11.04	1.2
T_3	21.43	1.8
T_4	42.22	2.6
T_5	83.79	3.8
T_6	125.362	5.5
T_7	166.9	8.1
T_8	250.07	12.4
T_9	333.21	22.1
T_{10}	416.35	50
T_{11}	500	500

Table 2.3: Proposed file size thresholds in Kbyte for FTP traffic, FPR and EPR schemes (Case A)

	FPR Scheme	EPR Scheme
T_0	0.5	0.5
T_1	6.23	0.8
T_2	13.24	1.3
T_3	25.69	2.0
T_4	50.6	2.9
T_5	100.45	4.5
T_6	150.3	7
T_7	200.26	11.1
T_8	300.01	20.0
T_9	399.49	45.8
T_{10}	500	500

Table 2.4: Proposed file size thresholds in Kbyte for FTP traffic, FPR and EPR schemes(Case B)

2.4 Proposed Scheduling Techniques

Scheduling algorithms are one of the most important aspects to guarantee QoS parameters such as delay and throughput. As mentioned before, the standards left the door open for developing efficient scheduling algorithms.

Many scheduling techniques were proposed for wireless links over the last few years [32] [33] [34]. While the cdma2000 1xEV-DV resource scheduling techniques are only discussed in [14] [30], scheduling on the R-PDCH is barely touched.

In here, we investigate scheduling techniques for minimizing the mobile stations packet delays, and maximizing the link throughput. In these schemes, the base station uses a greedy optimization algorithm to dynamically determine the optimum subset of users that are allowed to transmit on each time slot. Several constraints, such as the rise over thermal, frame error rate, fairness criterion, and the constraints imposed by H-ARQ on the minimum allowed delay between the user sub-packets, must be satisfied.

We will also study the effect of reducing the set of available rates to ten (Case B) on the scheduling process.

2.4.1 Scheduling Process: Base Station Model

In here, we assume that the scheduling process is done over a prespecified time window (TW). Although, theoretically, no maximum time applies, in order to minimize the new users' delay, we assume that the scheduling window is one time slot. On each slot, more than one user is allowed to transmit. The subset of users allowed to transmit will be chosen by the BS based on some optimization criterion using an efficient optimization technique. A slot will be allocated to the MS based on its current status reported on the R-PDCCH.

The optimization algorithm of the scheduling scheme can be based on different cost functions. The following schemes are considered:

Congested-Buffer-First: Minimize the number of packets accumulated in all active users' buffers. Hence, fairness is provided between MSs in the sense of service times.

Lowest-Rate-First: Maximize the number of users transmitting on each slot. In order to maximize the throughput.

2.4.2 Scheduling Process: Mobile Station Model

A slot will be allocated to each MS based on its current status reported on the R-PDCCH. When the MS finishes transmitting, it leaves the system immediately. The MS decides which sub-packet to send on the allocated slot based on the H-ARQ constraints discussed earlier.

Let $s = S(i, l, j)$ be the number of the slot allocated for the i^{th} user, where l $(1 \leq l \leq B_i)$ and j $(1 \leq j \leq 3)$ are the packet and sub-packet numbers the user decides to send on the allocated slot respectively. The H-ARQ operation Fig. 1.4 imposes the constraint that all sub-packets of the same packet have to be spaced by four slots (40 ms), i.e., we must have $S(i,l,j+1) - S(i,l,j) \geq 4$. The same user can not send more than one sub-packet per slot, i.e., $S(i, l+x, j+y) \neq S(i,l,j)$ for $x \neq 0$, $y \neq 0$.

2.4.3 Formulated Problem

The following notation will be used throughout this section:

- M: the number of active mobiles in the current sector
- $A(s)$: the number of users transmitting per slot
- $B_i(s)$: the number of packets in i^{th} user buffer at slot number s
- N_k: the number of active mobiles with data rate R_k transmitting at the current slot.
- RoT: the rise over thermal defined in equation (2.3)
- γ: the RoT threshold required in order to maintain a prespecified QoS
- $S(i,l,j)$: the slot allocated for the i^{th} user, where l $(1 \leq l \leq B_i)$ and j $(1 \leq j \leq 3)$ are the packet and sub packet numbers the MS uses to send on this slot.

Our proposed scheduling algorithm is equivalent to solving the following optimization problem on each slot:

Congested-Buffer-First: $\min \max B_i(s)$

Lowest-Rate-First: $\max A(s)$

Subject to the following constraints:

C1. $RoT \leq \gamma, \gamma = 7dB$.

C2. $S(i,l,j+1) - S(i,l,j) \geq 4$.

C3. $S(i,l+x,j+y) \neq S(i,l,j)$ for $x \neq 0, y \neq 0$.

C4. $\sum_{k=1}^{11} N_k \leq M$.

The above set of constraints can be re-stated as follows:

C1: The rise over thermal must be less than or equal to $7dB$.

C2: All sub-packets of the same packet have to be spaced by four slots.

C3: The same user can not send more than one sub-packet per slot.

C4: The number of users transmitting on each slot must be less than or equal to the number of active users.

2.5 Analysis and Results

In our simulation, we consider a distributed wireless star network in which each sector performs scheduling for its active set of MSs without prior coordination with other sectors. Throughout our simulations, the MSs are assumed to be stationary, i.e., the effect of different mobility parameters (e.g., speed of MSs, multipath channel parameters), have not been investigated by our simulation models. In particular, in all cases, the MSs are assumed to have enough power to support reliable transmission of its assigned data rate.

Fig. 2.2 shows the simulation flow chart we used to generate the results for our rate assignment and scheduling techniques. The system is assumed to be always loaded with a prespecified number of users, i.e., whenever a MS leaves the system another one will join. As mentioned above, the file length of the FTP traffic is modelled by a truncated lognormal distribution (Eq. 2.11). Each MS is assigned a rate, using rate assignment scheme. All the scheduling techniques satisfy the H-ARQ and RoT constraints. The values of $TPRC$, $TPRQ$, and $SINR$ are given in Table 2.1. The values of $TPRD_k$ are given in Table 1.2. Average throughput is calculated as the number of bits successfully transmitted over the total time of transmission averaged over all users. The average service time is calculated as the total number of time slots the user takes to transmit its file (including the time slots the user is not transmitting on) averaged over all users. The results are divided into three categories, proposed rate assignment techniques, proposed scheduling techniques, and combined results.

As mentioned in Section 1.3.1.2 the 3GPP2 [56] requires that all resource management techniques, for 1xEV standards, to satisfy the fairness criterion. Fairness is defined as providing all mobile stations with a minimal level of throughput and is evaluated by determining the cumulative density function (CDF) of the normalized throughput, with respect to the average user throughput, for all users. In order to satisfy the fairness criterion, the standard specifies that the CDF shall lie to the right of the curve given by the three points in Table 1.3. The first point in the table requires that the normalized user throughput of at least 90% of the users exceeds 0.1. Similarly, the second and third points require that the normalized user throughput of at least 80% and 50% of the users exceeds 0.2 and 0.5 respectively.

Throughout our results, we show that all our proposed techniques satisfy the fairness criterion.

2.5.1 Proposed Rate Assignment Techniques

To test our proposed rate assignment techniques the proportional fair scheduler [56] is used as the scheduling technique, and the subset of users chosen to transmit on each slot satisfies the RoT condition (Eq. 2.5) and H-ARQ constraints. The proposed techniques are compared with the highest rate assignment (HR) technique, which implies that each MS is assigned the highest rate its power and available data allow [56]. As we mentioned above the following two cases are investigated:

- Case A: $\mathbf{R} = \{R_k, 1 \leq k \leq 11\}$
- Case B: $\mathbf{R} = \{R_k, 1 \leq k \leq 10\}$

Case A: Figures 2.3(a), and 2.3(b) show that the proposed schemes outperform the highest rate scheme in terms of achieved throughput and service time. It is clear that the achieved system throughput is still far from the maximum achievable throughput bounds obtained in Section 2.2. From Figure 2.3(c), one can note that both the EPR rate assignment technique and highest rate schemes are

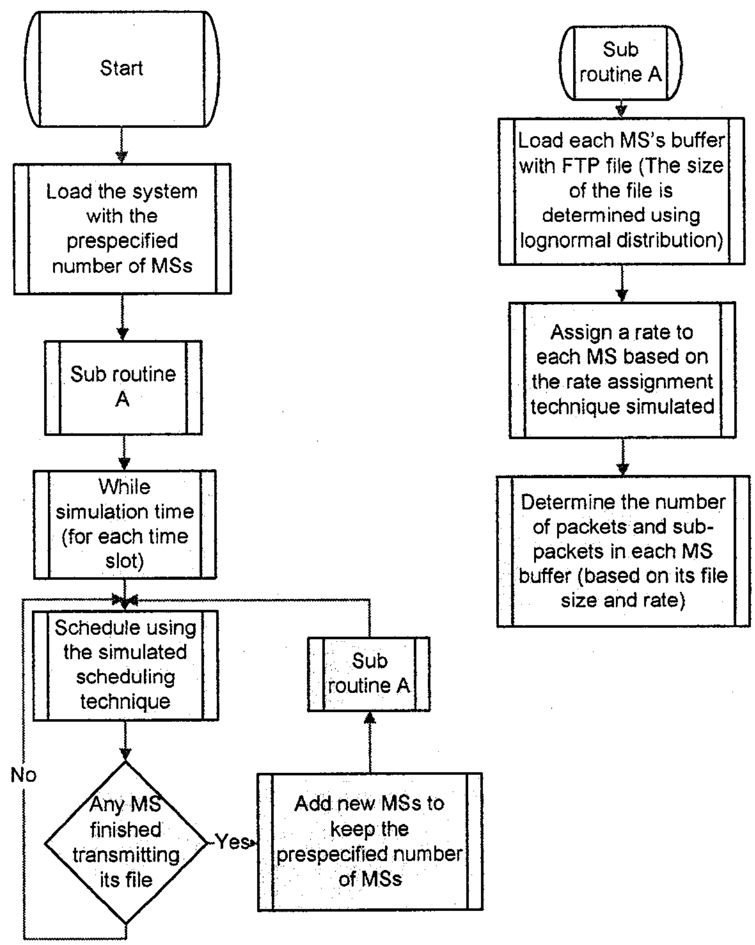


Figure 2.2: Simulation flowchart for the results in Section 2.5

not reaching the 7 dB RoT bound. This means that more improvement can be achieved for both schemes which will be further investigated.

Case B: From Figures 2.3(a), and 2.3(b), it is clear that a significant improvement is achieved when using the EPR rate assignment technique. As for FPR scheme, there is almost no improvement due to the RoT bound. Also it can be recognized that EPR technique outperforms the highest rate technique.

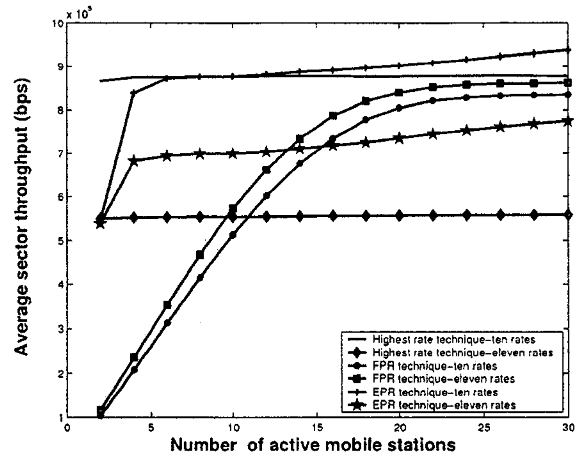
Figs 2.4(a), 2.4(b) show that both of the proposed techniques satisfy the fairness criterion required by the 3GPP2 (Section 1.3.1) as we can see that all the CDF curves are to the right of the fairness criterion line.

2.5.2 Proposed Scheduling Techniques

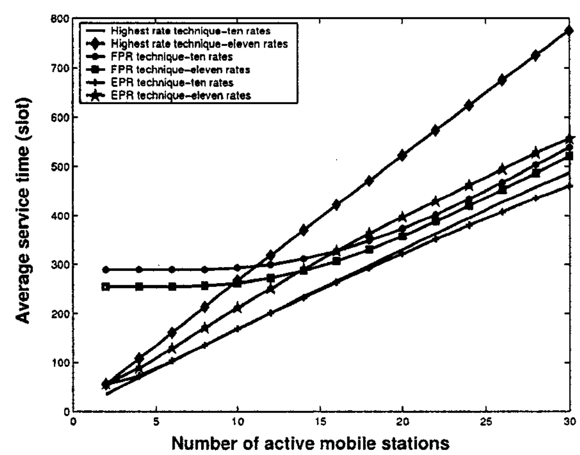
To fairly test our proposed scheduling techniques, highest rate assignment technique is used as the rate assignment technique. As mentioned above the two cases (A and B) are analyzed.

The proposed techniques are compared with the proportional fair scheduling technique and the subset of users chosen to transmit on each slot satisfies the RoT condition (Eq. 2.5) and H-ARQ constraints.

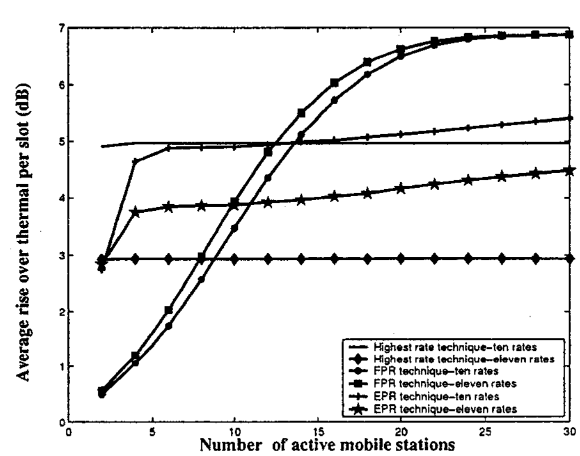
Case A: Fig 2.5(a) shows that the congested-buffer-first scheduling technique elevated the average sector throughput over the proportional fair scheduling algorithm and the lowest-rate-first scheduling technique. From Fig. 2.5(b), one can realize that the average service time for congested-buffer-first scheduling tech-



(a) Average sector throughput

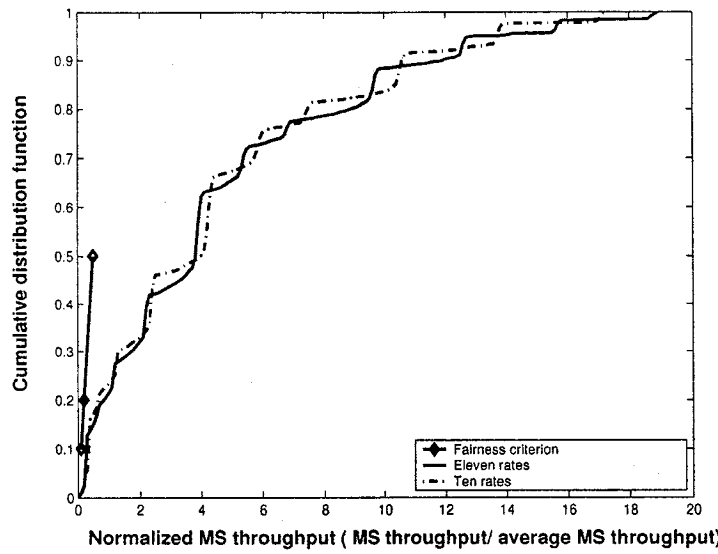


(b) Average service time

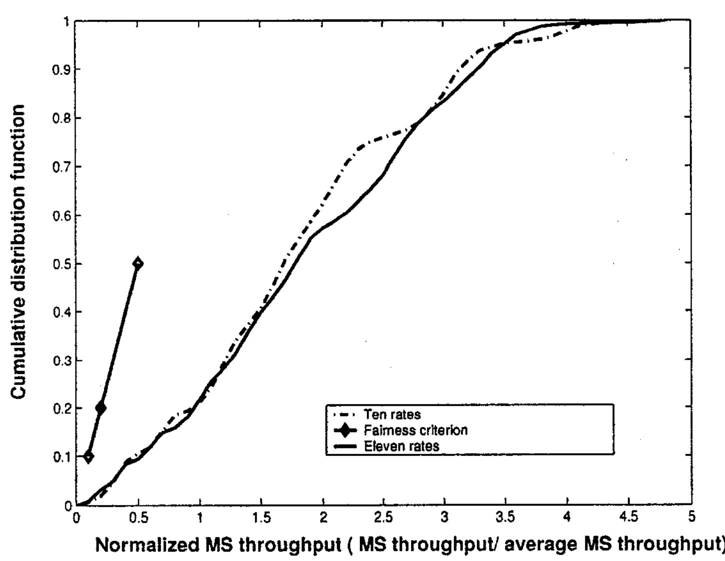


(c) Average RoT

Figure 2.3: Rate assignment techniques, channel BW 1.25 MHz



(a) FPR rate assignment technique



(b) EPR rate assignment technique

Figure 2.4: Cumulative density function (Case A and B) with thirty MSs per sector

nique is better than the other two techniques. From Fig. 2.5(c), it is clear that the RoT did not reach the 7 dB threshold, which leaves the door open for more improvements.

Case B: Fig. 2.5(a) shows that the congested-buffer-first scheduling technique elevated the average sector throughput up to 900 kpbs. Fig. 2.5(b), shows that the average service time for congested-buffer-first scheduling technique has been significantly improved over case A. From Fig. 2.5(c), it is clear that the RoT is still below the 7 dB threshold.

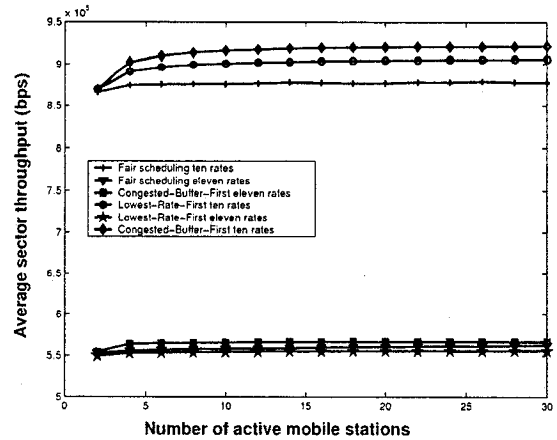
Figs. 2.6(b) and 2.6(a) show that the proposed scheduling techniques satisfy the fairness criterion required by 3GPP2.

The above results support the theory of using the highest rate $R11$ only if we have one active user per sector.

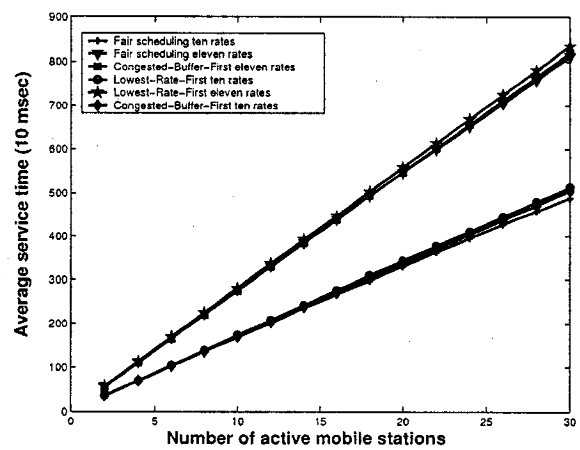
2.5.3 Proposed Scheduling and Rate Assignment Techniques

In this section we tested the affect of changing the rate assignment technique on the different scheduling algorithms. The EPR rate assignment technique (case B) is used as the rate assignment technique. From Fig. 2.7(a) we can see that the highest average throughput ($1050kbps$) is reached using the lowest-rate-first scheduling technique. Fig. 2.7(b) shows that the average service time remains the same. From Fig. 2.7(c), it is clear that the RoT constraint is satisfied.

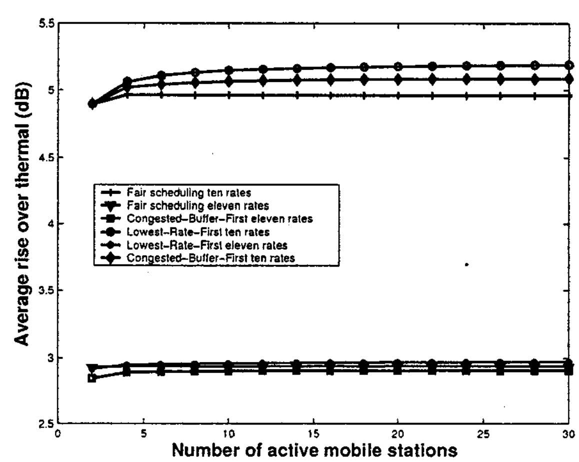
Fig. 2.8 shows that the proposed scheme satisfies the fairness criterion required



(a) Average sector throughput

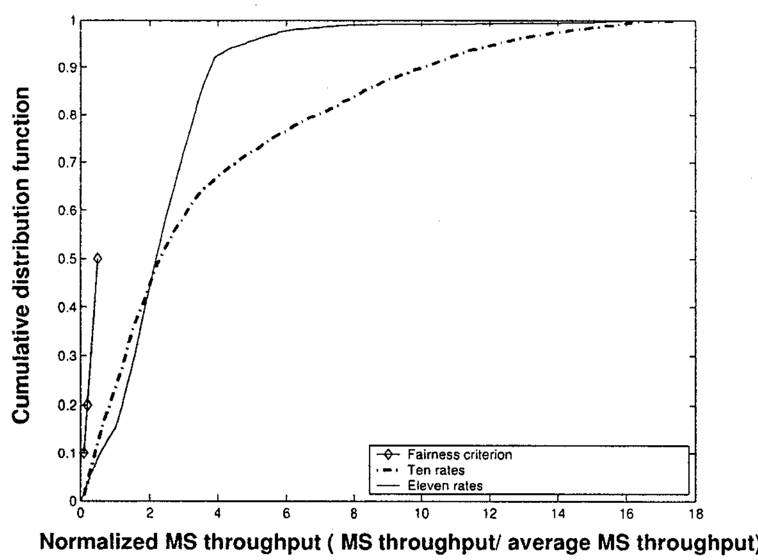


(b) Average service time

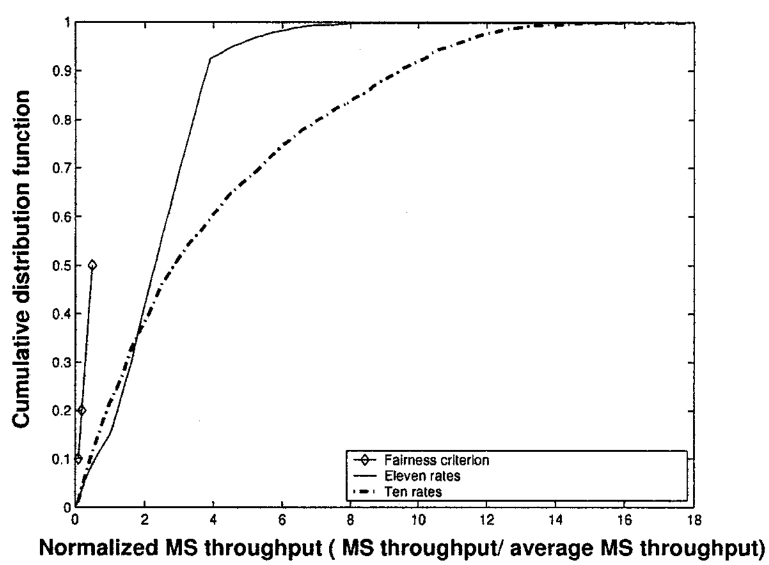


(c) Average RoT

Figure 2.5: Scheduling techniques, channel BW 1.25 MHz



(a) Congested-Buffer-First scheduling technique



(b) Lowest-Rate-First scheduling technique

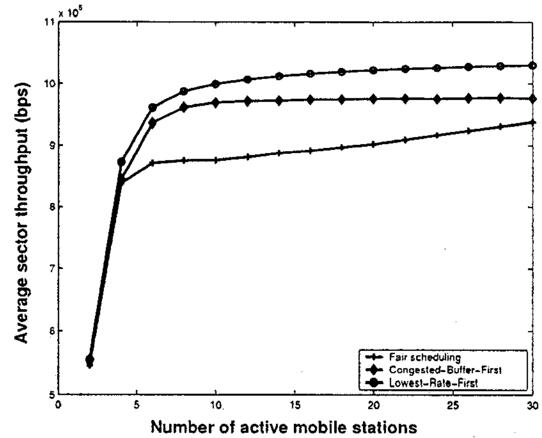
Figure 2.6: Cumulative density function (Case A and B) with thirty MSs per sector

by 3GPP2.

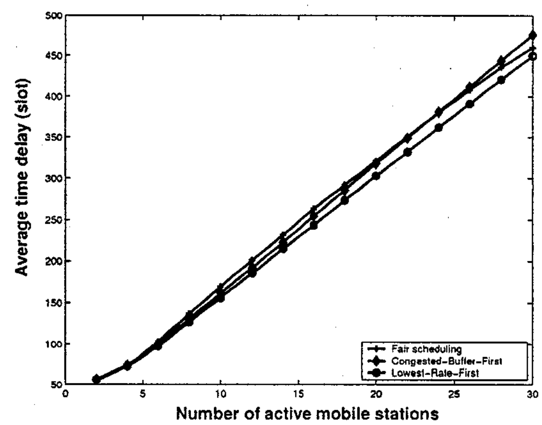
2.6 Summary

In this chapter, the developed RoT model is used to determine a bound on the maximum theoretical throughput that can be achieved over the R-PDCH. Our analysis shows that the highest rate, R_{11}, is adequate only if we have one active user in the sector. It is shown that the maximum instantaneous throughput can be achieved if only a subset of active users are allowed to transmit on each time slot. We proposed two new rate assignment techniques. Our simulation results show that higher throughput and lower service times are achieved using the new proposed techniques. We also proposed two new scheduling techniques, our simulation results show that the proposed algorithm enhances the sector throughput compared to the widely used proportional fair scheduling algorithm. It should be noted that the highest average throughput and lowest service time are achieved with the EPR rate assignment technique (case B) and lowest-rate-first scheduling technique.

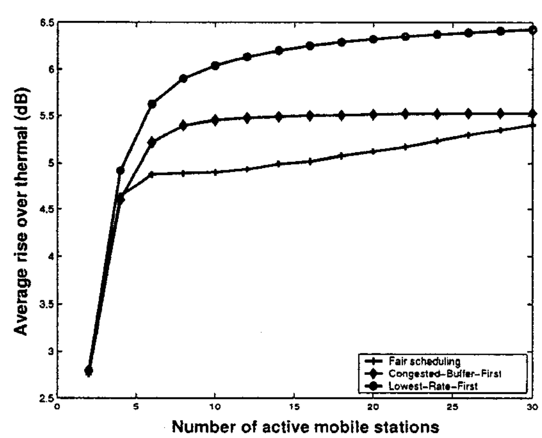
We also showed that the all the proposed techniques satisfy the fairness criterion and the RoT constraint required by the 3GPP2.



(a) Average throughput



(b) Average service time



(c) Average RoT

Figure 2.7: EPR rate assignment technique (Case B) with different scheduling algorithms, channel BW 1.25 MHz

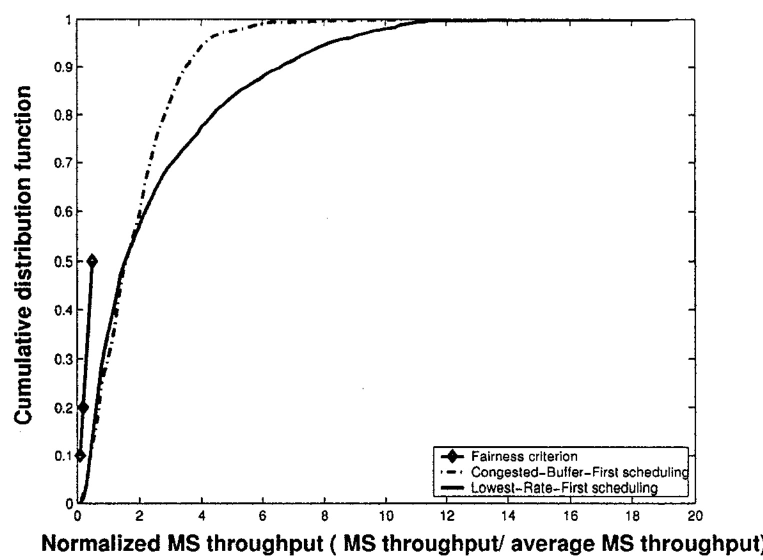


Figure 2.8: EPR rate assignment technique (Case B) (Cumulative density function with thirty MSs per sector)

Chapter 3

Analytical Models for Scheduling Techniques in cdma2000 1xEV Star Networks

Results of the previous chapter show that the second proposed rate assignment technique (Section 2.3) and the second proposed scheduling techniques (lowest-rate-first) (Section 2.4) outperform the other techniques.

In this chapter we compare the performance of the second proposed scheduling technique with other scheduling techniques. In particular, we develop an analytical model for lowest-rate-first, highest-rate-first priority scheduling techniques, and two round-robin fair scheduling techniques over the reverse data channel in cdma2000 1xEV. For these scheduling techniques, the distribution of the mobile stations among the possible data rates is modelled as a Markov process. An analytical expression for the steady state system throughput is derived from the steady state distribution of the above Markov process. The developed model is validated through simulations. Also, the developed Markov model is extended to include the effect of the cross layer design between the H-ARQ protocol, rate assignment, and scheduling and general expressions for the steady state system throughput and file transmission delay are derived.

In our model we assume that each MS keeps its assigned autonomous rate for the whole transmission session. The scheduler chooses a subset of MSs to transmit on each time frame based on its scheduling scheme, while satisfying the rise over thermal (RoT) constraint. Through out this chapter the second proposed rate assignment technique (Section 2.3.2) is used.

The rest of this chapter is organized as follows. The scheduling techniques models are developed in Section 3.1. The scheduling process and the cross layer design Markov models are developed in Sections 3.2 and 3.3 respectively. Finally a summary of the chapter is given in Section 3.4.

3.1 Scheduling Techniques Models

Scheduling algorithms are one of the most important mechanisms to guarantee QoS parameters such as delay and throughput. Many scheduling techniques were proposed for wireless links over the last few years [35] [41]. In this section we model and analyze the performance of the lowest-rate-first, highest-rate-first, and two round-robin scheduling techniques. We assume that all frames are accepted from the first transmission and the system is now loaded with M MSs. The standard [49] guarantees that the BS knows the current number of MSs to be scheduled, which is done by transmitting initial preambles from the MS to BS indicating the imminent arrival of data packets.

The scheduling algorithm allows a subset of MSs to transmit on each time slot, while satisfying the RoT constraint given by Eq. (2.7). Let $\mathrm{x} = (x_1, \cdots, x_r)$

denote the distribution of MSs among different data rates, where r is the number of allowable data rates. Let $a_k(\mathbf{x})$ denote the number of mobile stations assigned a data rate R_k and allowed (by the scheduling technique) to transmit at the t^{th} time slot given x. In what follows, we develop an analytical expression for $a_k(x), k = 1, \cdots, r$ for all the above scheduling techniques.

3.1.1 Lowest-Rate-First

In this scheme, users with an assigned rate R_k are given higher priority to transmit over users with assigned rate $R_{k'}$, $k' > k$. Satisfying the RoT constraint (Eq. 2.7) with the lowest-rate-first scheduling algorithm implies that $a_k(\mathbf{x})$ is given by

$$a_k(\mathbf{x}) = \begin{cases} x_k, & 0 < \sum_{i=1}^{k} c_i x_i \leq 0.8005 \\ \left\lfloor \frac{0.8005 - \sum_{i=1}^{k-1} c_i x_i}{c_k} \right\rfloor, & \sum_{i=1}^{k} c_i x_i > 0.8005, \text{and} \\ & 0 < \sum_{i=1}^{k-1} c_i x_i < 0.8005 \\ 0, & \sum_{i=1}^{k} c_i x_i > 0.8005, \text{and} \\ & \sum_{i=1}^{k-1} c_i x_i \geq 0.8005 \end{cases} \tag{3.1}$$

where c_i, $1 \leq i \leq r$, are given by Eq. (2.6). Taking the floor of the term $\frac{0.8005-\sum_{i=1}^{k-1} c_i x_i}{c_k}$ in the above equation reflects the fact that the number of MSs should always be an integer.

3.1.2 Highest-Rate-First

Similarly, in this scheme, users with an assigned rate R_k are given higher priority to transmit over users with assigned rate $R_{k'}$, $k' < k$. Satisfying the RoT constraints (Eq. 2.7) with the highest-rate-first scheduling algorithm implies that $a_k(\mathbf{x})$ is given by

$$a_k(\mathbf{x}) = \begin{cases} x_k, & 0 < \sum_{i=k}^{r} c_i x_i \leq 0.8005 \\ \left\lfloor \frac{0.8005-\sum_{i=k+1}^{r} c_i x_i}{c_k} \right\rfloor, & \sum_{i=k}^{r} c_i x_i > 0.8005, \text{and} \\ & 0 < \sum_{i=k+1}^{r} c_i x_i < 0.8005 \\ 0, & \sum_{i=k}^{r} c_i x_i > 0.8005, \text{and} \\ & \sum_{i=k+1}^{r} c_i x_i \geq 0.8005 \end{cases} \tag{3.2}$$

3.1.3 Round-Robin Scheme 1

In this scheme, users are divided into r groups according to their data rates. Depending on the distribution of MSs among different rates, for each time slot the scheduler loops $z(\mathbf{x})$ times (the value of $z(\mathbf{x})$ is determined below) through these groups starting from the one with the lowest data rate. Thus each scheduler loop is divided into r steps. During each step, the scheduler allows an additional user with rate R_k to transmit if this does not violate the RoT constraint (Eq. 2.7).

Let $\Delta_{k,i} \in \{0,1\}$ denote the increase, during the i^{th} loop of the scheduler, in the number of users with data rate R_k that are allowed to transmit. Satisfying the RoT constraint with this scheduling scheme implies that $\Delta_{k,i}$ is calculated by the following recursion:

$$\Delta_{k,i} = \begin{cases} 1, & \sum_{j=1}^{r} c_j(\sum_{l=1}^{i-1} \Delta_{j,l}) + \sum_{j=1}^{k} c_j \leq 0.8005,\ i \leq x_k \\ \\ 0, & otherwise \end{cases} \tag{3.3}$$

In the equation above, the term $(\sum_{l=1}^{i-1} \Delta_{j,l})$ denotes the number of rate R_j users that are added during the first $i-1$ scheduler loops.

Recall that in the 1xEV system (Sec. 1.3.1), each user is allowed to transmit at most once during each time slot. Thus, for each time slot, the scheduler keeps on looping until all users are allowed to transmit, i.e., for $\max_{k=1}^{r} x_k$ times, or until the following condition is satisfied during any k^{th} step of the scheduler loop

$$\begin{gathered}\Delta_{k,i} = 0,\ \Delta_{k-1,i} = 0, \cdots,\ \Delta_{1,i} = 0 \\ \text{and } x'_{k,i} \neq 0, 1 \leq k \leq r,\end{gathered} \tag{3.4}$$

where $x'_{k,i}$ denotes the number of users that have not been allowed to transmit yet in the rate R_k group before the i^{th} loop of the scheduler, i.e.,

$$x'_{k,i} = x_k - \sum_{j=1}^{i-1} \Delta_{k,j}. \tag{3.5}$$

Eq. 3.4 means that the scheduler quits looping whenever no additional users can be added to the pool of users that are allowed to transmit due to the RoT constraint. Let i_{min} denote the smallest positive integer i that satisfies the condition in Eq. 3.4. Then $z(\mathbf{x})$ is given by

$$z(\mathbf{x}) = \min\{\max_{k=1}^{r} x_k, i_{min}\}, \tag{3.6}$$

and the number of users with data rate R_k and allowed to transmit is given by

$$a_k(\mathbf{x}) = \sum_{i=1}^{z(\mathbf{x})} \Delta_{k,i} \tag{3.7}$$

3.1.4 Round-Robin Scheme 2

Similarly, in this scheme, users are divided into r groups according to their data rates. However, in this scheme, the scheduler starts looping from the group with the highest data rate. The scheduler keeps on looping until all users are allowed to transmit or until the following condition is satisfied

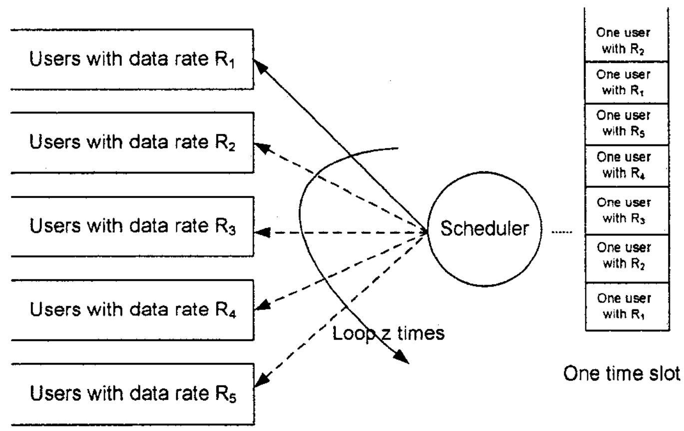


Figure 3.1: Round Robin Scheme 1 Fair Scheduler

$$\begin{gathered}\Delta_{k,i} = 0,\ \Delta_{k-1,i-1} = 0, \cdots,\ \Delta_{1,i-1} = 0 \\ \text{and } x'_{k,i} \neq 0, 1 \leq k \leq r.\end{gathered} \tag{3.8}$$

Eq. 3.8 means that the scheduler quits looping whenever no additional users could be added to the pool of users that are allowed to transmit (due to the RoT constraint).

Let i_{min} denote the smallest positive integer i that satisfies the condition in Eq. 3.8. Let $z(\mathbf{x})$ denote the required number of iterations for each time slot. Satisfying the above constraints implies

$$z(\mathbf{x}) = \min\{\max_{k=1}^{r} x_k, i_{min}\} \tag{3.9}$$

Also, satisfying the RoT constraint with the round-robin scheme 2 scheduling algorithm implies that $a_k(\mathbf{x})$ is given by

$$a_k(\mathbf{x}) = \sum_{i=1}^{z(\mathbf{x})} \Delta_{k,i} \tag{3.10}$$

where $\Delta_{k,i}$ is calculated by the following recursion:

$$\Delta_{k,i} = \begin{cases} 1, & \sum_{j=1}^{r} c_j(\sum_{l=1}^{i-1} \Delta_{j,l}) + \sum_{j=k}^{r} c_j \leq 0.8005,\ i \leq x_k \\ \\ 0, & otherwise \end{cases} \tag{3.11}$$

3.2 Scheduling Process Model

Markov models are some of the most powerful tools available for analyzing communication systems. This analysis yields results for the steady state of the system. A Markov chain is said to be irreducible if it is possible to get to any state from any state. The states of a finite-state, irreducible Markov chains are all recurrent. A state $\mathbf{x}$ is said to have a period d if any return to itself must occur in some multiple of d time steps. An irreducible Markov chain is said to be aperiodic if $d = 1$ [84].

If the Markov model is irreducible, aperiodic and positive recurrent, the state probabilities reach steady state values that are independent of the initial state probabilities. The steady state probabilities $\pi(\mathbf{x})$ for a state $\mathbf{x}$ could be obtained by solving the linear equations

$$\pi(\mathbf{y}) = \sum_{\forall \mathbf{x}} p_{\mathbf{xy}} \pi(\mathbf{x}) \tag{3.12}$$

and

$$\sum_{\forall \mathbf{x}} \pi(\mathbf{x}) = 1 \tag{3.13}$$

Let $S_k(t)$ denote the number of mobile stations assigned a data rate R_k at the t^{th} time slot. We define a state vector $\mathbf{S}(t) = (S_1(t), S_2(t), \cdots S_r(t))$. Note that a MS with an assigned rate R_k may not necessarily be allowed (by the scheduling algorithm) to transmit at a given time slot. The assumption that the system is always loaded with M MSs implies that $\sum_{k=1}^{r} S_k(t) = M$ and hence the number of valid states is given by $\binom{M+r-1}{r-1}$ [84].

Let $F_k(t+1), k = 1, \cdots, r$ denote the number of mobile stations assigned a data rate R_k and leaving the system (after finishing their file transmission) at the end of the t^{th} time slot, i.e., at the beginning of slot $t+1$. Similarly, let $J_k(t+1), k = 1, \cdots, r$ denote the number of MSs that join the system at time slot $(t+1)$ with an assigned rate R_k. Thus

$$\mathbf{S}(t+1) = \mathbf{S}(t) - \mathbf{F}(t+1) + \mathbf{J}(t+1). \tag{3.14}$$

Let $\mathbf{x} = (x_1, \cdots, x_r)$, and $\mathbf{y} = (y_1, \cdots, y_r)$ be the sample of $\mathbf{S}(t)$ and $\mathbf{S}(t+1)$ respectively. Similarly let $\mathbf{f} = (f_1, \cdots, f_r)$, and $\mathbf{j} = (j_1, \cdots, j_r)$ be the sample of $\mathbf{F}(t+1)$ and $\mathbf{J}(t+1)$ respectively.

Then the state transition probability $p_{\mathbf{xy}} = p(\mathbf{S}(t+1) = \mathbf{y} | \mathbf{S}(t) = \mathbf{x})$ is given by

$$p_{\mathbf{xy}} = \sum_{\mathbf{f},\mathbf{j}} p(\mathbf{S}(t+1) = \mathbf{y}|\mathbf{f},\mathbf{j},\mathbf{x}) \times p(\mathbf{J} = \mathbf{j}|\mathbf{f},\mathbf{x}) \times p(\mathbf{F} = \mathbf{f}|\mathbf{x}). \tag{3.15}$$

The conditional probability $p(\mathrm{S}(t+1) = \mathrm{y}|\mathrm{f},\mathrm{j},\mathrm{x})$ is given by

$$p(\mathbf{S}(t+1) = \mathbf{y}|\mathbf{f},\mathbf{j},\mathbf{x}) = \prod_{k=1}^{r} p(y_k = x_k - f_k + j_k) \tag{3.16}$$

where,

$$p(y_k = x_k - f_k + j_k) = \begin{cases} 1, & y_k = x_k - f_k + j_k, \\ 0, & y_k \neq x_k - f_k + j_k. \end{cases} \tag{3.17}$$

Example:

Let the total number of users $M = 2$, number of allowable rates $r = 3$. If $\mathrm{x} = (1,1,0)$, $\mathbf{f} = (0,1,0)$, and $\mathbf{j} = (1,0,0)$, then the new state $\mathbf{y} = (2,0,0)$.

Again, the assumption that the system is always loaded with M MSs implies that $\sum_{k=1}^{r} f_k = \sum_{k=1}^{r} j_k$.

Example:

Let $M = 2$, $r = 3$, $\mathrm{x} = (2,0,0)$, and $\mathrm{y} = (2,0,0)$. Then the possible sets of f and their corresponding sets of j are

$$\{\mathbf{f} = (0,0,0),\ \mathbf{j} = (0,0,0)\}, \{\mathbf{f} = (1,0,0),\ \mathbf{j} = (1,0,0)\}, \{\mathbf{f} = (2,0,0),\ \mathbf{j} = (2,0,0)\}.$$

Using the rate assignment technique described in section 2.3.2 implies that the new mobile stations joining the system at time slot $t+1$ will be distributed uniformly over the possible rates. Then $p(\mathbf{J} = \mathrm{j}|\mathrm{f},\mathrm{x})$ is given by

$$p(\mathbf{J} = \mathbf{j}|\mathbf{f}, \mathbf{x}) = \frac{1}{\binom{\sum_{k=1}^{r} f_k + r - 1}{r - 1}} \tag{3.18}$$

where the denominator in the equation above denotes the number of possible choices for $\mathbf{j}$ given $\mathbf{f}$.

Example:

Let $r = 3$, and $\mathbf{f} = (1, 1, 0)$. Then

$$\mathbf{j} \in \{(2,0,0), (1,1,0), (1,0,1), (0,2,0), (0,1,1), (0,0,2)\}.$$

Let $p(F_k = f_k|\mathbf{x})$ denote the probability that a MS with an assigned rate R_k will finish its transmission at time slot $t+1$. This probability is equal to the probability that the number of packets left at this MS buffer at the beginning of time slot t is equal to one and the scheduling algorithm allows this MS to transmit it.

Let $B_{u,k}$ and L_u denote the number of packets and the file size in bits for the u^{th} MS with rate R_k respectively. Hence

$$B_{u,k} = \frac{L_u}{b_k} \tag{3.19}$$

where b_k is a constant that denotes the number of bits per packet associated with rate R_k. Then $p(f_k|a_k(\mathbf{x}))$ can be approximated by

$$p(f_k|a_k(\mathbf{x})) \approx \begin{cases} 0, & a_k(\mathbf{x}) < f_k \\ \binom{a_k(\mathbf{x})}{f_k} (\frac{1}{B_k})^{f_k} (1 - \frac{1}{B_k})^{(a_k(\mathbf{x}) - f_k)}, & f_k \leq a_k(\mathbf{x}), \end{cases} \tag{3.20}$$

where

$$\overline{B}_k = \tfrac{1}{b_k} E[L_k], \tag{3.21}$$

where $E[\cdot]$ denotes the expected value of the enclosed argument and L_k is determined assuming the rate assignment technique discussed in section 2.3.2. Thus,

$$p(\mathbf{F} = \mathbf{f}|\mathbf{x}) = \prod_{k=1}^{r} p(f_k|a_k(\mathbf{x})) \tag{3.22}$$

where $a_k(\mathrm{x})$ denote the number of mobile stations assigned a data rate R_k and allowed (by the scheduling technique) to transmit at the t^{th} time slot given $\mathrm{S}(t) =$ x as shown in Section 3.1.

By noting that the above Markov model is irreducible, aperiodic, and positive recurrent [85], there will be a unique steady-state probability $\pi(\mathrm{x})$ for a state $\mathbf{x}$. The steady state system throughput is given by

$$\sum_{\forall \mathrm{x}} \sum_{k=1}^{r} a_k(\mathrm{x}) R_k \pi(\mathrm{x}) \tag{3.23}$$

where $\pi(\mathrm{x})$ denotes the steady state probability distribution of S.

3.2.1 Analysis and Simulation Results

The above model is applied to 1xEV-DO system, where $r = 5$. The values of $TPDRC$, and $SINR$ are given in table 2.1. The values of $TPRD_k$ are given in Table 1.1. The above scheduling and rate assignment techniques were applied to the FTP traffic in which the file length distribution is modelled as a truncated lognormal distribution Eq. 2.11.

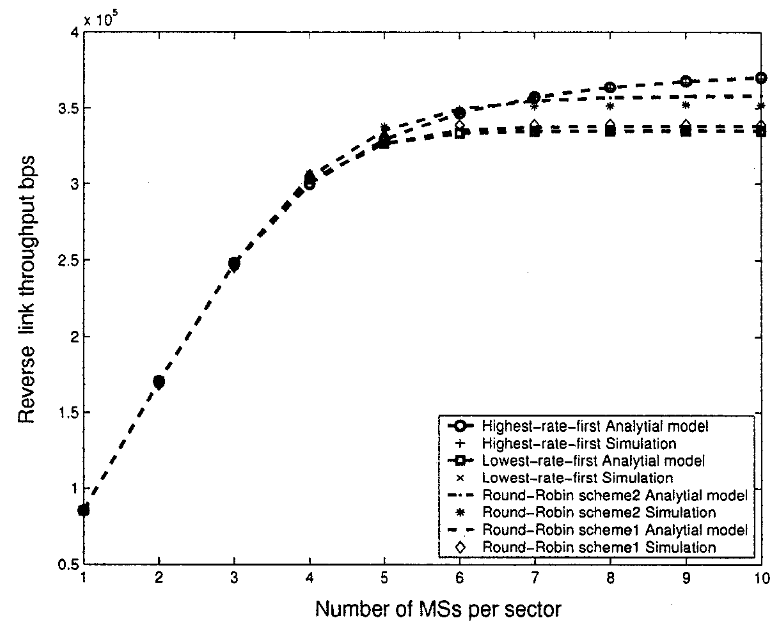
The simulation and analytical model results are shown in Fig. 3.2(a). Both scheduling and rate assignment techniques are adapted in simulation and analysis. One can realize that for $M < 5$ most of the users (with different rate distributions) can transmit in the same time slot while satisfying the RoT constraints, hence the achieved average throughput for $M < 5$ are the same for all four schemes.

In the course of computing the steady state probability vector $\pi(\mathbf{x})$ for Eq. 3.30, we noted that the states with large number of MSs and high data rates (waiting to transmit) dominate in the case of lowest rate first scheduling. This can be explained by noting that this scheduling technique allows a large number of low rate MSs to transmit first, which lowers the probability that MSs with higher rates can satisfy the RoT constraints, consequently they stay longer within the system waiting for service. This results in lowering the overall system throughput as compared to the highest-rate-first scheduling technique.

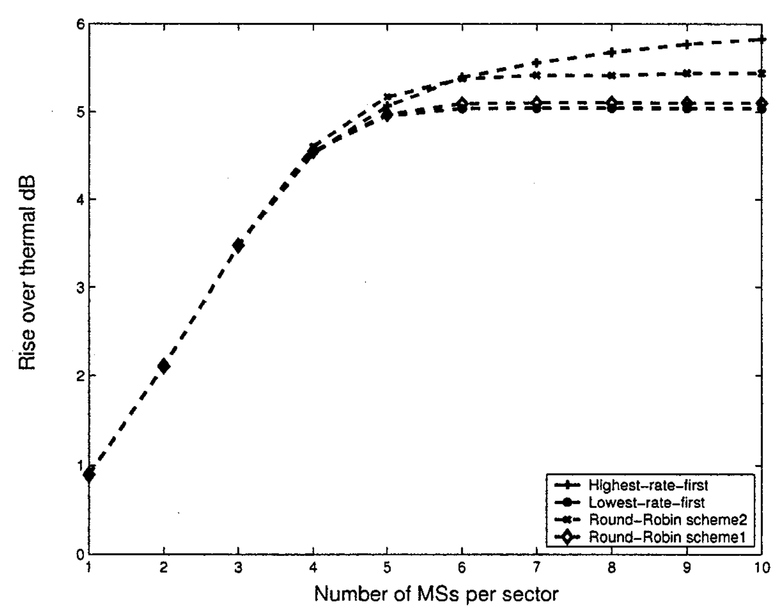
The same explanation also applies to the round-robin scheme 1 scheduling technique. Although it allows a combination of low rate MSs and high rate MSs to transmit, the priority in each round is given to low rate MSs.

The average RoT obtained by simulation for different values of M is shown in Fig. 3.2(b). It is clear that the RoT constraints are satisfied for all four schemes. One should also note that all the considered scheduling techniques do not reach the 7 dB RoT bound.

By evaluating the second largest eigenvalue for the transition matrices $\{p_{\mathbf{xy}}\}$,



(a) Reverse Link throughput



(b) Rise over Thermal

Figure 3.2: Performance Results for cdma2000 1xEV-DO reverse data channel

we have also analyzed the rate of convergence [85] of the Markov chain for all schemes. None of the schemes shows any clear advantage over the other with regard to this criterion.

3.3 Cross Layered Design Model

In this section we extend the model proposed in Section 3.2 to model and evaluate the overall performance of the hybrid ARQ, rate assignment, and time slot scheduling over the reverse packet data channel in cdma2000 1xEV. Our model is based on the analysis of the cdma2000 1xEV turbo encoder (Appendix A) and the probability that a mobile station is scheduled for transmission. Expressions for the steady state system throughput and file transmission delay are derived from the steady state distribution of the Markov process. The developed models are validated through simulations. The obtained results show how the adaptation of scheduling to the status of the physical layer and H-ARQ operation constitutes a cost effective cross layer design.

In order to determine a lower bound for the probabilities of sub-packets acceptance, it is inevitable to examine the details of the turbo decoding process. Let $ps_{1,k}$ and $ps_{2,k}$ denote the probabilities that a packet, transmitted with rate R_k, is accepted from the first or second transmissions respectively. Let b_k denote the number of data bits per packet associated with rate R_k. Thus

$$ps_{i,k} = (1 - (P_b(E))_{i,k})^{b_k} \tag{3.24}$$

where $i = 1, 2, 3$ correspond to the first, second and third sub-packets respectively, and $(P_b(E))_{i,k}$ denotes the BER for the i^{th} sub-packet with data rate R_k. In Appendix A, based on the analysis of the cdma2000 1XEV turbo encoder with code rates $R_{c_i} = 1/2, 1/3, 1/5$, $(P_b(E))_{i,k}$ is determined (Eq. A.9), and hence the values of $ps_{i,k}$ are obtained. We also show how different FEC coding rates R_{c_i} associated with the H-ARQ process affect the BER Fig. A.4.

As mentioned in section 1.3.1, the H-ARQ protocol has a 4 time slot cycle from the time the MS starts transmitting the sub-packet until it receives the ACK/ NAK and processes it. Therefore, while the first sub-packet is transmitted in one time slot, the time elapsed until the second and third sub-packets are transmitted is 5 and 9 time slots respectively (see Fig. 1.4). Therefore, from these delay constraints, the average number of time slots (without the affect of scheduling Fig. 3.3) to transmit a packet is given by

$$ps_{1,k} + 5(1 - ps_{1,k})ps_{2,k} + 9(1 - ps_{1,k})(1 - ps_{2,k}), \tag{3.25}$$

where $ps_{1,k}$ and $ps_{2,k}$ are the probabilities that the packet is accepted from the first or second retransmissions respectively. The probabilities $ps_{1,k}$ and $ps_{2,k}$ are evaluated by Eq. (3.24). When allowed by the scheduling algorithm, the MS can transmit different packets during the 40 msec cycle. Therefore the effective number of

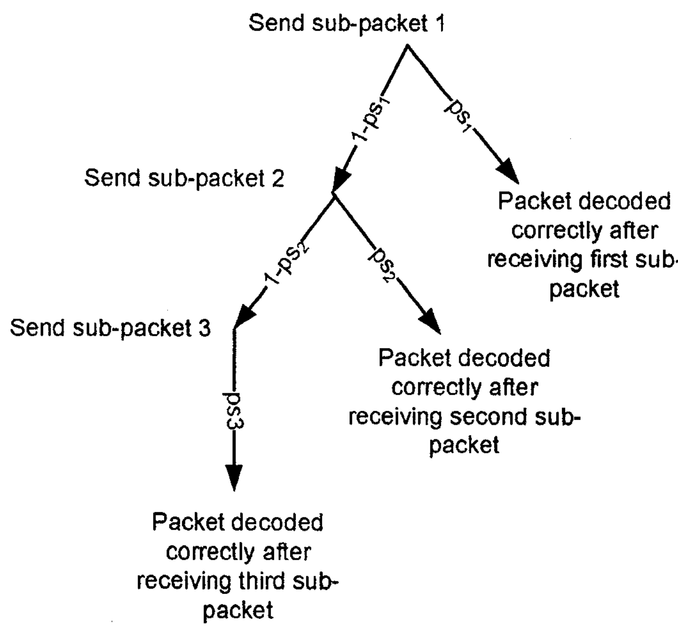


Figure 3.3: H-ARQ Probability Tree

time slots to transmit a file for a MS with data rate R_k including retransmissions due to H-ARQ $\overline{T}_k$ is given by

$$\overline{T}_k = \overline{B}_k(ps_{1,k} + 2(1 - ps_{1,k})ps_{2,k} + 3(1 - ps_{1,k})(1 - ps_{2,k})) \tag{3.26}$$

where $\overline{B}_k$ denotes the number of packets associated with rate R_k by Eq. 3.21.

Thus, the Mrakov model developed in section 3.2 can be extended to include the H-ARQ effect by changing Eq. 3.20. Hence $p(f_k|a_k(\mathbf{x}))$ can be approximated by

$$p(f_k|a_k(\mathbf{x})) \approx \begin{cases} 0, & a_k(\mathbf{x}) < f_k \\ \binom{a_k(\mathbf{x})}{f_k}(\frac{1}{\overline{T}_k})^{f_k}(1 - \frac{1}{\overline{T}_k})^{(a_k(\mathbf{x})-f_k)}, & f_k \leq a_k(\mathbf{x}), \end{cases} \tag{3.27}$$

Given the steady-state probability $\pi(\mathbf{x})$ for a state $\mathbf{x}$, the steady state average number of MS that is assigned data rate R_k is given by

$$\overline{x}_k = \sum_{\forall \mathbf{x}} x_k(\mathbf{x})\pi(\mathbf{x})$$

and the steady state average number of MSs that is assigned data rate R_k and is scheduled for transmission is given by

$$\overline{a}_k = \sum_{\forall \mathbf{x}} a_k(\mathbf{x})\pi(\mathbf{x})$$

Let $\overline{I}_k = \frac{\overline{a}_k}{\overline{x}_k}$ denote the steady state probability that a MS assigned data rate R_k is scheduled for transmission. One can show that the number of time slots needed to transmit the first sub-packet is

$$\begin{aligned}\overline{I}_k + 2(1-\overline{I}_k)\overline{I}_k + 3(1-\overline{I}_k)^2\overline{I}_k + \cdots\cdots &= \sum_{l=0}^{\infty}(1+l)(1-\overline{I}_k)^l\overline{I}_k \\ &= \frac{1}{\overline{I}_k}\end{aligned}$$

Similarly, number of time slots needed to transmit the second and third sub-packets are given by

$$\sum_{l=0}^{\infty}\sum_{i=0}^{\infty}(2+l+i)(1-\overline{I}_k)^{l+i}\overline{I}_k^2 = \frac{2}{\overline{I}_k}$$

and

$$\sum_{l=0}^{\infty}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(3+l+i+j)(1-\overline{I}_k)^{l+i+j}\overline{I}_k^3 = \frac{3}{\overline{I}_k}$$

respectively.

Let D_k denote the number of slots needed to transmit a file by a MS with an assigned rate R_k, including the effect of scheduling and H-ARQ. From Eq. (3.26) and by noting the number of time slots required for each transmission (obtained above), we obtain

$$\overline{D}_k = \tfrac{\overline{B}_k}{\overline{I}_k}(ps_{k,1} + 2(1 - ps_{k,1})ps_{k,2} + 3(1 - ps_{k,1})(1 - ps_{k,2}))$$

Therefore, the steady state average file transfer delay, including both transmission and access, is given by

$$\frac{1}{r}\sum_{k=1}^{r}\overline{D}_k T_s \tag{3.28}$$

where T_s is the length of each time slot, r is the number of allowable rates, and $\overline{B}_k$ is the average number of packets as defined in Eq. 4.27.

Due to the H-ARQ constraints, each packet with data rate R_k needs more than one time slot to be transmitted. Hence, the effective data rate is given by

$$R_{k(eff)} = \frac{R_k}{ps_{1,k} + 2(1 - ps_{1,k})ps_{2,k} + 3(1 - ps_{1,k})(1 - ps_{2,k})}, \tag{3.29}$$

and the steady state system throughput is given by

$$\sum_{k=1}^{r}\overline{a}_k R_{k(eff)} \tag{3.30}$$

3.3.1 Analysis and Simulation Results

We consider a system where the number of allowable rates r is set to seven $(R_1, \cdots, R_7)$. The values of $TPRC$, and $TPRQ$, and $TPRD_k$ are given in Tables 2.1 and 1.2. Also in all the results $W = 1.25\ MHz$ which implies single-carrier cdma2000, the extension of this work to multi-carrier system is a straightforward computation. The probability of sub-packet acceptance for different values of τ are determined based on the values E_b/N_o (Eq. A.22).

The above models were applied to the FTP traffic in which the file length distribution is modelled as a truncated lognormal Eq. 2.11. The system is assumed to be always loaded with a prespecified number of users, i.e., whenever a MS leaves the system another one will join. The rate assignment is performed whenever a new MS joins the system. Each MS keeps its assigned rate for the duration of its transmission. The simulation results presented in this section are obtained by simulating the system, for each point on the curves, over $500,000$ time slots.

The simulation and analytical results for the steady state throughput and file transfer delay of the lowest rate first scheduling technique are shown in Figs. 3.4 and 3.5, respectively, for different values of τ. The corresponding results for the highest rate first scheduling technique are shown in Figs. 3.6 and 3.7. In the case of well-conditioned channel, packets are assumed to be accepted from the first sub-packet transmission, i.e., $ps_{k,1} = 1$. In the case of an ill-conditioned channel, $ps_{k,1} \leq 1$, and the H-ARQ technique is applied.

For well-conditioned channel, it is clear that lower values of τ allow higher throughput and lower file transfer delay. This can be explained by noting that the scheduling technique allows more MSs to transmit on each time slot for lower values of τ. Basically, since the channel is well-conditioned, the throughput will not be affected because there would be no re-transmission. On the other hand, for ill-conditioned channel, lower values of τ result in a large degradation in both the

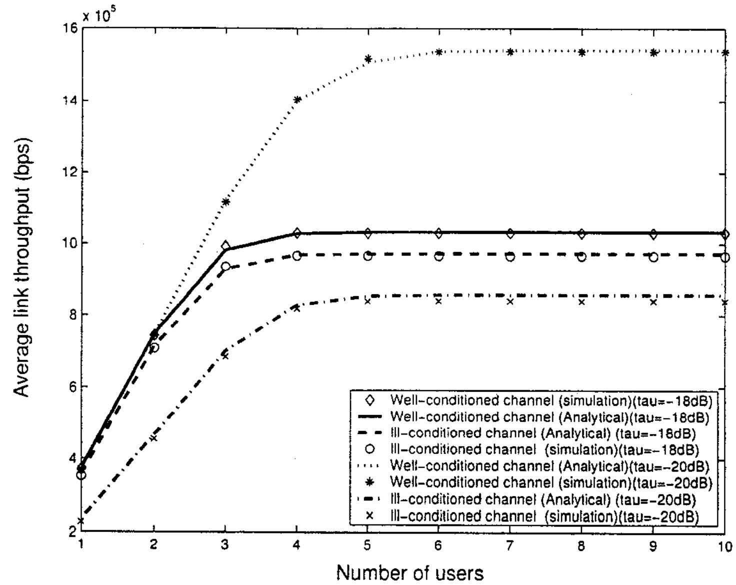


Figure 3.4: Steady state link throughput for lowest-rate-first scheduling technique.

system throughput and file transfer delay which can be explained by the required increase of the sub-packets retransmissions. In other words, higher values of τ help bring up the performance of the ill-conditioned channel to reach that of the well-conditioned channel but it will reduce the overall system throughput in order to satisfy the scheduling constraints.

3.4 Summary

In this chapter, four scheduling techniques for the reverse data channel of the 1xEV star network were modelled. The distribution of the MSs among the possible data rates is modelled as a Markov process. The effects of scheduling and

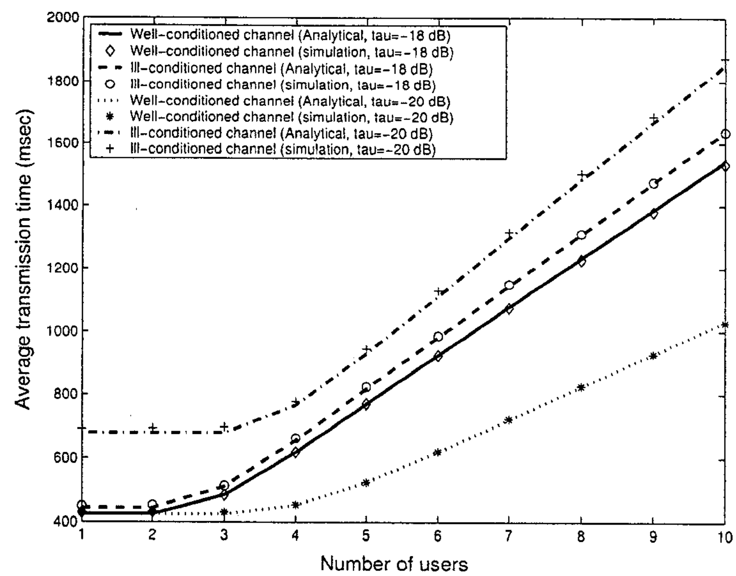


Figure 3.5: Steady state average file transfer time per user for lowest-rate-first scheduling technique.

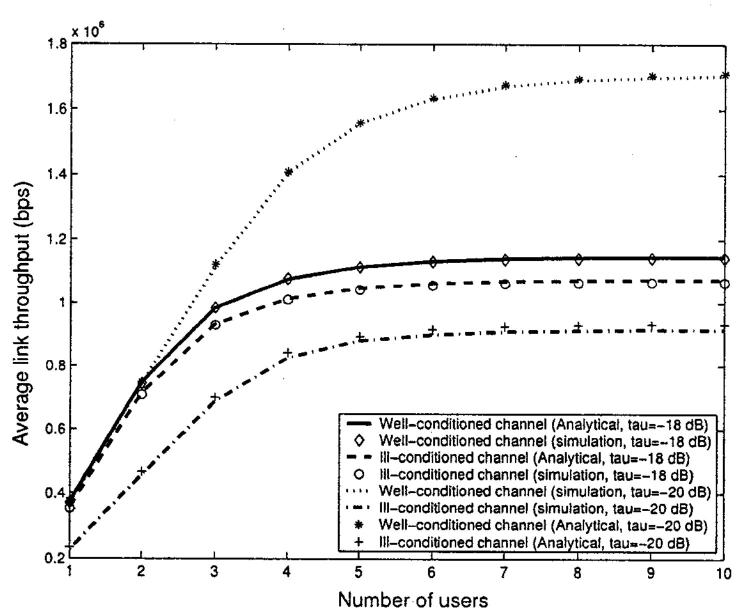


Figure 3.6: Steady state link throughput for highest-rate-first scheduling technique.

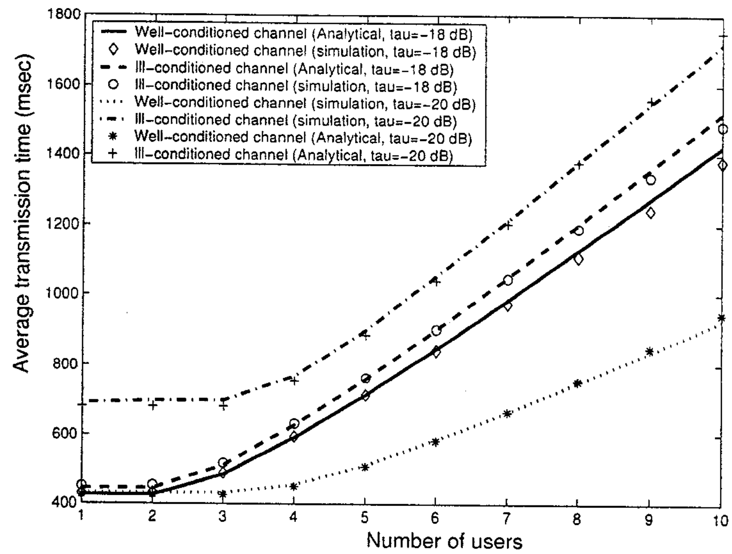


Figure 3.7: Steady state average file transfer time per user for highest-rate-first scheduling technique.

rate assignment techniques are included in the Markov model. The validity of the obtained models was confirmed with simulations. The results show that highest-rate-first scheduling technique outperforms all the other scheduling techniques in the sense of achieved throughput. The simulation results are fairly close to their analysis counterparts, which validates the assumptions of the analysis.

We extended our work to develop analytical expressions for the steady state link throughput and file transfer delay that are based on modelling the cross layer design involving the physical layer H-ARQ, network layer rate assignment, and MAC layer time slot scheduling techniques used over the reverse data channel. The results obtained show how the choice of the targeted signal to noise and interference ratio affect the overall system performance.

Chapter 4

Caching Effects on Cross-layered WiMAX Mesh Networks

The IEEE 802.16 backhaul mesh network consists of a base station (BS) and multiple service stations (SSs). The BS serves as a gateway for the SSs to the Internet, and each SS serves as a base station collecting aggregated traffic from end users in different star networks (see Fig. 4.1). In previous chapters we considered different rate assignment and scheduling techniques for star connected networks, the rate assignment and inherent cross layer design techniques that we adapted for star networks are naturally extended to backhaul mesh oriented networks.

In this chapter we investigate some techniques that could be used to improve the performance of backhaul mesh networks. A common characteristic of such networks is the use of mesh multi-hop networking to improve efficiency. IEEE 802.16 supports two modes of operation, point to multi-point (PMP) and mesh mode. In PMP each SS directly communicates with the BS through a single hop link, which requires all SSs to be within clear line of sight to the BS. In contrast, in mesh mode, the SSs can communicate with mesh BS and with each other through multi hop routes via other SSs. Mesh topology not only extends the network coverage and increases the capacity in non-line of sight environments, it also provides higher network reliability and availability when node or link failures occur,

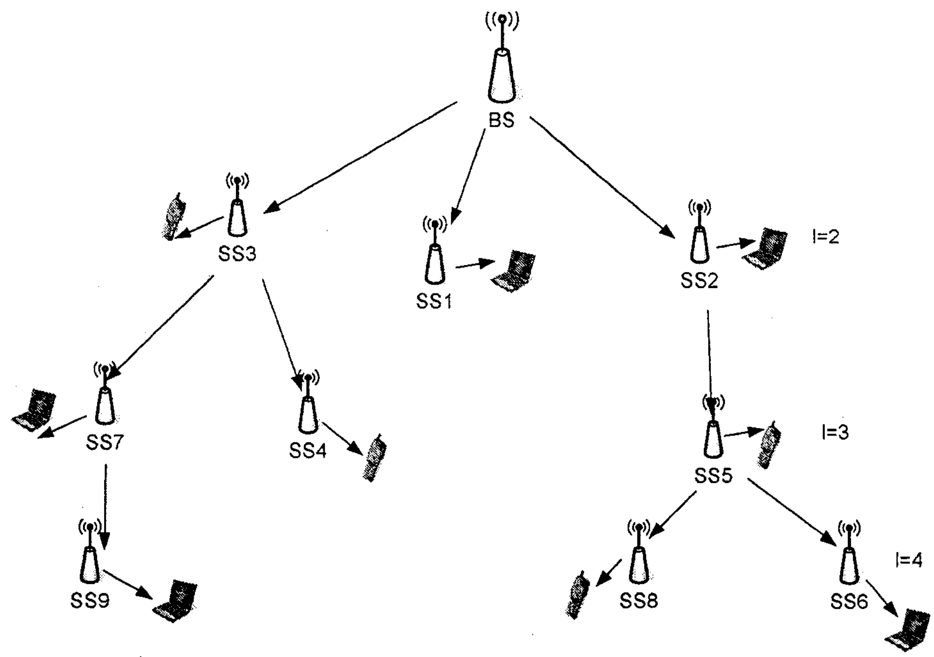


Figure 4.1: An example of backhaul mesh network

or when channel conditions are poor [80].

In the last few years, the importance of cross layer design resulted in many proposals. In [67] the author surveyed the cross layer designs based on how the layers are coupled. The layer coupling is divided into four categories: creation of new interfaces [68] and [69], merging of adjacent layers, design coupling without new interfaces [70], and vertical calibration across layers [71].

In this chapter, we provide cross layer techniques to improve the effects of forward error correction (FEC) coding amongst other measures on the overall network performance. Coding is one of the important measures to improve network performance. An opportunistic coding approach increases the throughput of wireless networks by minimizing the transmission delay. Turbo coding [89] is one of the important coding techniques used on the physical layer. While turbo codes have been employed for some time, recently upper layer FE codes such as LT codes [98], Raptor codes [97], and Fountain codes [96] have been introduced. Fountain codes are a class of erasure codes with the property that a potentially limitless sequence of encoding symbols can be recovered from any subset of size equal to or slightly larger than the number of source symbols.

In this chapter we relate the parameters of the higher layer Fountain codes to those of the physical layer codes to present a cross-layered coding technique. Based on the bit error rate of physical layer codes, the upper layer Fountain codes parameters are designed to obtain a prespecified performance. The performance of the proposed cross-layered coding technique is found to be comparable to that of the physical layer based Hybrid-ARQ technique. As an application, we study the performance of a WiMAX backhaul mesh network where caching is allowed at the service stations and the proposed coding technique is used instead of H-ARQ. In particular, the performance of the above network is modelled as a Markov process and analytical expressions for the steady state system performance are derived from its associated steady state distribution. The developed model is validated through simulations.

The rest of this chapter is organized as follows. In Section 4.1 the proposed cross layered coding technique is described, and its performance compared with that of the H-ARQ technique. In Section 4.2 we model and analyze the effect of caching on WiMAX backhaul mesh network using Markov process. An analytical expression for the steady state system transmission delays and throughput are derived from the steady state distribution of the above Markov processes. The developed model is validated through simulations. Finally a summary is given in Section 4.3.

4.1 A Cross-Layered Coding Technique

In this section we relate the parameters of the higher layer Fountain code to those of the physical layer turbo code to present a cross-layered coding technique. In our proposed technique we assume that each data packet is encoded using turbo code with coding rate 1/2. Based on the bit error rate of physical layer turbo code, the upper layer Fountain code parameters are designed to reach the required performance. The performance of the proposed cross-layered coding technique is compared with that of physical layer H-ARQ technique [56]. In what follows we describe both coding techniques.

4.1.1 Proposed Cross-Layered Code

Upper layer forward error correction (FEC) codes such as LT codes [98], Raptor codes [97], and Fountain codes [96] are mainly designed for erasure channels. Fountain code was proposed by Mackay [96]. Basically it produces packets that are random functions of the whole file. The transmitter sends these packets without any knowledge of the packets been received. If the original file size is K

packets the receiver needs N_R packets, slightly larger than K to restore the data file.

Let $N_R = K + E$ where E is the number of excess packet needed to restore a file of K packets. For any file size, the probability that the receiver fails to recover the file is bounded by $\delta(E) \leq 2^{-E}$ [96]. Hence, the probability that the file is recovered correctly is given by $p_f \geq 1 - 2^{-E}$ and the number of extra packets required at receiver is $E = \log_2 \frac{1}{1-p_f}$. Thus

$$N_R \simeq K + \log_2 \frac{1}{1 - p_f} \tag{4.1}$$

In the proposed cross-layered approach we assume that a turbo encoder with code rate $R_{c_1} = 1/2$ is used to encode each packet on the physical layer, this collaborates with the higher layer error concealing Fountain code. Let N_{CL} and $\overline{N}_R$ denote the number of packets transmitted and the average number of packets decoded correctly at the receiver. Thus, $\overline{N}_R$ is given by the mean of a binomial distribution i.e.,

$$\overline{N}_R = ps_1^L N_{CL}, \tag{4.2}$$

where ps_1^L is the probability that a packet is accepted after travelling L links, ps_1 is the probability that a packet is successfully decoded using rate $1/2$ turbo code (i.e. $i = 1$ in Eq. A.19). Therefore the number of packets that should be sent in order to satisfy the required p_f is given by

$$N_{CL} = \lceil \frac{K + \log_2 \frac{1}{(1-p_f)}}{ps_1^L} \rceil, \tag{4.3}$$

The Fountain code parameter N_{CL} is designed based on information received from the lower layer turbo decoder. We assume the existence of a mechanism at the lower layer that estimates ps_1, for example using some CRC attached to each packet and counting the statistics of CRC checking. It is interesting to see in Eq. (4.3) that the numerator represents the Fountain code functionality, while the dominator is the merit of turbo coding.

Fountain codes are suited between the application and transport layers. In this regard we recommend utilizing UDP as the transport vehicle for its simplicity and resilience to mobility. The utilization of Fountain codes on top of UDP more than compensates for occasional loss or error of UDP datagrams. On the other hand, intermittent and channel burst errors are better handled by turbo codes at the physical layer.

Let B_{CL} denote the average number of transmitted bits per file for the proposed technique. Hence B_{CL} is calculated based on details of turbo code,

$$B_{CL} = \frac{\Gamma N_{CL}}{R_{c_1}} \tag{4.4}$$

where Γ denotes the number of data bits per packet, and $R_{c_1} = 1/2$ is the coding rate of the physical layer turbo encoder. Define the efficiency as the number of information bits over the total number of transmitted bits. Therefore, the efficiency of the proposed cross-layered code is given by

$$\begin{aligned}\eta_{CL} &= \frac{\Gamma K}{B_{CL}} \\ &= \frac{K}{\frac{N_{CL}}{R_{c_1}}}\end{aligned} \tag{4.5}$$

and the average number of time frames at the physical layer required to transmit a file considering the cumulative effect of Fountain and turbo code,

$$D_{CL} = N_{CL} + L - 1 + \sum_{i=1}^{L-1} N_{Q_i}, \tag{4.6}$$

where $L-1$ reflects the arrival of the first packet at destination after travelling L links, one should note that all the other packets arrive in a pipeline after the first one. N_{Q_i} is the number of packet delay per queue as given in Eq. (4.12).

We are studying only one possible turbo code rate i.e. $1/2$ and then the upper layer Fountain code is designed accordingly, but there is nothing in principle against adjusting the lower layer turbo code to other code rate i.e. $1/3$. In this case the upper layer Fountain code design would automatically be adjusted.

4.1.2 Hybrid ARQ

H-ARQ protocol is usually applied in order to improve the efficiency of the transmission channel (see Chapter 1). In this section we consider a turbo encoder with a final coding rate of $R_c = 1/5$. Each encoded packet is divided into three sub-packets. The first sub-packet has encoding rates of $R_{c_1} = 1/2$, the first and second sub-packets combined will have an effective lower encoding rate of $R_{c_2} =$

1/3, the three sub-packets combined will reach the final encoding rate of $R_{c3} = 1/5$.

Let ps_1, ps_2 and ps_3 denote the probabilities that a packet is decoded correctly after the first, second and third transmissions respectively. Let p_p denote the probability that a packet is finally decoded correctly (see Fig 3.3). Thus,

$$p_p = ps_1 + (1 - ps_1)ps_2 + (1 - ps_1)(1 - ps_2)ps_3 \tag{4.7}$$

where $ps_i, i = 1, 2, 3$ is obtained from Eq. (A.19).

Hence, the end-to-end probability that the requested file is decoded correctly is given by

$$p_f \;= p_p^{LK} \tag{4.8}$$

where K and L denote the number of packets per data file and the number of links each packet has to travel respectively.

Eq. (4.8) effectively says that the total file is successfully decoded at the receiver if all packets successfully propagate through all links to destination.

Let B_{HA} denote the average number of bits required to transmit a file if the H-ARQ technique is used. Hence, B_{HA} is given by

$$B_{HA} = \Gamma K(\tfrac{ps_1}{R_{c_1}} + \tfrac{(1-ps_1)ps_2}{R_{c_2}} + \tfrac{(1-ps_1)(1-ps_2)}{R_{c_3}}); \tag{4.9}$$

where ps_1, ps_2 denote the probability that the packet is decoded correctly after first, and second sub-packet transmission respectively. Γ/R_{c_1}, Γ/R_{c_2}, and Γ/R_{c_3}

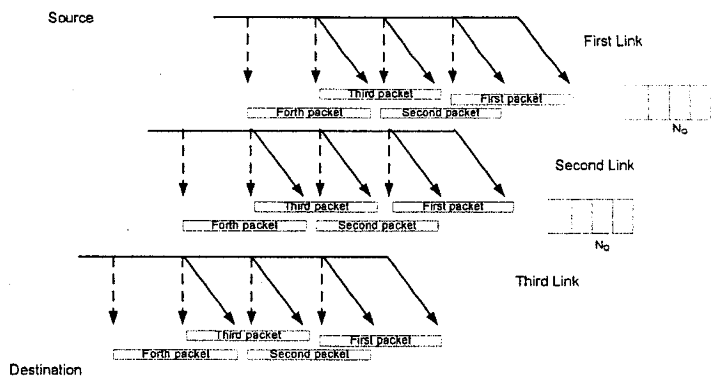


Figure 4.2: Delays due to propagation and queueing over number of links

denote the effective number of bits transmitted if a packet is accepted after the first, second, and third transmission respectively.

Thus, the efficiency of the H-ARQ technique is given by

$$\begin{aligned} \eta_{HA} &= \frac{\Gamma K}{B_{HA}} \\ &= \frac{1}{\frac{ps_1}{R_{c_1}} + \frac{(1-ps_1)ps_2}{R_{c_2}} + \frac{(1-ps_1)(1-ps_2)}{R_{c_3}}} \end{aligned} \tag{4.10}$$

As shown in Fig. 1.4 the packets are sent in a pipe. Assume that each sub-packet is transmitted over one time frame. Therefore, the average number of time frames required to transmit a file over one link, while using the H-ARQ technique is $K(ps_1 + 2(1-ps_1)ps_2 + 3(1-ps_1)(1-ps_2))$ and the end-to-end average number of time frames required to transmit a file taking into account the number of links and queueing effect (see Fig. 4.2) is given by

$$
\begin{aligned}
D_{HA} &= K(ps_1 + 2(1-ps_1)ps_2 + 3(1-ps_1)(1-ps_2)) \\
&+(L-1)(ps_1 + 5(1-ps_1)ps_2 + 9(1-ps_1)(1-ps_2)) \\
&+\sum_{i=1}^{L-1} N_{Q_i},
\end{aligned} \tag{4.11}
$$

where the term $(L-1)(ps_1+5(1-ps_1)ps_2+9(1-ps_1)(1-ps_2))$ reflects the average delay in time frames at which the first packet arrives at destination (Assuming the same transmission policy used by 1xEV standards Section 1.3.1). Remaining packets of the file arrive sequentially at the receiver in a time equal to the first term. N_{Q_i} denotes the delay in packets (time frames) associated with each queue and the $\sum_{i=1}^{L-1} N_{Q_i}$ represents the total queueing delay over all links.

4.1.3 Analysis Results

In this section efficiency and file transmission delays are obtained for both H-ARQ and the proposed cross-layered coding techniques. For analysis purpose, we consider M/D/1 queues at all SSs. The expected number of packets at the i^{th} SS queue is given by [99]

$$
N_{Q_i} = \frac{\rho_i^2}{2(1-\rho_i)} \tag{4.12}
$$

where $\rho_i = \frac{\lambda_i}{\mu_i} = 0.5$ denotes utilization factor (the ratio between the arrival λ_i and service rate μ_i) for the i^{th} SS and it is assumed constant for all the queues.

We assume that the original file size is one thousand packets. The results are obtained for $E_b/N_0 = 0, 1, \cdots, 4$ dB and the number of links $L = 1, 2, \cdots, 5$.

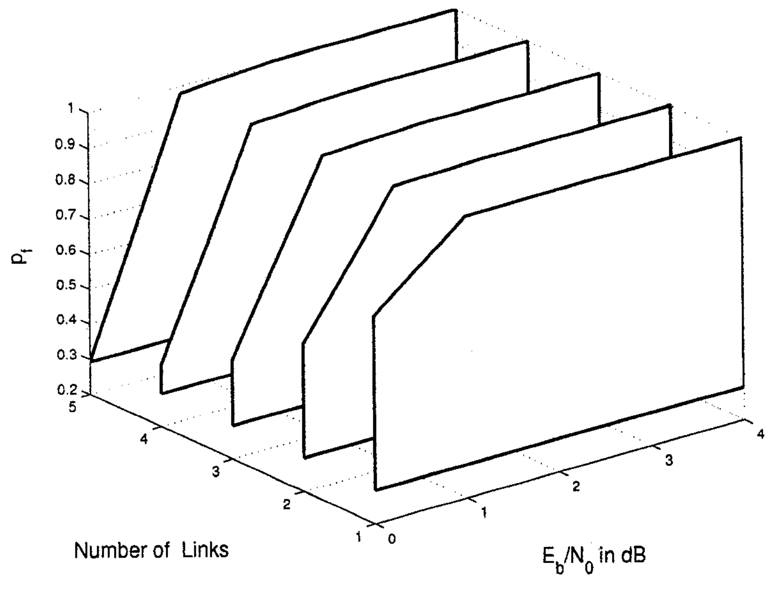


Figure 4.3: Probability of file acceptance p_f

For the H-ARQ case Fig. 4.3 shows the probability of file acceptance p_f in Eq. (4.8) for different E_b/N_0 and different number of links. The same p_f is used as an input parameter to conduct the analysis of the cross-layered coding technique so that we have a fair comparison. The figure shows that p_f improves as E_b/N_0 gets higher and the effect of the number of links almost vanishes for high E_b/N_0.

Figs. 4.4 and 4.5 show the efficiency of the H-ARQ and cross-layered coding techniques as in Eqs. (4.10 and 4.5) respectively. It is clear that for $E_b/N_0 \geq 3\ dB$ the efficiency of the cross-layered coding technique is almost the same as this of the H-ARQ.

The effective number of time frames required to transmit a file for both H-ARQ and cross-layered coding techniques as in Eqs. (4.11 and 4.6) are shown in Figs. 4.6 and 4.7 respectively. One can notice that for $E_b/N_0 \geq 1$ dB the performance of

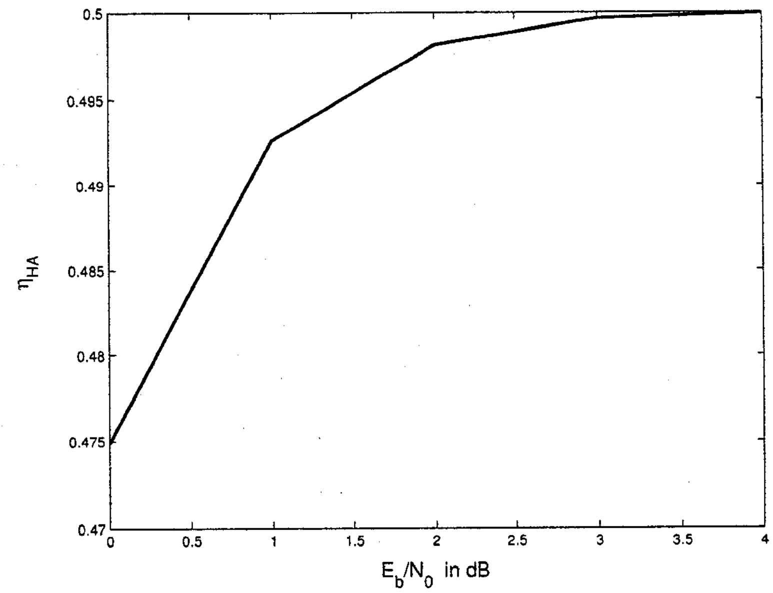


Figure 4.4: Efficiency of H-ARQ technique η_{HA}

the cross-layered coding technique is comparable with that of the H-ARQ technique even for large number of links. Note that the cross-layered FEC approach is an end to end forward approach that does not suffer the ARQ dialog necessary at each link as in the case of H-ARQ. Thus, by using the cross-layered FEC we can reach almost the same performance of H-ARQ, with the privilege that the overhead of the H-ARQ technique is removed.

It is also important to mention that the proposed cross-layered coding technique fits better for files with large number of packets otherwise the efficiency will get lower. As an application, the proposed cross-layered FEC technique is used to encode data files while we study the effects of caching data on WiMAX backhaul mesh network.

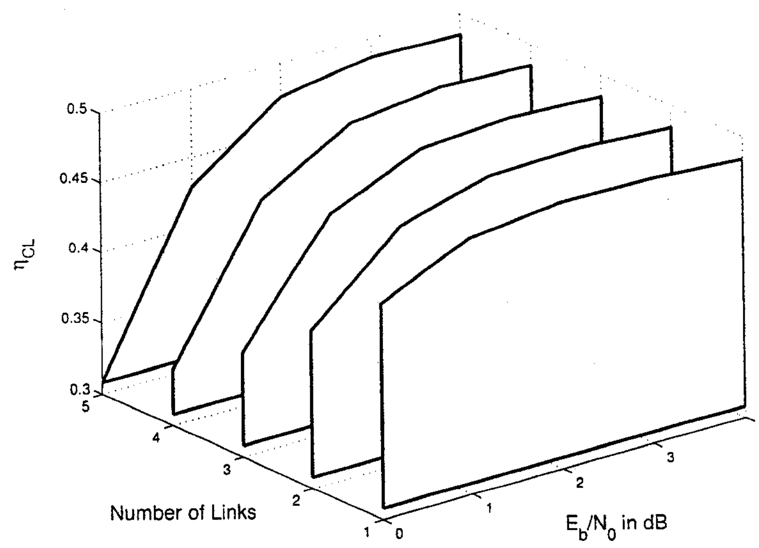


Figure 4.5: Efficiency of cross-layer coding technique η_{CL}

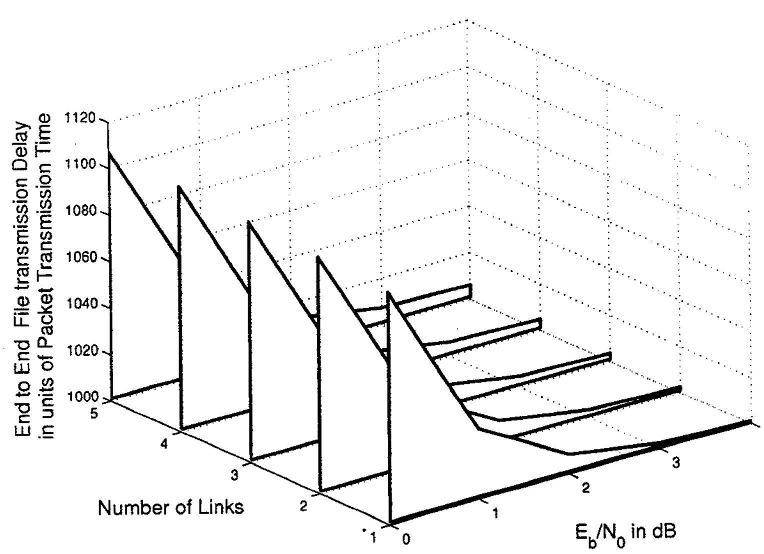


Figure 4.6: Delay of H-ARQ technique in number of required time frames to transmit a file D_{HA}

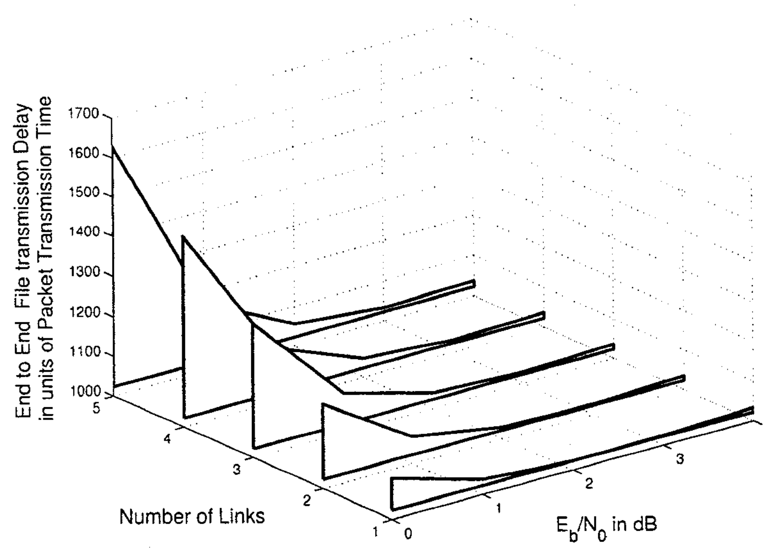


Figure 4.7: Delay of cross-layered coding technique in number of required time frames to transmit a file D_{CL}

4.2 Network Performance Model

In this section we model and study the effect of caching on the performance of WiMAX backhaul mesh network. As mentioned above the proposed cross-layered coding technique is used as the WiMAX coding scheme.

We study a distributed WiMAX backhaul mesh network where each local SS has cached some of the information required by its end users including actual web-pages, files, network topology, and other control parameters. Each end user node has the option to contact the BS directly or a SS, it will select either route based on the required QoS. If the request is processed through a SS, the SS tries to locate the required data in its cache or in other SSs caches along its route to the BS. Hence, the delay is reduced. Typically, the data rate assigned to an end user node served directly through the BS is lower than the one assigned to an end user node served by a SS. An example of the used setting is given in Fig. 4.1.

4.2.1 Caching Effect Model

Let p_{miss} denote the probability that a file requested by the end user node is not found within its local SS cache, Z_{SS} denote the total number of SSs in the mesh network. Assume a network with a maximum depth of L_{max} links, where $2 \leq L_{max} \leq Z_{SS} + 1$. Let K_{BS} and K_{SS} denote the number of data packets per file if the end user node is served by BS and SS respectively. Let N_{BS} denote the total number of packets required to transmit a file if the end user node is served directly by the BS. We can also notice that if the user is served by the BS $L = 1$, hence substitute in Eq. (4.3)

$$N_{BS} = \lceil \frac{K_{BS} + \log_2 \frac{1}{\delta}}{ps_1} \rceil, \tag{4.13}$$

where $(1 - \delta = p_f)$ denote the required probability of successful file recovery.

As for the requests served by a SS the total number of packets depends on the number of the links the requested file has to traverse until the end user node. Let $N_{SS}(l)$ denote the total number of packets if the file travels l links.

$$N_{SS}(l) = \lceil \frac{K_{SS} + \log_2 \frac{1}{\delta}}{ps_1^l} \rceil, \tag{4.14}$$

Hence, from Fig. 4.1 on average N_{SS} is given by

$$N_{SS} = \lceil \sum_{i=2}^{L_{max}} \frac{N_{SS}(i)p_{miss}^{i-1} + \sum_{j=1}^{i-1} N_{SS}(j)p_{miss}^{j-1}(1-p_{miss})}{L_{max}-1} \rceil \quad (4.15)$$

To clarify the above equation, assume the case when $i = 3$, i.e. the end user node is located at distance of three links from the BS. Hence, the request may travel only one link with a probability $(1 - p_{miss})$ (the requested file is cached at the first SS along its route to the BS). The other scenario is the file is not cached at the first SS but it is cached at the second SS along its route, therefore the file travels two links with probability $p_{miss}(1 - p_{miss})$. Finally, if the file is not cached at any of the SSs along its route the request is served by the BS and the file has to travel three links with a probability p_{miss}^2.

Define $\mathbf{x} = \{x_1, x_2\}$, where x_1 and x_2 denote the number of requests directly served by the BS or by all SSs respectively. $W_{BS}(\mathbf{x})$ and $W_{SS}(\mathbf{x})$ denote the available BW in packets per time frame for each request if it is processed by the BS or by a SS respectively. In a simple policy the end user node should decide to connect to the BS or SS based only on the availability of the requested file in the cache of the SS. In a more elaborate policy the end user node estimates the total file transmission time if it hooks to the BS or a SS and make its decision accordingly. Let $T_{BS}(\mathbf{x}) = \frac{N_{BS}}{W_{BS}(\mathbf{x})}$ and $T_{SS}(\mathbf{x}) = \frac{N_{SS}}{W_{SS}(\mathbf{x})}$ denote the number of time frames required to transmit a file if the request is served directly by the BS or a SS respectively. The end user node selects the shorter of the two with a certain probability. In this case the probability of the file being processed by BS is given by

$$p_{BS}(\mathrm{x}) = \frac{T_{SS}(\mathrm{x})}{T_{BS}(\mathrm{x})+T_{SS}(\mathrm{x})} \tag{4.16}$$

The above equation implies that if $T_{BS}(\mathrm{x}) < T_{SS}(\mathrm{x})$, the probability that a request is served by BS is greater than the probability that it is served by a SS.

Let W denote the maximum WiMAX BW in packets per time frame for BS or any SS, hence $W_{BS}(\mathrm{x})$ is given by

$$W_{BS}(\mathrm{x}) = \begin{cases} \frac{W}{x_1+x_{BS_{miss}}} & x_1 + x_{BS_{miss}} \geq 1, \\ \\ W & x_1 + x_{BS_{miss}} < 1, \end{cases} \tag{4.17}$$

The term $x_{BS_{miss}}$ represents the average number of requests that are assigned to SSs and reached the BS due to cache missing. Assume that all SSs are distributed uniformly over $L_{max} - 1$ links and x_2 is distributed uniformly over all SSs (see Fig. 4.1). Hence, $\frac{x_2}{L_{max}-1}$ requests could reach the BS with probability p_{miss}, another $\frac{x_2}{L_{max}-1}$ requests could reach the BS with probability p^2_{miss} and so on. Therefore $x_{BS_{miss}}$ is given by,

$$x_{BS_{miss}} = \sum_{k=1}^{L_{max}-1} \frac{x_2}{L_{max}-1} p^k_{miss} \tag{4.18}$$

Assume that each SS will serve on average $\frac{x_2}{Z_{SS}}$ requests, where Z_{SS} denote the total number of SSs, plus the number of requests assigned to other SSs and reach it due to p_{miss} denoted by $x_{SS_{miss}}$. Let $W_{SS}(\mathrm{x}, l)$ denote the average available BW at any SS distant $l-1$ links from BS. Therefore

$$W_{SS}(\mathbf{x},l) = \begin{cases} \frac{W}{\frac{x_2}{Z_{SS}}+x_{SS_{miss}}(l)} & \frac{x_2}{Z_{SS}} + x_{SS_{miss}}(l) \geq 1 \\ W & \frac{x_2}{Z_{SS}} + x_{SS_{miss}}(l) < 1; \end{cases} \tag{4.19}$$

Similar to Eq. 4.18, $x_{SS_{miss}}(l)$ is given by,

$$x_{SS_{miss}}(l) \approx \sum_{k=1}^{L_{max}-l} \frac{x_2}{L_{max}-1} p_{miss}^k \tag{4.20}$$

Similar to Eq. 4.15, if the requested node is located at a distance of 3 links from the BS, the available BW for this request is $W_{SS}(\mathbf{x},3)$ with probability $(1-p_{miss})$, $W_{SS}(\mathbf{x},2)$ with probability $p_{miss}(1-p_{miss})$ or $W_{BS}(\mathbf{x})$ with probability p_{miss}^2. Hence, the average BW available for any request served by a SS is given by

$$W_{SS}(\mathbf{x}) = \sum_{i=2}^{Lmax} \frac{W_{BS}(\mathbf{x})p_{miss}^{i-1} + \sum_{j=2}^{i} W_{SS}(\mathbf{x},j)p_{miss}^{i-j}(1-p_{miss})}{L_{max}-1} \tag{4.21}$$

4.2.2 Transmission Process Model

As discussed previously, Markov models are among the most powerful tools available for analyzing communications systems. If the Markov model is irreducible, aperiodic and positive recurrent, the state probabilities reach steady state values that are independent of the initial state probabilities. The steady state probabilities $\pi(\mathbf{x})$ for a state $\mathbf{x}$ could be obtained by solving the linear equations

$$\pi(\mathbf{y}) = \sum_{\forall \mathbf{x}} p_{\mathbf{xy}} \pi(\mathbf{x})$$

and

$$\sum_{\forall \mathbf{x}} \pi(\mathbf{x}) = 1$$

where $p_{\mathbf{xy}}$ is the transition probability from state $\mathbf{x}$ to state $\mathbf{y}$.

In this section, using Markov models we derive the steady state distribution of the number of requests served by either the BS or SSs in a WiMAX backhaul network where caching is allowed at SSs. Hence, the steady state performance of the network is determined.

Define a state vector $\mathbf{R}(t) = (R_1(t), R_2(t))$, where $R_1(t)$ and $R_2(t)$ are the number of requests served directly by the BS and by other SS nodes respectively at time t. Without loss of generality, we assume that the system can handle R_{max} requests which implies that $R_1(t) + R_2(t) = R_{max}$ at any time t. Hence, the number of valid states is given by $R_{max} + 1$.

The distance between the embedded points of the considered Markov chain is one frame time. Let the vector $\mathbf{F}(t+1) = (F_1(t+1), F_2(t+1))$, denote the number of requests that are already finished downloading at the end of the t^{th} time slot, i.e., at the beginning of slot $t+1$. Similarly, Let $\mathbf{A}(t+1) = (A_1(t+1), A_2(t+1))$ denote the number of requests that arrived at time slot $(t+1)$. Thus

$$\mathbf{R}(t+1) = \mathbf{R}(t) - \mathbf{F}(t+1) + \mathbf{A}(t+1). \tag{4.22}$$

Let $\mathbf{x} = (x_1, x_2)$, and $\mathbf{y} = (y_1, y_2)$ be the sample of $\mathbf{R}(t)$ and $\mathbf{R}(t+1)$ respectively.

Similarly let $\mathbf{f} = (f_1, f_2)$, and $\mathbf{a} = (a_1, a_2)$ be the sample of $\mathbf{F}(t+1)$ and $\mathbf{A}(t+1)$ respectively.

Therefore the state transition probability $p_{\mathbf{xy}} = p(\mathbf{R}(t+1) = \mathbf{y}|\mathbf{R}(t) = \mathbf{x})$ is given by

$$p_{\mathbf{xy}} = \sum_{\mathbf{f},\mathbf{a}} p(\mathbf{R}(t+1) = \mathbf{y}|\mathbf{a},\mathbf{f},\mathbf{x}) \times p(\mathbf{A} = \mathbf{a}|\mathbf{f},\mathbf{x}) \times p(\mathbf{F} = \mathbf{f}|\mathbf{x}) \tag{4.23}$$

The conditional probability $p(\mathbf{R}(t+1) = \mathbf{y}|\mathbf{f},\mathbf{a},\mathbf{x})$ is given by

$$p(\mathbf{R}(t+1) = \mathbf{y}|\mathbf{f},\mathbf{a},\mathbf{x}) = \prod_{i=1}^{2} p(y_i = x_i - f_i + a_i) \tag{4.24}$$

where,

$$p(y_i = x_i - f_i + a_i) = \begin{cases} 1, & y_i = x_i - f_i + a_i, \\ 0, & y_i \neq x_i - f_i + a_i. \end{cases} \tag{4.25}$$

where the above equation represents the legitimate transitions.

Let p_a denote the probability that a new arrival is served by the BS. As mentioned before the end user node chooses to be served by the BS or by SS based on the estimated downloading time. Hence, refer to Eq. (4.16) p_a is given by

$$p_a = p_{BS}(\mathbf{x} - \mathbf{f}) \tag{4.26}$$

which depends on the current distribution at the beginning of time slot $t+1$ once $\mathbf{f} = \{f_1, f_2\}$ requests finished downloading. Hence, the probability $p(\mathbf{A} = \mathbf{a}|\mathbf{f},\mathbf{x})$ is given by the following binomial distribution,

$$p(\mathbf{A}=\mathbf{a}|\mathbf{f},\mathbf{x}) = \binom{a_1+a_2}{a_1} p_a^{a_1}(1-p_a)^{a_2} \tag{4.27}$$

which is the probability that a_1 requests choose to be directly served by BS, and a_2 requests choose to be served by all SSs.

Let $p_1(\mathbf{x}) \approx \frac{1}{T_{BS}(\mathbf{x})}$, and $p_2(\mathbf{x}) \approx \frac{1}{T_{SS}(\mathbf{x})}$ denote the probability that the number of packets left to be downloaded for a request served by the BS or a SS respectively at the beginning of time slot t is equal to the available BW. Where, $T_{BS}(\mathbf{x})$ and $T_{SS}(\mathbf{x})$ denote the required number of time frames to transmit a file if it is served by the BS or by a SS respectively as defined in Section 4.2. Let $p(f_k|\mathbf{x})$ denote the probability that f_k requests that are served by either BS ($k = 1$) or SSs ($k = 2$) finishes downloading at time slot $t+1$. The above probabilities are modelled as binomial distribution i.e.,

$$p(f_1|\mathbf{x}) = \binom{x_1}{f_1} (p_1(\mathbf{x}))^{f_1}(1-p_1(\mathbf{x}))^{(x_1-f_1)}, \tag{4.28}$$

and,

$$p(f_2|\mathbf{x}) = \binom{x_2}{f_2} (p_2(\mathbf{x}))^{f_2}(1-p_2(\mathbf{x}))^{(x_2-f_2)}, \tag{4.29}$$

The probability $p(\mathbf{F}=\mathbf{f}|\mathbf{x})$ is therefore obtained by multiplying the above probabilities

$$p(\mathbf{F}=\mathbf{f}|\mathbf{x}) = \prod_{k=1}^{2} p(f_k|\mathbf{x}) \tag{4.30}$$

By noting that the above Markov model is irreducible, aperiodic, and positive recurrent [85], there will be a unique steady-state probability $\pi(\mathbf{x})$ for a state $\mathbf{x}$ and the steady state average available BW if the request is served by BS and by a SS is given by

$$\overline{W}_{BS} = \sum_{\forall \mathbf{x}} W_{BS}(\mathbf{x})\pi(\mathbf{x})$$

$$\overline{W}_{SS} = \sum_{\forall \mathbf{x}} W_{SS}(\mathbf{x})\pi(\mathbf{x})$$

The average effective steady state throughput (in bits per frame time) is given by

$$\overline{TH}_{BS} = \frac{F}{\overline{D}_{BS}}$$

$$\overline{TH}_{SS} = \frac{F}{\overline{D}_{SS}}$$

where F is the number of information bits per file. $\overline{D}_{BS} = \frac{N_{BS}}{\overline{W}_{BS}}$ and $\overline{D}_{SS} = \frac{N_{SS}}{\overline{W}_{SS}}$ denote the steady state average file transfer delay in time frames T_f, if the request is served by BS or a SS respectively.

4.2.3 Analytical and Simulation Results

In our analysis we consider a 5 MHz WiMAX backhaul mesh network with one BS and nine SSs. We assume each end user node has enough power to reach $ps_1 = 0.8$. The probability that the receiver fails to recover the file is set to $\delta = 10^{-5}$

and hence $p_f = 1 - 10^{-5}$. The values of W_{BS} and W_{SS} are allowed to be fractions which reflect the effect of time scheduling.

The above analysis is applied to end users downloading web pages. Let F_{main} and F_{emb} denote the number of bits per main object and embedded object respectively. Let N_{emb} denote the number of embedded objects per page as shown in Table 4.1 [56]. Hence, the total page size F in bits is given by

$$F = F_{main} + N_{emb}F_{emb} \tag{4.31}$$

We average for the purpose of analysis, i.e.

$$F = \overline{F}_{main} + \overline{N}_{emb}\overline{F}_{emb} \tag{4.32}$$

We assume that each packet occupies one sub-channel for the whole frame duration (one row in Fig. 1.5). Since the request served directly by the BS it is assigned lower data rate than that served through a SS, thus we assume that $K_{BS} = 2K_{SS} = \frac{F}{2N_cN_{sym}}$ data packets, where N_c is the number of sub-carriers per sub-channel, and N_{sym} is the number of symbols for the whole frame duration (Table 1.5).

Throughout the simulation different network topologies that are generated randomly were considered. For each topology the simulation runs for a large number of time frames (100000). The results are averaged over all topologies and time frames.

Figs. 4.8 and 4.9 show the average steady state BW per request with different

Component	Distribution	Parameters	PDF
Main object	Truncated lognormal	Min=100 bytes Max=2 Mbytes mean=10710 bytes Std. dev=25032 bytes	$g(x) = \frac{1}{\sqrt{2\Pi}\sigma x} \exp \frac{-(\ln x-\mu)^2}{2\sigma^2}$, $x \geq 0$ $\sigma = 1.37, \mu = 8.35$
Embedded object	Truncated lognormal	Min=50 bytes Max=2 Mbytes mean=7758 bytes Std. dev=126168 bytes	$g(x) = \frac{1}{\sqrt{2\Pi}\sigma x} \exp \frac{-(\ln x-\mu)^2}{2\sigma^2}$, $x \geq 0$ $\sigma = 2.36, \mu = 6.17$
Number of Embedded objects	Truncated Pareto	mean=5.64 Max=53	$g(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}, k \leq x < m$ $g(x) = (\frac{k}{m})^\alpha, x = m$ $\alpha = 1.1, k = 2, m = 55$

Table 4.1: HTTP traffic model parameters

p_{miss} and R_{max} respectively. The results are obtained for both BS and SS. It is clear that the size of cache per SS which defines p_{miss} and the total number of requests have a great impact on the average available BW per request. Studying the effect of the cache size on the available BW per request we notice that the average BW for a request served directly by BS is smaller than that of a request served by a SS. This reflects the fact that BS receives missed requests from all SSs in the network, as for the SSs the missed requests and the number of users are distributed among them. One can also recognize that as p_{miss} gets larger the average BW per request served by a SS approaches that of a request served by BS, this simply happens because for smaller cache sizes requested files are not found at SSs and hence most of the requests reach the BS. Therefore the effective BW per request is constrained with that of the BS (see Eq. 4.21).

Fig. 4.10 shows the average number of packets for a request served by BS or SS (see Eqs. 4.13 and 4.15). One may notice that while p_{miss} has no effect on N_{BS}, the number of packets N_{SS} grow as p_{miss} gets larger. This can be clarified by noting that N_{SS} is a function of the number of links and as p_{miss} gets larger the number of links each packet has to travel gets higher. As for the average number of packet for a request served by BS it is only a function in ps_1.

In Fig. 4.11 we can see the effect of the p_{miss} and on the average throughput per request if served by BS and SS. The impact of caching is clearly seen through these results. Fig. 4.12 shows the change of the average throughput per request with the total number of requests if served by BS or SS. It is obvious that the through-

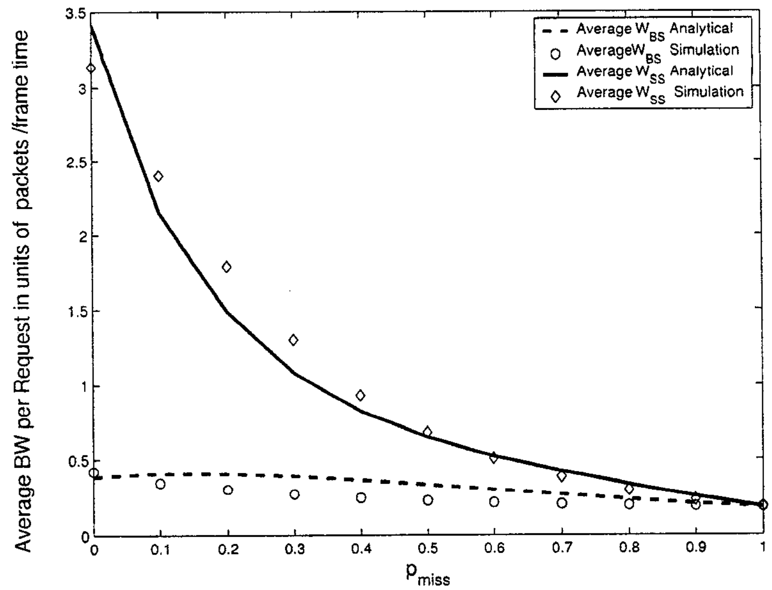


Figure 4.8: Average BW per request. $R_{max} = 90$, $L_{max} = 4$, $p_{s1} = 0.8$, $\delta = 10^{-5}$, and $Z_{SS} = 9$

put gets higher for higher number of requests, however the average throughput per user is reduced.

Figs. 4.13 and 4.14 show the steady state distribution of the requests with different p_{miss} and R_{max} respectively. It should be noted that when $p_{miss} = 0$ some of the requests will still be served by the BS, even though this requires longer downloading time because of the probabilistic decision (Eq. 4.16), and hence will stay longer in the system.

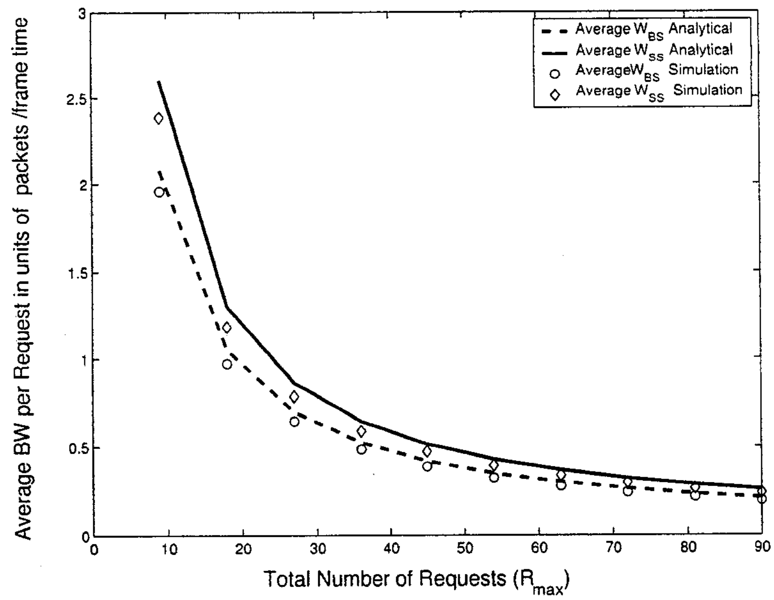


Figure 4.9: Average BW per request. $p_{miss} = 0.9$, $p_{s1} = 0.8$, $L_{max} = 4$, $\delta = 10^{-5}$, and $Z_{SS} = 9$

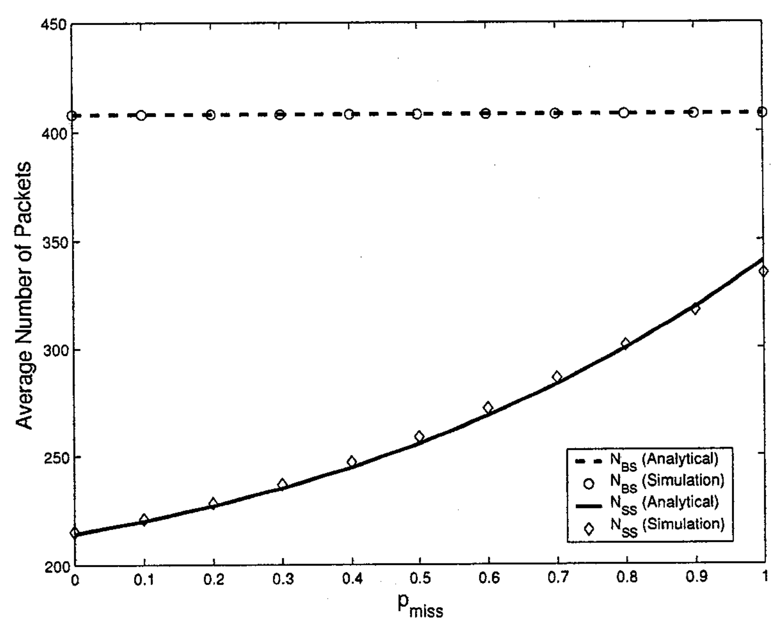


Figure 4.10: Average number of packets per request. $L_{max} = 4, \delta = 10^{-5}, p_{s1} = 0.8$, and $Z_{SS} = 9$

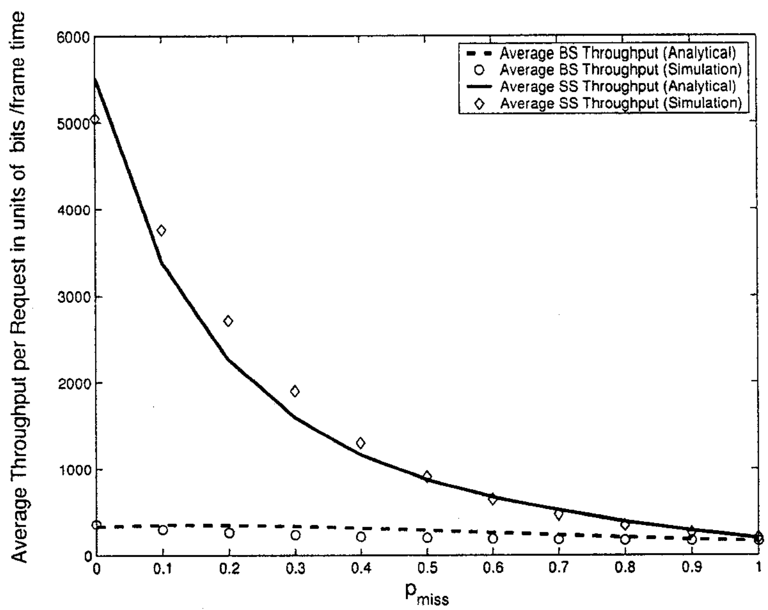


Figure 4.11: Average throughput per request. $R_{max} = 90$, $L_{max} = 4$, $\delta = 10^{-5}$, and $Z_{SS} = 9$

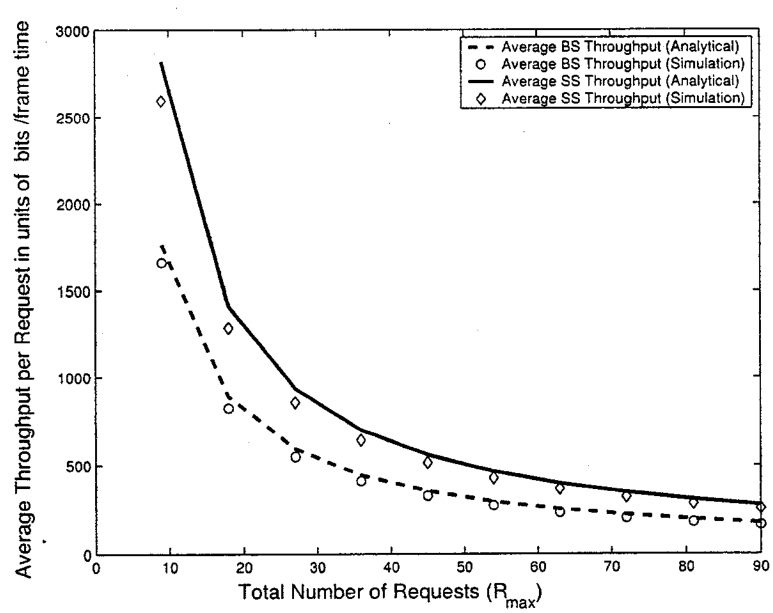


Figure 4.12: Average throughput per request for BS. $p_{miss} = 0.9$, $L_{max} = 4$, $\delta = 10^{-5}$, $p_{s1} = 0.8$, and $Z_{SS} = 9$

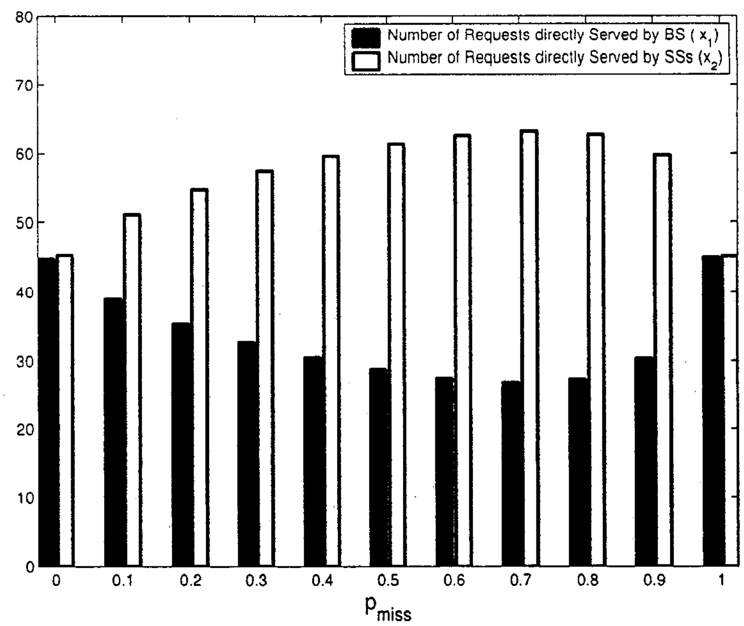


Figure 4.13: Steady state distribution. $R_{max} = 90$, $L_{max} = 4$, $\delta = 10^{-5}$, $p_{s1} = 0.8$, and $Z_{SS} = 9$

4.3 Summary

In this chapter we proposed a cross-layered coding technique in which upper layer Fountain codes is applied on top of physical layer turbo code. The results showed that the performance of the proposed technique is comparable to that of the H-ARQ technique. We also considered a WiMAX backhaul mesh network in which caching is allowed at SSs and UDP is used as its transport protocol. The network is modelled and analyzed using Markov process. An analytical expression for the steady state system transmission delays and throughput are derived from the steady state distribution of the above Markov process. The developed model is validated through simulations.

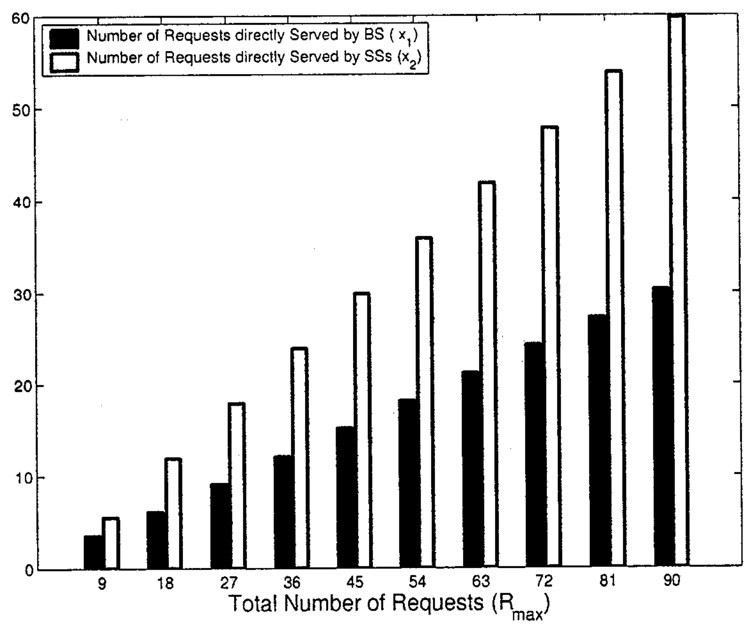


Figure 4.14: Steady state distribution. $p_{miss} = 0.9$, $L_{max} = 4$, $\delta = 10^{-5}$, $p_{s1} = 0.8$, and $Z_{SS} = 9$

Chapter 5

Conclusion

In this thesis, the developed RoT model is used to determine a bound for the maximum theoretical throughput that can be achieved over the reverse data channel in cdma2000 1xEV star network. It is shown that the maximum instantaneous throughput can be achieved if only a subset of active users are allowed to transmit on each time slot. We proposed two new rate assignment techniques. Our simulation results show that higher throughput and lower service times are achieved using the new proposed techniques. We also proposed two new scheduling techniques; our simulation results show that the proposed algorithm enhances the sector throughput compared to the widely used proportional fair scheduling algorithm.

We also modelled four scheduling techniques for the reverse data channel of the 1xEV star network. The distribution of the MSs among the possible data rates is modelled as a Markov process. The effects of scheduling and rate assignment techniques are included in the Markov model. The validity of the obtained models was confirmed with simulations. The results show that highest-rate-first scheduling technique outperforms all the other scheduling techniques in the sense of achieved throughput. The simulation results are fairly close to their analysis counterparts, which validates the assumptions of the analysis.

We extended the above model to develop analytical expressions for the steady state link throughput and file transfer delay over the reverse data channel in cdma2000 1XEV where the H-ARQ protocol is applied. These expressions are based on modelling the cross layer design involving the physical layer H-ARQ, network layer rate assignment, and MAC layer time slot scheduling technique used over the reverse data channel. The results obtained show how the choice of the targeted signal to noise and interference ratio affect the overall system performance.

Finally, we proposed a cross layered coding technique in which upper layer Fountain code is applied on top of the physical layer turbo encoder. The results showed that the performance of the proposed technique is comparable to that of the H-ARQ technique. We also considered a WiMAX backhaul mesh network in which caching is allowed at SSs. The network is modelled and analyzed using Markov process. An analytical expression for the steady state system transmission delays and throughput are derived from the steady state distribution of the Markov process. The developed model is validated through simulations. The results shows that, both the cash size at each SS and the total number or requests being handled have a great effect on the network performance.

5.1 Contributions

Throughout this thesis we considered resource management for 3G-4G technology in star and mesh networks. In particular we have

- obtained maximum theoretical throughput that can be achieved over the reverse packet data channel in cdma2000 1xEV star network,
- shown that the maximum instantaneous throughput can be achieved if only a subset of the active users is allowed to transmit on each time slot,
- proposed two new rate assignment techniques,
- proposed two new scheduling techniques,
- derived an analytical models for four scheduling techniques,
- modelled the transmission process over the reverse data channel as Markov process,
- analyzed the cdma2000 1xEV turbo encoder and obtained values for the probability of packet acceptance after first, second and third sub-packet transmissions,
- extended the transmission process model to include the effects of the cross layer design between H-ARQ, rate assignment, and scheduling,
- proposed new cross layered coding technique which allows the use of UDP with a prespecified performance,

- introduced the idea of caching at service station in WiMAX backhaul mesh networks,
- modelled the performance WiMAX backhaul mesh networks as a Markov process,
- provided expressions for the steady state performance for all the models developed throughout the thesis,
- and validated all the developed models through simulation.

It should be noted that while, in this thesis, we considered specific standards, the techniques used throughout our work are generic and can be applied to other types of star and mesh networks.

5.2 Future Work

In a typical cellular system, the maximum allowable transmit power of the mobile stations limit its maximum permissible data rates. It would be interesting to consider how this limitation would affect the overall system throughput and service time.

The rise over thermal model in this thesis considers only the intra-cell interference. The effects of the inter-cell interference on the performance of data transmissions should be further investigated.

The probability of cache missing has a great impact on the performance of

WiMAX backhaul mesh network. This probability mainly depends on the caching policy used at the service station. Studying the effect of different caching policies on the WiMAX backhual mesh networks should be explored.

During the course of this thesis the impact of handovers are not been considered. For future research it is important to study and model the impact of handovers on all the proposed techniques.

BIBLIOGRAPHY

[1] CITA, *CTIAs Semi-Annual Wireless Industry Survey,* Dec. 2007.

[2] S. Anand, and A. Chockalingam, *Performance Analysis of Voice/Data Cellular CDMA with SIR-Based Admission Control,* IEEE Journal on Selected Areas in Communications, vol. 21, pp. 1674-84, Dec. 2003.

[3] M.H. Ahmed, *Call Admission Control in Wireless Networks: a Comprehensive Survey,* IEEE Communications Surveys and Tutorials, vol. 7, pp. 49 - 68, First Qtr. 2005.

[4] Z. Liu, and M. El Zarki, *SIR-based Call Admission Control for DS-CDMA Cellular Systems,* IEEE Journal on Selected Areas in Communications, vol. 12, pp. 638-44, May 1994.

[5] Z. Dziong, M. Jia, and P. Mermelstein, *Adaptive Traffic Admission for Integrated Services in CDMA Wireless Access Networks,* IEEE Journal on Selected Areas in Communications, vol. 14, pp. 1737-47, Dec. 1996.

[6] Y. Ishikawa, and N. Umeda, *Capacity Design and Performance of Call Admission Control in Cellular CDMA Systems,* IEEE Journal on Selected Areas in Communications, vol. 15, pp. 1627-35, Oct. 1997.

[7] Y. Guo, and B. Aazhong, *Call Admission Control in Multiclass Traffic CDMA Cellular System Using Multiuser Antenna Array Receiver,* Proceedings of IEEE Vehicular Technology Conference (VTC 2000), vol. 1, pp. 365-69, May 2000.

[8] J. Evans, and D. Everitt, *Effective Bandwidth-based Admission Control for Multi-service CDMA Cellular Networks*, Proceedings of IEEE Vehicular Technology Conference (VTC 1999), vol. 48, pp. 36-46, Jan. 1999.

[9] M. Andersin, Z. Rosberg, and J. Zander, *Soft and Safe Admission Control in Cellular Networks*, IEEE Transaction on Networks, vol. 5, pp. 255-65, April 1997.

[10] J. Kuri, and P. Mermelstein, *Call Admission on the Uplink of a CDMA System Based on Total Received Power Communications*, Proceedings of IEEE International Conference on Communications (ICC 1999), vol. 3, pp. 1431-36, 1999.

[11] W. Yang, and E. Geraniotis, *Admission Policies for Integrated Voice and Data Traffic in CDMA Packet Radio Networks*, IEEE Journal on Selected Areas in Communications, vol. 12, pp. 654-64, May 1994.

[12] Lei Wang, and Weihua Zhuang, *A Call Admission Control Scheme for Packet Data in CDMA Cellular Communications*, IEEE Transactions on Wireless Communications, vol. 5, no. 2, pp. 406-16, Feb. 2006.

[13] Hwanjoon Kwon, Younsun Kim, Jin-Kyu Han, Donghee Kim, Hyeon Woo Lee, and Young Kyun Kim, *Performance Evaluation of High-Speed Packet Enhancement on cdma2000 1xEV-DV*, IEEE Communications Magazine, vol. 43, pp. 67-73, April 2005.

[14] H. Kwon, Y. Kim, J. Han, and D. Kim, *An Efficient Radio Resource management Technique for the Reverse Link in cdma2000 1XEV-DV,* Proceedings of IEEE Wireless Communications and Networking Conference (WNCS 2005), vol. 1, pp. 364-68, March 2005.

[15] T. Shu, and Z. Niu, *A Dynamic Rate Assignment Scheme for Data Traffic in Cellular Multi-code CDMA Networks,* Proceedings of IEEE Vehicular Technology Conference (VTC 2002), vol. 2, pp. 811-15, May 2002.

[16] W. Yeo, and D. Cho, *An Analytical Model for Reverse Link Rate Control in cdma2000 1xEV-DO Systems,* IEEE Communications Letters, vol. 9, no.3, pp. 270-72, March 2005.

[17] H. Lee, W. Yeo, and D. Cho, *New Rate Control Scheme Based on Adaptive Rate Limit for in 1xEV-DO Reverse Link Traffic Channels,* IEEE Communications Letters, vol. 9, pp. 903-5, October 2005.

[18] H. Lee, W. Yeo, and D. Cho, *Adaptive Reverse Link Rate Control Scheme for cdma2000 1xEV-DO Systems,* Proceedings of IEEE Vehicular Technology Conference (VTC 2005), vol. 3, pp. 1441-45, May-June 2005.

[19] Woon-Young Yeo, and Dong-Ho Cho, *Markovian Approach for Modeling IS-856 Reverse Link Rate Control,* Proceedings of IEEE International Conference on Communications (ICC 2004), vol. 6, pp. 3236-40, June 2004.

[20] W. Yeo, and D. Cho, *Enhanced Rate Control Scheme for 1xEV-DO Reverse Traffic Channels,* IEEE Communications Letters, vol. 39, pp. 1677-79, Nov. 2003.

[21] R. Vannithamby, and E. Sousa, *Resource Allocation and Scheduling Schemes for WCDMA Downlinks,* Proceedings of IEEE International Conference on Communications (ICC 2001), vol. 5, pp. 1406-10, June 2001.

[22] F. Rodriguez, and D. Goodman, *Power and Data Rate Assignment for Maximal Weighted Throughput in 3G CDMA,* Proceedings of IEEE Wireless Communications and Networking (WCNC 2003), vol. 1, pp. 525-31, March 2003.

[23] F. Rodriguez, D. Goodman, and Z. Marantz, *Power and Data Rate Assignment for Maximal Weighted Throughput in 3G CDMA: A Global Solution with Two Classes of Users,* Proceedings of IEEE Wireless Communications and Networking (WCNC 2004), vol. 4, pp. 2201-06, March 2004.

[24] P. Bjorklund, P. Varbrand, and Di Yuan, *PA Dynamic Programming Technique for Downlink Bandwidth Allocation in WCDMA Networks,* Proceedings of IEEE Vehicular Technology Conference (VTC 2004), vol. 4, pp. 2007-11, May 2004.

[25] C. Carciofi, and P. Grazioso, *Radio Resource Management Strategies for Packet Data Services in UMTS,* Proceedings of IEEE Vehicular Technology Conference (VTC 2001), vol. 2, pp. 1012-16, May 2001.

[26] J. Price, and T. Javidi, *Cross-Layer (Mac and Transport) Optimal Rate Assignment in CDMA-Based Wireless Broadband Networks,* Proceedings of IEEE Conference of Signals, Systems and Computers, vol. 1, pp. 1044-48, Nov. 2004.

[27] J. Price, and T. Javidi, *Decentralized and Fair Rate Control in a Multi-Sector CDMA System,* Proceedings of IEEE Wireless Communications and Networking (WCNC 2004), vol. 4, pp. 2189-94, March 2004.

[28] J. Price, and T. Javidi, *Decentralized Rate Assignments in a Multi-Sector CDMA Network,* IEEE Transactions on Wireless Communications, vol. 5, pp. 3537-47, December. 2006.

[29] H. Fattah, and C. Leung, *Load-Based Transmission Rate Assignment Scheme for Integrated Voice/Data DS-CDMA System,* IEE Electronics Letters, vol. 39, pp. 1011-13, June 2003.

[30] S. Ci, and M. Guizani, *A Dynamic Resource Scheduling Scheme for CDMA2000 Systems,* Proceedings of third ACS/ IEEE International Conference on Computer Systems and Applications, 2005.

[31] S. Ci, M. Guizani, and G. Brahim, *A Dynamic Resource Allocation Scheme for Delay-Constrained Multimedia Services in CDMA 1xEV-DV Forward Link,* IEEE Journal on Selected Areas in Communications, vol. 24, pp. 46-53, Jan. 2006.

[32] A. Jalali, R Padovani, and R. Pankaj, *Data Throughput of CDMA-HDR a High Efficiency - High Data Rate Personal Communication Wireless Systems,* Proceed-

ings of IEEE Vehicular Technology Conference (VTC 2000), pp. 1854-58, May 2000.

[33] S. Shakkottai, and A.L. Stolyar, *Scheduling Algorithms for a Mixture of Real Time and Non-Real-Time Data in HDR,* Proceedings of the 17th International Teletraffic Congress (ITC-17), Salvador da Bahia, Brazil, September 2001.

[34] O. Shin, and K. Lee, *Packet Scheduling over a Shared Wireless Link for Heterogeneous Classes of Traffic,* Proceedings of IEEE International Conference on Communications (ICC 2004), vol. 1, pp 58-62, June 2004.

[35] A. Garcia, and I. Widjaja, *Communication Networks, Fundamental Concepts and Key Architectures,* Mc Graw Hill, 2004.

[36] L. Almajano, and J. Prez-Romero, *Packet Scheduling Algorithms for Interactive and Streaming Services under QoS Guarantee in a CDMA System,* Proceedings of IEEE Vehicular Technology Conference (VTC 2002), vol. 3, pp. 1657-61, Sept. 2002.

[37] E. Esteves, *On the Reverse Link Capacity of cdma2000 High Rate Packet Data Systems,* Proceedings of IEEE International Conference on Communications (ICC 2002), vol. 3, pp. 1823-28, April-May 2002.

[38] I. Lopez, P. J. Ameigeiras, J. Wigard, and P. Mogensen, *Downlink Radio Resource Management for IP Packet Services in UMTS,* Proceedings of IEEE Vehicular Technology Conference (VTC 2001), vol. 4, pp. 2387-91, May 2001.

[39] S. Malik, and D. Zeghlache, *Improving Throughput and Fairness on the Downlink Shared Channel in UMTS WCDMA Networks,* Proceedings of European Wireless 2002, Florence, Italy, February 2002.

[40] M. A. Haleem, and R. Chandramouli, *Adaptive Downlink Scheduling and Rate Selection: A Cross-Layer Design,* IEEE Journal on Selected Areas in Communications, vol. 23, pp. 1287-97, June 2005.

[41] H. Fattah, and C. Leung, *An Overview of Scheduling Algorithms in Wireless Multimedia Networks,"* IEEE Wireless Communications, pp. 76-83, October 2002.

[42] Jeongrok Yang, Insoo Koo, Yeongyoon Choi, and Kiseon Kim, *Dynamic Resource Allocation Scheme in Multi-Service CDMA Systems,* IEICI Transactions on Communications, vol. E87-B, pp. 2634-37, Sept. 2004.

[43] B. Al-Manthari, N. Naser, and H. Hassanein, *Fair and Efficient Channel Dependent Scheduling Algorithm for HSDPA System,* 2^{nd} International Symposium on Wireless Communication Systems (ISWCS2005), Siena, Italy, September 2005.

[44] F. Long, G. Feng, and C. Siew, *Channel States Dependent Fair Service: A New Packet Scheduling Algorithm for CDMA,* ELSEVIER Computer Networks, vol. 49, pp. 201-16, March 2005.

[45] Li Wang, Yu-Kwong Kwok, Wing-Cheong Lau, and V. K. N. Lau, *Channel Adaptive Fair Queuing for Scheduling Integrated Voice and Data services in Multicode CDMA Systems,* ELSEVIER Computer Communications, vol. 27, pp. 809-20, 2004.

[46] George Xylomenos, George C. Polyzos, Petri Mahonen, and Mika Saaranen, *TCP Performance Issues over Wireless Links,* IEEE Communications Magazine, vol. 39, no. 4, pp. 52-58, 2001.

[47] 3GPP2, *cdma2000 Standard for Spread Spectrum Systems,* Revision C, May 2002.

[48] 3GPP2, *cdma2000 Standard for Spread Spectrum Systems,* Revision D, Version 2.0, Sep. 2005.

[49] 3GPP2, *cdma2000 High Rate Data Packet Data Air Interface Specification,* Version 4.0, Oct. 2002.

[50] IEEE Standerds, *IEEE Standard for Local and Metropolitan Area Networks,* IEEE Std. 802.16e, 2005.

[51] Vijay K. Garg, *IS-95 CDMA and cdma 2000: Cellular/PCS Systems Implementation,* Prentice Hall PTR; first edition, 1999.

[52] 3GPP2, *cdma2000 Standard for Spread Spectrum Systems,* Revision 0, June 2000.

[53] 3GPP2, *cdma2000 Standard for Spread Spectrum Systems,* Revision A, June 2000.

[54] 3GPP2, *cdma2000 Standard for Spread Spectrum Systems,* Revision B, April 2002.

[55] D. Comstok, R. Vannithamby, S. Balasubramanian, L. Hsu, and M. Cheng, *Reverse Link High-Speed Packet Data Support in cdma2000 1xEV-DV Upper Layer Protocols,* IEEE Communications Magazine, vol. 43, no. 4, pp. 48-56, April 2005.

[56] 3GPP2, *cdma2000 Evaluation Methodology,* Dec. 2004.

[57] R. Thomas Derryberry, and Zhouyue Pi, *Reverse High-speed Packet Data Physical Layer Enhancements in cdma2000 1xEV-DV,* IEEE Communications Magazine, vol. 43, pp. 41-47, April 2005.

[58] A. C. K. Soong, Seong-Jun Oh, A.D. Damnjanovic, and Y.C. Yoon, *Forward High-speed Wireless Packet Data Service in IS-2000 1xEV-DV,* IEEE Communications Magazine, vol. 41, pp. 170-77, Aug. 2003.

[59] Soonyil Kwon, Kijun Kim, Youngwoo Yun, S. G. Kim, and B. K. Yi, *Power Controlled H-ARQ in cdma2000 1xEV-DV,* IEEE Communications Magazine, vol. 43, pp. 77-81, April 2005.

[60] Y. Kim, and B. Yi, *3G Wireless and cdma2000 1x Evolution in Korea,* IEEE Communications Magazine, vol. 43, pp. 36-40, April 2005.

[61] Yong-Hoon Choi, Jaesung Park, Beomjoon Kim, and M. A. Shayman, *A Framework for Elastic QoS Provisioning in the cdma2000 1xEV-DV Packet Core Network,* IEEE Communications Magazine, vol. 43, pp. 82-88, April 2005.

[62] Stawomir Pietrzyk, *OFDMA for Broadband Wireless Access.* Mobile communcation series, Boston, Artch House, 2006.

[63] J.G. Proakis, *Digital Communications,* McGraw Hill, 2001.

[64] Wimax Fourm, *Mobile WiMAX Part I: A Technical Overview and Performance Evaluation,* August 2006.

[65] Liang Peng, Tang Cailin, Ma Jie, Chang Yongyu, and Yang Dacheng, *Experimental Study on Traffic Model of Wireless Internet Services in CDMA Network,* Proceedings of IEEE Vehicular Technology Conference (VTC 2005), vol. 4, pp. 2137-41, May-June 2005.

[66] H. Gharavi, R. Wyatt-Millington, and F. Chin, *CDMA2000 Reverse-Link Simulation Model Design and Evaluation,* International Conference on Third Generation Wireless and Beyond, May-June 2001.

[67] Vineet Srivastava, *Cross-Layer Design: A Survey and the Road Ahead,* IEEE Communications Magazine, pp. 112-19, Dec. 2005.

[68] S. Shakkottai, T. S. Rappaport, and P. C. Karlsso, *Cross-Layer Design for Wirelss Networks,* IEEE Communications Magazine, pp. 74-80, Oct. 2003.

[69] George Xylomenos, and George C. Polyzos, *Quality of Service Issues in Multiservice Wireless Internet Links,* Computer Networks, vol. 37, no. 5, pp. 601-15, Nov. 2001.

[70] Lang Tong, Vidyut Naware, and Parvathinathan Venkitasubramaniam, *Signal Processing in Random Access,* IEEE Signal Processing Magazine, Special issue on signal processing in networking: An integrated approach, vol. 21, no. 5, Sept. 2004.

[71] Qingwen Liu, Shengli Zhou, Georgios B. Giannakis, *Cross-Layer Combining of Adaptive Modulation and Coding with Truncated ARQ over Wireless Links,* IEEE Transactions on Wireless Communications, vol. 3, no. 5, pp. 1746-55, Sept. 2004.

[72] M. Kemal Karakayali, D. Yates, and Leonid V. Razoumov, *Downlink Throughput Maximization in CDMA Wireless Networks,* IEEE Transactionson Wireless Communications, vol. 5, pp. 3492-500, Dec. 2006.

[73] Peshala Pahalawatta, Randall Berry, Thrasyvoulos Pappas, and Aggelos Katsaggelos, *Content-Aware Resource Allocation and Packet Scheduling for Video Transmission over Wireless Networks,* IEEE Journal on Selected Areas in Communications, vol. 25, pp. 749-59, May 2007.

[74] Jia Tang, and Xi Zhang, *Cross-Layer Resource Allocation Over Wireless Relay Networks for Quality of Service Provisioning,* IEEE Journal on Selected Areas in

Communications, vol. 25, pp. 645-56, May 2007.

[75] TIA, *An Over view of the Application of Code Division Multiple Access (CDMA) to Digital Cellular Systems and Personal Cellular Netwoks,* May 21, 1992.

[76] Mahmudur Rahman, Halim Yanikomeroglu, Mohamed H. Ahmed, and Samy Mahmoud, *Opportunistic Nonorthogonal Packet Scheduling in Fixed BroadbandWireless Access Networks,* EURASIP Journal on Wireless Communications and Networking, vol. 2006, pp. 1-11, 2006.

[77] Sofia N. Hertiana, Ida Wahidah, Rita Magdalena, and Muhammad Ary Murti, *PF Scheduler Algorithm and Open Loop Rate Control For Performance Improvement of CDMA 2000 1xEV-DO Network,* Wireless and Optical Communications Networks (WOCN '07), pp. 1-6, July 2007.

[78] Sem Borst, Ken Clarkson, John Graybeal, Harish Viswanathan, and Phil Whiting, *User-Level QoS and Traffic Engineering for 3G Wireless 1xEV-DO Systems,* Bell Labs Technical Journal, vol. 8(2), pp. 33-47, 2003.

[79] Harish Viswanathan, and Sayandev Mukherjee, *Throughput-Range Tradeoff of Wireless Mesh Backhaul Networks,* IEEE Journal on Selected Areas in Communications, vol. 24, no. 3, pp. 593-602, March 2006.

[80] Min Cao, Vivek Raghunathan, and P. R. Kumar, *A Tractable Algorithm for Fair and Efficient Uplink Scheduling of Multi-hop WiMax Mesh Networks.* Second

IEEE Workshop on Wireless Mesh Networks (WiMesh 2006), pp. 101-8, Sept. 2006.

[81] Min Cao,Wenchao Ma, Qian Zhang, Xiaodong Wang, and Wenwu Zhu, *Modelling and Performance Analysis of the Distributed Scheduler in IEEE 802.16 Mesh Mode,* The 6th ACM International Symposium on Mobile Ad-hoc Networking and Computing, pp. 78-89, 2005.

[82] Gautam Kulkarni, Sachin Adlakha, and Mani Srivastava, *Subcarrier Allocation and Bit Loading Algorithms for OFDMA-Based Wireless Networks,* IEEE Transactions on Mobile Computig, vol. 4, no. 6, pp. 652-62, Nov./Dec. 2005.

[83] Farid Dowla, *Handbook of RF and Wireless Technologies,* Elsevier, Oct 2003.

[84] Alberto Leon-Garcia, *Probability and Random Processes for Electrical Engineering,* Addison-Wesley, 2nd edition, July 1993.

[85] Pierre Bremaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer Verlag, 1999.

[86] R.C. Eberhart and Y.Shi, *Comparison between Genetic Algorithms and Particle Swarm Optimization,* Proc. IEEE International Conference on Evolutionary Comp., pp 611-16, 1998.

[87] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization,* Athena Scientific, 1997.

[88] *Special Issue on Particle Swarm Optimization,* IEEE Transactions on Evolutionary Computation, vol. 8, no. 3, June 2004.

[89] Shu Lin, and Daniel J. Costello, *Error Control Coding,* Pearson Prentice Hall, 2nd edition, 2004.

[90] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation, Space-Time Coding,* Wiley, 2004.

[91] M. Reza Solymani, Yingzi Gao, and U. Vilaipornsawai, *Turbo Coding for Satellite and Wireless Comunications,* Kluwer Academic Pubishers, 2002.

[92] Douglas N. Rowitch, and Laurence B. Milstein, *On the Performance of Hybrid FEC/ARQ Systems Using Rate Compatible Punctured Turbo (RCPT) Codes,* IEEE Transactions on Communications, vol. 48, pp. 948-59, June. 2000.

[93] D. Divsalar, *A Simple Tight Bound on Error Probability of Block Codes with Application to Turbo Codes,* TMO Progress Report 42-139, 1999.

[94] Sergio Benedetto, and Guido Montorsi, *Unveiling Turbo Codes: Some Results on Parallel Concatenated Coding Schemes,* IEEE Transactions on Information Theory, vol. 42, pp. 409-28, March 1996.

[95] Ioannis Chatzigeorgiou, Miguel R. D. Rodrigues, Ian J. Wassell, and Rolando Carrasco, *A Union Bound Approximation for Rapid Performance Evaluation of Punctured Turbo Codes,* Conference on Information Sciences and Systems, Baltimore, USA, March 2007.

[96] D.J.C. Mackay, *Fountain Codes,* IEE Transactions Proceedings on Communication, vol. 152, no. 6, pp. 1062-68, Dec. 2005.

[97] Amin Shokrollahi, *Raptor Codes,* IEEE Transactions on Information Theory, vol. 52, no. 6, pp. 2551-67, June 2006.

[98] M. Luby, *LT Codes,* Proc. of the 43^{rd} IEEE Symp. Foundations of Computer Science, FOCS 2002, pp. 271- 280, Nov. 2002.

[99] Dimitri Bertsekas, and Robert Gallager, *Data Networks,* Prentice Hall, New Jersey, 1987.

[100] Hiroshi Harada, and Ramjee Prasad, *Simulation and Software Radio for Mobile Communication,* Artech House, Boston, 2002.

[101] Ayda Basyouni, Ahmed Elhakeem, and Anjali Agarwal, *Efficient Resource Management for Packet Mode cdma2000,* European Transactions on Telecommunications (Accepted for publication).

[102] Ayda Basyouni, Anjali Agarwal, and Ahmed Elhakeem, *An Analytical Model for Reverse Data Channel Scheduling Techniques in cdma2000 1xEV-DO,* Journal of Wireless Communications and Mobile Computing (Accepted for publication).

[103] Ayda Basyouni, Anjali Agarwal, and Ahmed Elhakeem, *Performance Analysis of the cdma2000 Reverse Packet Data Channel,* International Journal of Communication Networks and Distributed Systems (Accepted for publication).

[104] Ayda Basyouni, Ahmed Elhakeem, and Anjali Agarwal, *A Dynamic Scheduling Scheme for the Reverse Packet Data Channel in cdma2000 1xEV-DV*, Proc. of the IEEE Canadian conference in Electrical and Computer Engineering, CCECE'2006, Ottawa, Canada, May 2006.

[105] Ayda Basyouni, Ahmed Elhakeem, and Anjali Agarwal, *On Rate Assignment Schemes for the Reverse Packet Data Channel in cdma2000 1xEV-DV*, Proc. of the 64^{th} IEEE Vehicular Technology conference (VTC'06), September 2006.

[106] Ayda Basyouni, Ahmed Elhakeem, and Anjali Agarwal, *Maximizing the Reverse Link Throughput for cdma2000 1xEV-DO Using Particle Swarm Optimization*, Proc. of the Seventh International Conferences on Wireless and Optical Communications (WOC 2007), Montreal, Canada, 2007.

[107] Ayda Basyouni, Anjali Agarwal, and Ahmed Elhakeem, *Performance Analysis of cdma2000 Reverse Packet Data Channel*, Proc. of the Second International Conference on Access Networks (ACCESSNETS 2007), Ottawa, Canada, 2007.

[108] Ayda Basyouni, Anjali Agarwal, and Ahmed Elhakeem, *A Slot Allocation Technique for WiMAX Backhaul Networks*, 24th Queen's Biennial Symposium on Communications, 2008, Kingston, Canada.

Appendix A

Bit-Error Bounds for cdma2000 1x Turbo Encoder

In what follows, based on the properties of the 1xEV turbo encoder [48], lower bounds on the probabilities of data sub-packets acceptance are derived.

A.1 The cdma2000 1xEV Turbo Encoder

As described in section 1.3.1, each encoder packet is fed into a turbo encoder with code rate $1/5$. The structure of cdma2000 1xEV turbo encoder is shown in Fig. A.1. It consists of two parallel concatenated constituent encoders, each of the encoders has a transfer function which is given by

$$\mathbf{G}(D) = [1 \; \frac{n_1(D)}{d(D)} \; \frac{n_2(D)}{d(D)}], \tag{A.1}$$

where

$$\begin{aligned} d(D) &= 1 + D^2 + D^3, \\ n_1(d) &= 1 + D + D^3, \\ n_2(D) &= 1 + D + D^2 + D^3. \end{aligned}$$

Each encoded packet is divided into three sub-packets. The first sub-packet has an encoding rate of $1/2$. As the second redundancy sub-packet is transmitted, the combined received data packets will have a higher encoding rate of $1/3$.

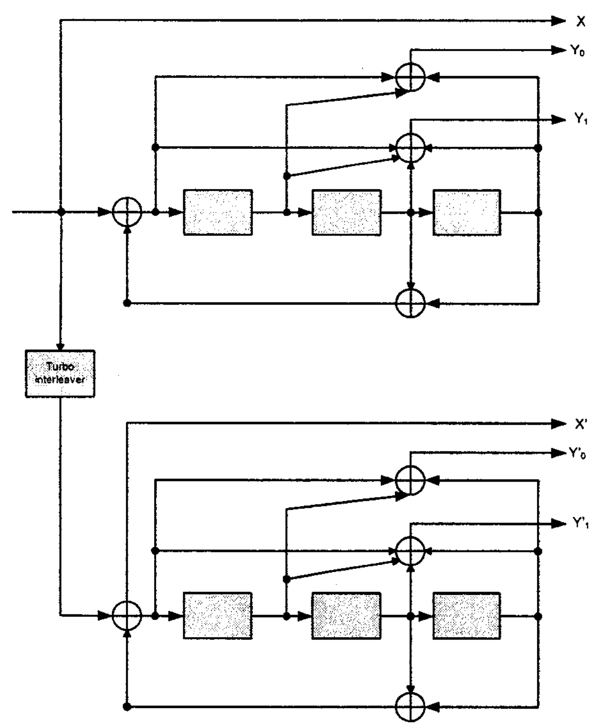


Figure A.1: cdma2000 1X turbo encoder

output	$R_c = 1/2$		$R_c = 1/3$	$R_c = 1/5$
	First codeword	Second codeword		
x	1	1	1	1
y_0	1	0	1	1
y_1	0	0	0	1
x'	0	0	0	0
y'_0	0	1	1	1
y'_1	0	0	0	1

Table A.1: Output punctured code for cdma2000 1X encoder

As the third sub-packet is transmitted, it reaches the original encoding rate of 1/5. As shown in Table A.1 the total codeword with encoding rate 1/5 has a codeword structure $\{x, y_0, y_1, y'_0, y'_1\}$. The first sub-packet with code rate $R_c = 1/2$ is considered to be a punctured output of the encoder shown in Fig. A.1 with codeword structure $\{x, y_0\}\{x, y'_0\}$. The first and second sub-packets combined are treated as a punctured output of the encoder in Fig. A.1 with codewords structure $\{x, y_0, y'_0\}$ [48].

A.2 BER Upper bound

In what follows, we briefly review the procedure for determining upper bounds for the BER of parallel concatenated constituent codes (PCCCs) as outlined in [89], [94].

For a PCCCs with large interleaver of size L and $(n, 1, v)$ systematic feedback constituent encoder, an approximate BER upper bound is given by [89], [94]

$$P_b(E) \approx \sum_{w_{min} \leq w \leq L} \frac{w}{L} W^w A_w^{PC}(Z)|_{W=Z=e^{-R_c E_b/N_0}}, \tag{A.2}$$

where R_c is the code rate, $\frac{E_b}{N_o}$ is the SNR per information bit, w is the codeword weight, w_{min} is the minimum codeword weight, and $A_w^{PC}(Z)$ is the codeword conditional enumerating function for parallel concatenated turbo encoder which is given by

$$A_w^{PC}(Z) = \frac{A_w^{C_1}(Z) A_w^{C_2}(Z)}{\binom{L}{w}}, \tag{A.3}$$

where $A_w^{C_1}(Z)$ and $A_w^{C_2}(Z)$ are the codeword conditional enumerating function for the first and second constituent encoders respectively.

Let h denote the number of error events. For large L, we can approximate the number of codewords containing h error events by $\binom{L}{h}$. Therefor $A_w(Z)$ can be defined as

$$A_w(Z) \approx \sum_{1 \leq h \leq h_{max}} \binom{L}{h} A_w^{(h)}(Z), \tag{A.4}$$

where $A_w^{(h)}(Z)$ is the h-error enumerator for input weight w, and h_{max} is the maximum number of error events. Substituting into Eq. (A.3), we get

$$A_w^{PC}(Z) \approx \sum_{1\le h_1 \le h_{max}} \sum_{1\le h_2\le h_{max}} \frac{\binom{L}{h_1}\binom{L}{h_2}}{\binom{L}{w}} A_w^{(h_1)}(Z) A_w^{(h_2)}(Z). \tag{A.5}$$

For $L >> h$, the approximation $\binom{L}{h} \approx \frac{L^h}{h!}$ can be used. Hence Eq. (A.5) can be rewritten as

$$A_w^{PC}(Z) \approx \sum_{1\le h_1 \le h_{max}} \sum_{1\le h_2\le h_{max}} \frac{w!}{h_1! h_2!} L^{(h_1+h_2-w)} A_w^{(h_1)}(Z) A_w^{(h_2)}(Z). \tag{A.6}$$

A further approximation can be achieved by considering only the most significant terms in the double summation. Let $h_1 = h_2 \approx h_{max}$. therefor

$$A_w^{PC}(Z) \approx \frac{w!}{h_{max}!^2} L^{(2h_{max}-w)} [A_w^{(h_{max})}(Z)]^2 \tag{A.7}$$

By substituting into Eq. (A.2) we get

$$P_b(E) \approx \sum_{w_{min}\le w \le L} w \frac{w!}{h_{max}!^2} L^{(2h_{max}-w-1)} W^w [A_w^{(h_{max})}(Z)]^2 |_{W=Z=e^{-R_c E_b/N_0}}. \tag{A.8}$$

For $(n, 1, v)$ systematic feedback constituent encoders, $w_{min} = 2$, and $h_{max} = \lfloor \frac{w}{2} \rfloor$. Therefor the term $L^{2h_{max}-w-1} = L^{-2}$ for odd w, and $L^{2h_{max}-w-1} = L^{-1}$ for even w. Thus for large L the term L^{-2} can be neglected, and only the terms of the type $A_{2w}^{(w)}(Z)$ are considered. Further, for any $(n, 1, v)$ systematic feedback constituent encoder $A_{2w}^{(w)}(Z) = [A_2^{(1)}(Z)]^w$. By substituting into (A.8) we get

$$P_b(E) \approx \sum_{1 \le w \le \lfloor \frac{L}{2} \rfloor} 2w \binom{2w}{w} L^{-1} W^{2w} [A_2^{(1)}(Z)]^{2w} |_{W=Z=e^{-R_c E_b/N_0}} \tag{A.9}$$

where, $A_2^{(1)}$ is the parity weight enumerator for single error event with input weight 2.

From the above cdma2000 1xEV encoder description, it is clear that a codeword with code rate $R_c = 1/3$ and $1/5$ can be generated using constituent encoder with a transfer functions $\mathbf{G}(D) = [1\ n_1(D)/d(D)\ n_2(D)/d(D)]$, and $\mathbf{G}(D) = [1\ n_1(D)/d(D)]$ respectively. Also one can recognize that for code rates $R_c = 1/3$ and $1/5$ the parity outputs from both constituent encoders are the same, and hence Eq. (A.9) can be applied. As for code rate $R_c = 1/2$ half of the codewords are generated only by the first convolutional encoder $\{x, y_0\}$, and the other half is generated by puncturing the output of the $R_c = 1/3$ turbo encoder $\{x, y_0'\}$. Applying Eq. (A.5), we get

$$A_w^{PC}(Z) \approx \sum_{1 \le h_1 \le h_{max}} \sum_{1 \le h_2 \le h_{max}} \frac{\binom{L}{h_1}\binom{L}{h_2}}{\binom{L}{w}} (A_w^{(h_1)}(Z) A_w^{(h_2)}(Z))_{\{x,y_0\}\{x,y_0'\}}. \tag{A.10}$$

By inspecting the turbo encoder (see Fig. A.1), we can see that for the codeword $\{x, y_0\}$, there is no output from the second convolutional encoder. As for the codeword $\{x, y_0'\}$ the output from the turbo encoder is the same as the output of first convolutional encoder. Note that the first part of Eq. (A.10) already takes care of the effect of the interleaver. Thus

$$\begin{aligned} (A_w^{(h_1)}(Z)A_w^{(h_2)}(Z))_{\{x,y_0\}} &= (A_w^{(h_1)}(Z)A_w^{(h_2)}(Z))_{\{x,y_0'\}} \\ &= (A_w^{h_1}(Z)), \end{aligned} \tag{A.11}$$

and Eq. (A.10) can be rewritten as

$$A_w^{PC}(Z) \approx \sum_{1 \leq h_1 \leq h_{max}} \sum_{1 \leq h_2 \leq h_{max}} \frac{\binom{L}{h_1}\binom{L}{h_2}}{\binom{L}{w}} [A_w^{(h_1)}(Z)] \tag{A.12}$$

Following the steps as in Eqs. A.5-A.7, one can show that

$$A_w^{PC}(Z) \approx \frac{w!}{(h_{max}!)^2} L^{(2h_{max}-w)} [A_w^{(h_{max})}] \tag{A.13}$$

Thus the BER for code rate $R_c = 1/2$ is given by

$$P_b(E) \approx \sum_{1 \leq w \leq \lfloor \frac{L}{2} \rfloor} 2w \binom{2w}{w} L^{-1} W^{2w} [A_2^{(1)}(Z)]^w |_{W=Z=e^{-R_c E_b/N_0}} \tag{A.14}$$

where, $A_2^{(1)}$ is the parity weight enumerator for single error event with input weight 2 for code rate $R_c = 1/3$.

In what follows, we determine the weight enumerator functions of the cdma2000 1xEV encoder with different code rates. Table A.2 shows the states of cdma2000 1xEV encoder and their corresponding input/output bits for constituent encoder. Figs. A.2 and A.3 show the state diagram for cdma2000 1xEV constituent encoder with $R_c = 1/5$, and $R_c = 1/3$ respectively.

Form Table A.2 and Figs. A.2 and A.3 one can notice that the shortest path for single error event with input weight 2 is given by: $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_5 \rightarrow S_3 \rightarrow$

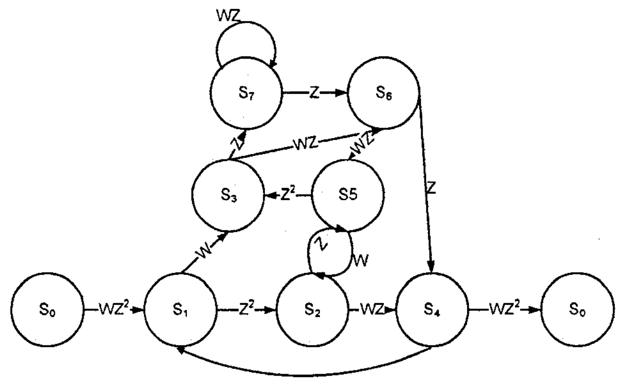


Figure A.2: State diagram for cdma2000 1xEV encoder with $R_c = 1/5$

$S_7 \rightarrow S_6 \rightarrow S_4 \rightarrow S_0$ with one cycle around the loop ($S_4 \rightarrow S_1 \rightarrow S_2 \rightarrow S_5 \rightarrow S_3 \rightarrow S_7 \rightarrow S_6 \rightarrow S_4$). Therefor,

$$A_2^{(1)}(Z) = \begin{cases} Z^{12} + Z^{20} + Z^{28} \ldots\ldots = \frac{Z^{12}}{1-Z^8}, & R_c = 1/5 \\ Z^6 + Z^{10} + Z^{14} \ldots\ldots = \frac{Z^6}{1-Z^4}, & R_c = 1/3 \end{cases}$$

A.3 Probability of Packet Acceptance

In this section we derive an approximate expression for the probability of packet acceptance that is a function of coding rate R_c and E_b/N_0.

Let $ps_{1,k}$ and $ps_{2,k}$ be the probabilities that the packet is accepted after the first or second transmissions respectively. Let b_k denote the number of bits per packet associated with data rate R_k. Hence,

		Input/Output $R_c = 1/5$	Input/Output $R_c = 1/3$	next state
S_0	000	0/000	0/00	000 (S_0)
S_0	000	1/111	1/11	100 (S_1)
S_1	100	0/011	0/01	010 (S_2)
S_1	100	1/100	1/10	110 (S_3)
S_2	010	0/010	0/01	101 (S_5)
S_2	010	1/101	1/10	001 (S_4)
S_3	110	0/001	0/00	111 (S_7)
S_3	110	1/110	1/11	011 (S_6)
S_4	001	0/000	0/00	100 (S_1)
S_4	001	1/111	1/11	000 (S_0)
S_5	101	0/011	0/01	110 (S_3)
S_5	101	1/100	1/10	010 (S_2)
S_6	101	0/010	0/01	001 (S_4)
S_6	101	1/101	1/10	101 (S_5)
S_7	111	0/001	0/00	011 (S_6)
S_7	111	1/110	1/11	111 (S_7)

Table A.2: States for the cdma2000 1X encoder

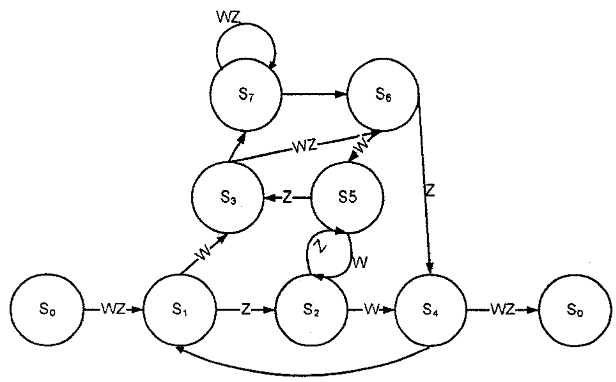


Figure A.3: State diagram for cdma2000 1xEV encoder with $R_c = 1/3$

$$\begin{aligned} ps_{i,k} &= (1-(P_b(E))_{i,k})^{b_k} \\ &= 1 - \binom{b_k}{1}(P_b(E))_{i,k} + \binom{b_k}{2}(P_b(E))^2_{i,k} - \cdots + (P_b(E))^{b_k}_{i,k} \end{aligned} \tag{A.15}$$

where $(P_b(E))_{i,k})$ is the probability of BER (Eqs. A.9 and A.14) associated with the i^{th} transmission ($i = 1, 2$)and $L \approx b_k$. By noting that $P_b(E)_{i,k} \leq 1$, we consider only the first few dominant terms in Eq. A.15. Thus

$$ps_{i,k} \approx 1 - \binom{b_k}{1}(P_b(E))_{i,k} + \binom{b_k}{2}(P_b(E))^2_{i,k} \tag{A.16}$$

By applying the binomial coefficient approximation used in Eq. (A.6), we get

$$ps_{i,k} \approx \quad 1 - b_k(P_b(E))_{i,k} + \tfrac{b_k^2}{2!}(P_b(E))^2_{i,k} \tag{A.17}$$

Eqs (A.9) and (A.14), $(P_b(E))_{i,k}$ can be rewritten in the following form:

$$(P_b(E))_{i,k} \approx \tfrac{c_i}{L_k}, \quad i = 1, 2, 3. \tag{A.18}$$

where, $L_k \approx b_k$ is the interleaver size corresponds to data rate R_k and

$$c_1 = \sum_{1 \le w \le 10} 2w \binom{2w}{w} W^{2w} [A_2^{(1)}(Z)]^{w} |_{W=Z=e^{-1/2E_b/N_0}},$$

$$c_2 = \sum_{1 \le w \le 10} 2w \binom{2w}{w} W^{2w} [A_2^{(1)}(Z)]^{2w} |_{W=Z=e^{-1/3E_b/N_0}},$$

and

$$c_3 = \sum_{1 \le w \le 10} 2w \binom{2w}{w} W^{2w} [A_2^{(1)}(Z)]^{2w} |_{W=Z=e^{-1/5E_b/N_0}},$$

By substituting into Eq. A.17, we get

$$ps_{i,k} \approx 1 - c_i + \tfrac{1}{2!}(c_i)^2 \stackrel{\text{def}}{=} ps_i, \quad i = 1, 2, 3. \tag{A.19}$$

which is function of only the coding rate R_c and E_b/N_0.

Based on the above analysis, Fig. A.4 shows how the lower bounds for the probabilities of sub-packets acceptance vary with E_b/N_0. It is clear that, even for low E_b/N_0, the probability of sub-packets acceptance is well above 90%, and for $E_b/N_0 \ge 4dB$, the data packet is practically accepted from the first transmission.

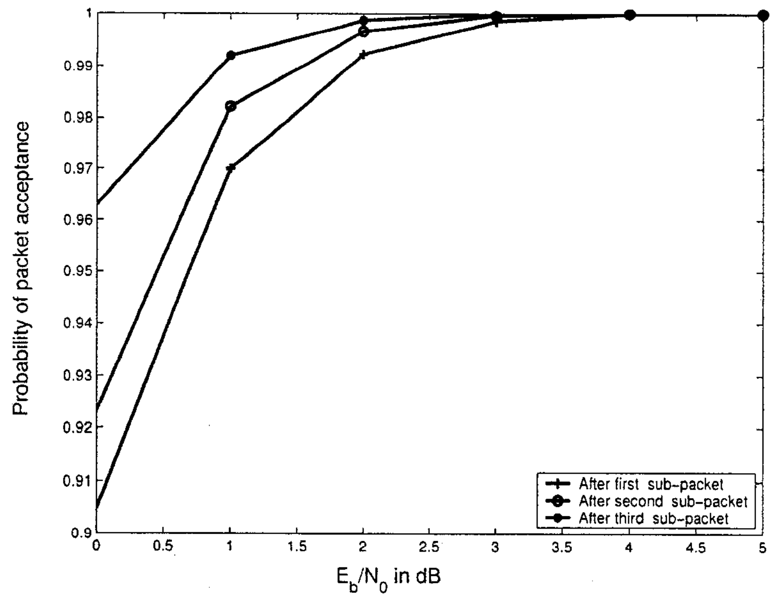


Figure A.4: Lower bounds for the probabilities of packet acceptance

A.4 Eb/No

In this section we show the relation between E_b/N_o and the required signal to noise and interference ratio for the pilot channel τ.

Let P_T be the total power received at the BS. Then $\frac{E_b}{N_o}$ is given by

$$\frac{E_b}{N_o} = \frac{P_i}{(P_T - P_i)} \times \frac{W}{R_i} \tag{A.20}$$

where P_i and R_i are the received power at the BS, and the data rate assigned to the i^{th} mobile respectively. W is the available bandwidth.

Let $\overline{P_i}$ be the pilot channel power for the i^{th} mobile. Define the traffic to pilot ratio for the i^{th} mobile as $TPR_i = P_i/\overline{P_i}$. Denote the targeted signal to noise and interference ratio for the pilot channel as $\tau = \overline{P_i}/(P_T - P_i)$, where $(P_T - P_i)$ is the total interference and the noise power for the i^{th} mobile. Therefore Eq. (A.20), can be rewritten as,

$$\begin{aligned}\frac{E_b}{N_o} &= \frac{P_i}{(P_T - P_i)} \times \frac{\overline{P_i}}{\overline{P_i}} \times \frac{W}{R_i} \\ &= \frac{TPR_i}{\frac{1}{\tau}} \times \frac{W}{R_i} \\ &= \tau \times TPR_i \times \frac{W}{R_i}\end{aligned} \tag{A.21}$$

The traffic to pilot ratio TPR_k corresponding to data rate R_k, is given by Eq. 2.4. Hence, the E_b/N_o corresponding to data rate R_k is given by,

$$\left(\frac{E_b}{N_o}\right)_k = \tau \times TPR_k \times \frac{W}{R_k} \tag{A.22}$$

One can notice that $(\frac{E_b}{N_o})_k$ depends on the value of τ as TPR_k, W, and R_k have prespecified values (see Section 1.3.1).