On Optimization of Wireless Multimedia Broadcasting

by

Çağdaş Atıcı

A Dissertation Submitted to the
Graduate School of Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy

in

Electrical & Electronics Engineering

Koç University

January, 2009

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a doctoral dissertation by

Çağdaş Atıcı

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

_____

Assoc. Prof. M. Oğuz Sunay

_____

Prof. A. Murat Tekalp

_____

Prof. Reha Civanlar

_____

Assist. Prof. Özgür Gürbüz

_____

Assist. Prof. Özgür Erçetin

Date: _____

There is no thing endowed with life—from man, who is enslaving the elements, to the nimblest creature—in all this world that does not sway in its turn. Whenever action is born from force, though it be infinitesimal, the cosmic balance is upset and the universal motion results.

Nikola Tesla, 1925.

To my family,

Hakan, Zerrin and Efehan

ABSTRACT

This dissertation is concerned with the development of a cross-layer framework for wireless broadcasting systems. Both cellular and ad hoc networks are considered. It has been well-established that cross-layer designs provide significant gains in the performance of point-to-point wireless communication systems. However, the point-to-multipoint nature of the broadcast system has so far eluded most researchers into thinking that a cross-layer solution is not feasible. This dissertation aims to dispell this reasoning. Indeed, throughout the thesis, we show that practical cross-layer solutions for wireless broadcast systems exist and they provide significant performance gains.

First, we consider a CDMA-based cellular broadcast system where the video content to be broadcast is SVC encoded prior to transmission. For this scenario, we develop a solution where we aim to simultaneously maximize the number of users receiving the video broadcasting service and the average successfully decoded video data rate. The proposed broadcast system aims to optimally divide the code space amongst the multiple layers of the SVC content using multiple-objective optimization. This framework finds the best compromise between the objectives of maximizing the average decodable video data rate and minimizing the basic quality video outage probability. Simulations conducted for the ITU Pedestrian A and Vehicular B channels show that high data rates with low outages are possible when a 3G system such as 1xEV-DO is used for video broadcasting, however, it may not be desirable to use scalable video coding for this purpose.

Second, we consider a cellular broadcast system where the video encoding is conducted using the state-of-the-art video codec, H.264/AVC. In this scenario, we optimize the number of users receiving the broadcasting service and the average video quality of the received stream. To solve this problem optimally, we adjust the physical and the application layer parameters in a cross-layer fashion while utilizing the characteristics of the video that is being transmitted as well as the wireless channel. We propose a content-adaptive, multi-objective optimized, cross-layer video broadcasting framework for a wireless system capable

of supporting a multitude of transmission data rates using the H.264/AVC. The optimization problem aims to find the H.264/AVC as well as the physical layer system parameters jointly to reach the optimal compromise between the objectives of maximizing the average received video PSNR and minimizing the video broadcast service outage probability. Simulations conducted for the ITU Pedestrian A and Vehicular B channels show that significant gains in system performance can be achieved for video broadcasting using such a cross-layer design is used.

Third, we consider a wireless ad hoc system where any of the participating nodes may cooperate in the broadcast initiated by the source node. In such a scenario, significant difficulties exist in terms of the selection of participating relay nodes if the aim is to achieve the optimal capacity for a given total energy constraint. For a given number of active users and the CSI between each user, finding the optimal broadcast capacity achieving subset of cooperating nodes satisfying the energy constraint is an NP-complete problem. To alleviate the problem, we consider the case of a circular ad hoc network where there are infinitely many nodes. In this scenario, we find critical points that determine the optimal broadcast capacity for a given total energy constraint. Following this analysis, we propose two algorithms, one distributed and one centralized, both based on heuristics computed by the theoretical model. Not only the proposed methods present a practical solution to the optimal broadcast capacity achieving relay selection problem but they also provide a significant increase in the overall system performance. We discuss the behaviors of the broadcast capacity for the two proposed models under a variety of system settings based on realistic assumptions. The results show that using a randomized and distributed relay selection in broadcasting yields significant gains in the system performance, even in the worst case scenario.

ÖZETÇE

Bu tez raporu, hem hücresel hem de tasarsız ağlardaki kablosuz yayın sistemleri için katmanlar arası taslakların geliştirilmesiyle ilgilidir. Katmanlar arası tasarımın, noktadan noktaya kablosuz iletişim sistemlerinde önemli kazanımlar sağladığı zaten bilinmektedir. Ancak, yayın sistemlerinin doğasında bulunan noktadan çoklu noktaya iletim, birçok araştırmacının bu konuda katmanlar arası bir çözüm uygulanabileceğini düşünmesini engellemiştir. Bu tez raporu, bu düşünceyi değiştirmeyi amaçlamaktadır. Tez raporunda, kablosuz yayın sistemleri için pratik katmanlar arası çözümlerin var olduğunu ve önemli performans kazanımları sağladıklarını göstermekteyiz.

İlk olarak, yayınlanacak video içeriğinin iletimden önce SVC ile kodlanmış olduğu bir CDMA tabanlı hücresel yayın sistemini ele almaktayız. Bu senaryoda, video yayını servisini alan kullanıcı sayısını ve başarıyla kodçözülmüş ortalama video veri hızını birlikte eniyilemeyi amaçlayan bir çözüm önermekteyiz. Önerilen yayın sistemi çoklu amaç eniyilemesi kullanarak kod uzayını SVC içeriğinin katmanları arasında en iyi şekilde bölmeyi amaçlamaktadır. Bu sistem kodçözülmüş ortalama video veri hızını enbüyülten ve temel video kalitesindeki kesinti olasılığını enküçülten eniyi ödünleşim noktasını bulmaktadır. ITU Yaya A ve Taşıt B kanalları için yapılan benzetimler, 1xEV-DO gibi bir 3G sistemi video yayını için kullanıldığında, yüksek veri hızlarında düşük hizmet kesinti olasığı ile video yayını yapılabildiğini ancak ölçeklenebilir video kodlayıcısının bu amaca uygun olmayabileceğini göstermektedir.

İkinci olarak, en son teknoloji video kodlayıcısı H.264/AVC ile video kodlamasının yapıldığı bir hücresel yayın sistemini ele almaktayız. Burada, video yayını servisini alan kullanıcı sayısı ile alınan ortalama video kalitesi eniyilenmektedir. Bu problemi en iyi şekilde çözebilmek için, fiziksel ve uygulama katmanlarındaki değişkenler, katmanlar arası şekilde, hem gönderilen videonun özelliklerine ve hem de kablosuz kanala göre ayarlanmaktadır. H.264/AVC kodlayıcısı kullanılarak birden çok iletim hızına destek veren bir kablosuz sistem için, içerik uyarlamalı, çoklu amaç eniyilenmiş katmanlar arası video yayını şeması

önermekteyiz. Eniyileme problemi, alınan ortalama video PSNR değerini enbüyülten ve video yayını servisindeki kesinti olasılığını enküçülten en iyi ödünleşim noktasına ulaşmak için, hem H.264/AVC kodlayıcısındaki hem de fiziksel katmandaki sistem değişkenlerini bulmayı amaçlamaktadır. ITU Yaya A ve Taşıt B kanalları için yapılan benzetimler, katmanlar arası tasarım ile sistem başarımında önemli kazanımlar elde edildiğini göstermektedir.

Üçüncü olarak, kaynak kullanıcı tarafından başlatılan ve katılan kullanıcıların işbirliğinde bulunabildiği bir kablosuz tasarsız tümegönderim sistemini ele almaktayız. Bu senaryo, verilen toplam enerji kısıtı ile en iyi kapasite değerine ulaşmak için hangi kullanıcıların röle olarak seçileceği gibi önemli sorunları içermektedir. Belirli bir kullanıcı sayısı ve bu kullanıcılar arasındaki kanal durum bilgisi verildiğinde, enerji kısıtını sağlayan ve en iyi tümegönderim kapasitesini sağlayan işbirlikçi kullanıcı kümesinin bulunması polinom zamanda çözülemeyecek kadar zor bir problemdir. Bu problemi kolaylaştırmak için, sınırsız sayıda kullanıcının olduğu dairesel tasarsız bir ağ üzerinde çalışılmaktadır. Bu durumda, verilen enerji kısıtı için en iyi tümegönderim kapasitesini belirleyen kritik noktalar bulunmaktadır. Bunun ardından, kuramsal model ile hesaplanan buluşsal değerlere dayalı iki algoritma, bir dağıtımlı ve bir merkezi, sunulmaktadır. Önerilen yöntem, sadece "en iyi tümegönderim kapasitesini elde eden röle seçim problemi"ne pratik bir çözüm önermekle kalmayıp, sistem başarımında da önemli artışlar sağlamaktadır. Her iki önerilen yöntemin tümegönderim kapasitesinin, gerçekçi kabullere bağlı farklı sistem ayarlarındaki davranışları incelenmektedir. Sonuçlar, tümegönderim için rasgele ve dağıtımlı röle seçimi yönteminin kullanılmasıyla, en kötü durumda bile sistem başarımında önemli artışlar elde edildiğini göstermektedir.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and profound respect to Assoc. Prof. M. Oğuz Sunay for his reliable guidance, inspiration, encouragement and valuable support over the past years. I have benefited tremendously from his wisdom and advices he has given me not only during my Ph.D. study but also during my undergradute years. I want to thank him for his patience, consistent support and belief in me. Not only he has helped me to reach one of the most important milestones in my academic career, but also he prepared me to reach many more.

I am also grateful to members of my thesis committee for their valuable comments. I would like to thank Prof. A. Murat Tekalp, Prof. Reha Civanlar, Assist. Prof. Özgür Gürbüz and Assist. Prof. Özgür Erçetin for the critical reading of this thesis and serving on my advisory committee.

The financial support of the Turkish National Science Foundation (TÜBİTAK) is also sincerely acknowledged.

Last but not the least, I want to express my sincere thanks to my family for their love and support. They have always inspired me to pursue an academic career. I thank them for their unflagging encouragement and understanding that helped me in hard days. I want to thank my father, Hakan, for guiding me patiently and I want to thank my mother, Zerrin, for listening me and providing me an endless support. Also, I want to thank my brother, Efehan, for his love, respect and sincereness. This dissertation is lovingly dedicated to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

1xEV-DO	1xEvolution-Data Optimized
AMC	Adaptive Modulation and Coding
ARQ	Automatic Repeat reQuest
AVC	Advanced Video Coding
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
CLO	Cross-Layer Optimization
CSI	Channel State Information
DF	Decode-and-Forward
DVB-H	Digital Video Broadcasting-Handheld
FEC	Forward Error Correction
GOP	Group of Pictures
HARQ	Hybrid Automatic Repeat reQuest
HDR	High Data Rate
IMT	International Mobile Telecommunications
IP	Internet Protocol
ITU	International Telecommunication Union
MBMS	Multimedia Broadcast Multicast Service
MOO	Multiple-Objective Optimization
MRC	Maximal-Ratio-Combining
NAL	Network Abstraction Layer
OSI	Open Systems Interconnection
PER	Packet Error Rate
PSNR	Peak Signal-to-Noise Ratio
QAM	Quadrature Amplitude Modulation

QP	Quantization Parameter
QPSK	Quadrature Phase Shift Keying
RS	Random Selection
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
ST	Selection by Triangularization
SVC	Scalable Video Coding
TDM	Time Division Multiplexing
UMTS	Universal Mobile Telecommunications System
VCL	Video Coding Layer

Chapter 1

INTRODUCTION

Today, wireless communications is a hot and still growing topic in the communications field. Although its growth is proliferated by the introduction of cellular phones, wireless communications includes many application fields other than mobile communications, such as wireless sensor networks, wireless ad hoc networks, wireless computer peripherals, satellite television, global positioning system, just to name a few. The basic idea behind all of these applications is the information transmission between a source and a destination without the use of electrical conductors. Over the last decades, many researchers put their effort on improving the rate of this information transmission between the source-destination pair by developing newer signal processing techniques, protocols, system designs, hardwares and algorithms.

The tremendous growth in research and development as well as the current marketing demands facilitate the merger of three fundamental technologies, namely, the Internet, TV and wireless communications. Such a merger will not only enable interactive multimedia delivery to end users via the use of the IP as a common transmission protocol, but also will provide mobile users to access multimedia content from many different venues, such as broadcast stations as well as the Internet. Many standards and proposals exist to enable multimedia broadcasting over wireless systems. While some of these systems require an entirely new physical layer, others have been designed to make use of existing wireless systems. It has been well-established that cross-layer designs provide significant gains in the performance of point-to-point wireless communications systems. However, the point-to-multipoint nature of the broadcast system has so far eluded most researchers into thinking that a cross-layer solution is not possible. Therefore, cross-layer adaptation of system resources for wireless broadcasting has not been well-studied in the literature so far. The main goal of this dissertation is to investigate the feasibility and the impact of a cross-layer

framework for wireless broadcasting of multimedia content. The work presented herein provides solutions for improving the wireless broadcast transmission in both cellular and ad-hoc networks.

The basic form of transmission includes a source-destination pair and it is called unicast transmission. Although much of the work in the literature concentrated on improving the performance of unicast transmission, researchers started to investigate the performance of networks with more than one source and/or destination in recent years. The information transmission in a network can be in the form of dissemination from a single source (or multi sources) to a number of destinations. This type of transmission is called multicast if destination set includes a group of users and it is called broadcast if each user in the network is a destination. For a broadcast system, the performance of the network is based on each source-destination pair (or only the worst pair, depending on the performance criteria) as there are more than one channel in the system. Thus, system resources can not be allocated by using an adaptation to one channel's condition as it is done in a unicast scenario. As a result, there is a need for optimal allocation of system resources according to the form of transmission used in the network.

Multimedia became an indispensable part of the transmitted data in today's networks, following the rapid developments in image and video processing in the past decades. Today, people are video conferencing over the Internet, video calling on cell phones and watching video clips on the web. Less than 10 years ago, people are using text-based chat engines instead of video chat via webcameras, using SMS (Short Message Service) instead of video call or MMS (Multimedia Message Service) and we were exploring the Internet only by reading text instead of watching a video clip or a full-length movie. At the same time, multimedia transmission gained enormous interest in the networking literature. A lot of work is done on developing new system designs and algorithms for multimedia transmission over both wired and wireless networks. In a wired network, the adaptation of the system resources for a multimedia content includes a lot of degrees of freedom due to its slow varying nature and almost always available duplex communication between source-destination pair. However, multimedia transmission poses significant difficulties in a wireless environment due to the presence of fading, interference and noise. On the other hand, video quality at the receiver side can be a degraded version of the original video when transmission errors

are considered. The quantity of this degradation varies according to which part of the video is lost during the transmission. For instance, assume a full frame is lost during the transmission. If there is a high correlation between the previous frame and the lost frame (e.g., slowly varying content and/or low details), then video decoder can use a simple error concealment to recover this error. However if there is a low correlation (e.g., fast motion and/or high details), then some artifacts are unavoidable at the received video regardless of the performance of the error concealment technique used at the decoder. So, the loss in video quality is dependent on the video content. Therefore, a successful multimedia transmission can be achieved by adjusting system parameters according to the properties of both the transmission medium and the multimedia content, in an optimal way.

When wireless broadcasting includes multimedia transmission, all of the above mentioned factors affect the system performance in different ways. Thus, an optimization of system resources and parameters is needed. This thesis focuses on wireless broadcasting of multimedia and optimization of system performance. It is constituted of three parts, each of which includes an optimization of a wireless network for a predefined set of objectives. In the following section, the motivation for optimization of wireless multimedia broadcasting is given. Then, we summarize the contributions of this thesis and give the thesis outline.

1.1 Motivation

With the rapid dissemination of wireless networks and demand for a vast number of applications, an increasing number of users are expected to enjoy a variety of services provided by the future wireless systems. Considering the share of the multimedia transmission among those services, wireless system solutions for mobile multimedia services are of major importance [28]. Among those services, there are a variety of solutions proposed for wireless video broadcasting, which are reviewed in Section 2.2. On the other hand, video coding techniques are developed and performances of video codecs are increased substantially. Today, H.264/AVC is the state-of-the-art video encoder and it presents several useful properties for wireless transmission of multimedia [118]. H.264/AVC, which is explained in Section 2.5.1, realizes the existing idea of scalable video coding due to its effective rate-distortion performance. Although [135] examines the scalable coding without using H.264/AVC over wireless networks, following this significant development in video processing, it is worth

to investigate the performance of scalability extension of H.264/AVC in a wireless environment [105]. But improving multimedia transmission over wireless networks needs not only an optimal allocation of system resources [43] and but also an optimal setting of video coding parameters [120]. When these two ideas, wireless multimedia transmission and scalable video coding, are combined in a broadcasting service, a question exists whether it is beneficial to use scalability extension of the H.264/AVC in a broadcasting application over existing 3G wireless systems. The first part of this thesis answers this question by analyzing the feasibility of this scenario.

Today, widely used wireless systems are based on the layered architecture, which is one of the major causes of the expeditious growth of the Internet. Traditionally, in a layered stack design, each OSI layer is allowed to communicate with the layer above or the layer below. As an evolution, CLO [110], [115] breaks the rigid and strict structure of the traditional design by allowing interactions between layers. At this point, CLO seems to be an alternative technique for improving the multimedia transmission performance of the existing wireless networks instead of developing a revolutionary design from scratch. Although traditional approach is to optimize the performance of each layer in isolation from the other layers, this approach presents a suboptimal solution when compared with a joint optimization of layers. Another cause of this benefit comes from the utilization of the feedback information from the receiver side. However, when a broadcasting service is considered, there is a multitude of feedback information from each user with a variety of channel conditions. Thus, the work on cross-layer design for wireless video broadcasting is scarce in the literature. To fill this gap, the second part of this thesis presents a framework for cross-layer optimization in a wireless video broadcasting scenario.

As wireless ad hoc networks proliferate, the problem of broadcasting a common information from a source node to all the other nodes in the network gains particular importance. In a wireless ad hoc network, one of the important objectives of broadcasting is to deliver all the information to all nodes successfully in the least amount of time. To achieve this goal, some of the nodes can be used as relays so that the capacity is improved substantially by increasing the total energy consumption in the transmission. Besides, allowing the existence of relay nodes creates an interference since wireless transmission using an omnidirectional antenna is broadcast in nature. However, given a network topology, CSI between each user

and total energy constraint, selecting relays that maximizes the broadcast capacity becomes difficult with the introduction of the interference. Therefore, it is both necessary and challenging to improve the broadcast capacity of a wireless multihop interference network under a total energy constraint. In the last part of this thesis, we analyze how broadcast capacity of a wireless multihop network scales when interference is taken into account and propose two heuristics as a practical solution to this problem.

Finally, it is worthwhile to mention the research trend on system optimization in recent years. In the past, system parameters of a novel design are adjusted according to heuristics or common assumptions in the literature. However, many studies [77], [132], [133], [60] show that operating at optimal system parameters gains further benefits to the overall performance. Therefore, each wireless system mentioned in this thesis is optimized according to a set of objectives and constraints.

1.2 Contributions & Thesis Outline

In this thesis, we present three different system designs on improving wireless multimedia broadcasting. First two cases can readily be implemented using the existing 3G infrastructure while third one is a theoretical study on achievable broadcast capacity under realistic assumptions and it can also be implemented in a wireless ad hoc network. The basic premise of this thesis is, in all of the three cases, optimization of system resources are done by taking into account the practical considerations of environment, transmission system and data type that is being transmitted. Therefore, the simulation results reflect the values and behaviors that can be achievied using a practical wireless system.

In this thesis, we try to answer open questions in the literature on wireless multimedia broadcasting by presenting three different system designs. We contribute to the literature on the following aspects, given in three parts.

The first part of this thesis presents;

1. A feasibility study on optimization of wireless broadcasting of pre-encoded video content using SVC,

2. An optimal division in code space among scalable video layers and its effect on the performance of wireless video broadcasting over 3G wireless networks using CDMA

as a channel access method,

3. The optimal number of scalable video layers for users with different mobility characteristics in this framework,

4. A multi-objective optimization technique for two conflicting objectives; maximization of the received video quality and minimization of the service outage.

The second part of the thesis introduces;

1. A cross-layer scheme, which has not been considered for a broadcasting scenario in the literature, for wireless video broadcasting,

2. An investigation of the optimal operating PER for wireless multimedia transmission since this value has not been questioned in the literature for wireless video broadcasting and inherited as 1%, which is an optimized value for voice transmission (i.e., achieving satisfactory quality from vocoders),

3. Joint optimization of a variety of parameters in order to achieve a best compromise operating point satisfying two conflicting objectives given above,

4. The effect of transmitted video content on the system performance since a joint source-channel coding approach is taken in this work.

The third part investigates;

1. The broadcast capacity of a wireless multihop network when interference and total energy constraint are taken into account,

2. The scaling of the broadcast capacity with system parameters, such as bandwidth, power, network density and number of fingers at the receiver,

3. The theoretical achievable broadcast capacity ratio between a multihop and direct transmission when there are infinitely many users in the network,

4. The performance comparison between a centralized and a distributed relay selection strategy since relay selection problem can not be solved optimally in practice.

The outline of this dissertation is as follows. Chapter 2 gives the necessary background information on digital video broadcasting solutions, cdma2000 1xEV-DO, wireless channel model and H.264/AVC. A novel, multi-objective optimized high data rate wireless video broadcasting solution over 3G wireless networks is presented in Chapter 3. Then, another wireless video broadcasting solution, utilizing a cross-layer design, is described in Chapter 4. Chapter 5 examines the achievable broadcast capacity limits of multihop interference networks under a total energy constraint. Finally, Chapter 6 summarizes our conclusions.

Chapter 2

BACKGROUND

2.1 Introduction

Following Heinrich Hertz's demonstration of electromagnetic radiation, Marconi was the first person to successfully transmit a signal over a distance of eighteen miles without using wires [11] [124]. Guglielmo Marconi's successful experiment of a radiotelegraph system was an important milestone in wireless communications history. After this revolutionary breakthrough, it took more than 80 years to establish the first cellular system, Advanced Mobile Phone System (AMPS) [76]. Since then, the pace of development in wireless communications has increased substantially following rapid improvements in microelectronics and digital signal processing.

AMPS, a first-generation (1G) cellular technology, was analog in nature and focused on voice communications. The division of geography into cells and utilization of frequency reuse was first used in a large scale using AMPS. Afterwards, cellular wireless technologies evolved very rapidly. 2G systems, such as cdmaOne (IS-95) and GSM (Global System for Mobile communications), benefited digital communication technologies. As the goal was to provide voice transmission in an efficient manner, these systems were circuit-switched. They had more enhanced voice capability, wider coverage and better spectrum management than their predecessors. However, as voice service was aimed in GSM, the maximum data rate was 14.4 kbps, which is satisfactory for wireless telephony but inadequate for data services.

With the dissemination of the Internet in wireline networks, advancements were made to second-generation wireless systems to provide data service in addition to the voice service. These systems, such as HSCSD (High-Speed Circuit-Switched Data), GPRS (General Packet Radio Service), EDGE (Enhanced Data rates for GSM Evolution) and IS-95B, are categorized as 2.5G technologies in the literature. HSCSD allows assignment of up to four consecutive timeslots to a single user, so it is four times faster than GSM. GPRS implements a packet-switching layer for data services on top of circuit-switching layer of GSM. It allows

assignment of all eight timeslots of GSM to a single user and it achieves a maximum data rate of 171.2 kpbs. Data rates up to 384 kbps are available by EDGE, where GPRS is further enhanced using variable-rate modulation and coding. On the other hand, a maximum data rate of 115.2 kbps is available in IS-95B, where a maximum of eight orthogonal Walsh functions can be assigned to a single user. On the whole, 2.5G wireless systems enable data services by making enhancements to the initial voice-centric systems.

In the late 1990s, ITU defined IMT-2000, a global standard including specifications for the desired third-generation (3G) digital cellular systems. In IMT-2000, broadband data services, such as Internet access, audio and video streaming and gaming, were added to the initial voice services. As a response to ITU IMT-2000 requirements, there were two competing standards, W-CDMA and cdma2000. While 3GGP1 supported W-CDMA, which is backward compatible with GSM and IS-136, 3GPP2 (Third Generation Partnership Project 2) supported cdma2000, which is backward compatible with cdmaOne. Although both standards use CDMA with power control and RAKE receivers, they are not compatible with each other due to the discrepancies between chip rates and other minor differences. Initial 3G wireless systems use a combination of circuit-switching and packet-switching since providing data services using a system optimized for voice services is not spectrally efficient due to the nonoverlapping objectives and constraints.

The core of cdma2000 wireless air interface standard is known as cdma2000 1xRTT (1xRadio Transmission Technology), where '1x' indicates the use of same RF (Radio Frequency) bandwidth as IS-95, one pair of 1.25 MHz radio channel. Therefore, cdma2000 is backward compatible with the IS-95 system. The evolutions made to cdma2000 include cdma2000 1xEV-DO, cdma 2000 1xEV-DV and cdma2000 3X. cdma2000 1xEV-DO, also known as IS-856 or HDR, enhances the core technology by using a separate of 1.25 MHz channel dedicated to data services. In cdma2000 1xEV-DO Rev. 0 [121], a peak data rate of 2457.6 kbps in downlink and 153.2 kbps in uplink is possible. This is established by AMC, fast packet scheduling and ARQ or HARQ. In 1xEV-DO Rev. A [52], the uplink channel capacity is increased substantially to 1.8 Mbps while the downlink capacity is 3.1 Mbps. In addition, further improvements on coverage and number of simultaneously supported users are achieved in this revision [9]. These advancements enable operators to meet the demand for popular services, such as VoIP, mobile gaming, file transfer and video streaming.

The second evolution, cdma2000 1xEV-DV supports a maximum data rate of 4.8 Mbps to all 1xRTT and 1xEV-DO users within the same channel by defining up to two downlink packet data channels that are separated with orthogonal Walsh codes. The latest evolution, cdma2000 3X [53], also known as cdma2000 1xEV-DO Rev. B, is a multi-carrier enhancement to 1xEV-DO Rev. A specification. It aggregates three 1.25 MHz radio channel into one 3.75 MHz channel and is expected to support a peak data rate of 14.7 Mbps for enabling of new services such as high definition video streaming.

W-CDMA, as the competing standard to cdma2000, is referred to as UMTS, being one of the 3G cellular technologies. UMTS is the continuation of GSM to 3G standards and therefore sometimes called as 3GSM. Compared to cdma2000 standards, a maximum data rate of 2.4 Mbps is supported in W-CDMA by using 5 MHz channels. Similar to cdma2000, evolutions are made to W-CDMA to provide higher data rates. HSDPA (High Speed Downlink Packet Access) is an enhancement to W-CDMA, where a peak data rate of 14.4 Mbps is possible. A further increase in data rates is possible with HSPA+, also known as Evolved HSPA (High Speed Packet Access) or HSPA Evolution, where a data rate of 42 Mbps in the downlink and 22 Mbps in the uplink is provided via the use of MIMO and higher order modulation techniques. Recently, as [23] indicates, the work on LTE (Long Term Evolution) was started to improve the performance of 3G wireless networks to meet the further demands and expectations in the near future.

On the whole, starting from the first generation analog systems to second and third generations, wireless communications evolved tremendously in the last 3 decades. After the deployment of 3G systems, many researchers and engineers started working on 4G wireless systems. Although many companies have made prototype systems for 4G, also known as Beyond 3G, these systems can not be called as the 4G cellular technology of choice until the 4G standards are defined. As the debate goes on about 4G, the 4G working group has defined the objectives of this new technology. An all-IP packet-switched network, a data rate of at least 100 Mbps, smooth handoff between different networks, seamless connectivity, combatibility with the existing standards and high spectral efficiency can be listed as the important objectives of the upcoming 4G wireless communications standard.

The remainder of this chapter is organized as follows. In the next section, we will provide a brief information on current digital video broadcasting solutions. Then, 1xEV-DO Rev.

0 (IS-856), which is used in the simulations as a reference 3G wireless system, is reviewed in Section 2.3. Wireless channel models and state-of-the-art video coding technology used in the simulations in this thesis is described in the subsequent sections.

2.2 Digital Video Broadcasting Solutions

Following the rapid growth of wireless networks and the success of Internet video, wireless video services are expected to be widely deployed in the near future. Of such services, wireless video broadcasting has gained significant attention recently thanks to two highly publicized proposals: DVB-H [30] and MediaFlo [99]. While DVB-H technology uses already available DVB-T (Digital Video Broadcasting-Terrestrial) infrastructure and spectrum, MediaFlo requires a new infrastructure for deployment. Two classes of solutions exist for wireless video broadcasting, namely, high and low power solutions. The two above mentioned proposals both require transmission powers much greater than those used in today's 2G and 3G wireless systems. Subsequently, these proposals require the deployment of brand new cellular networks where cell sizes are much larger than those currently in use (tens of kilometers of cell radii).

Alternatively, one can provide a low-power, lower bandwidth video broadcasting solution using the existing 3G wireless network infrastructure. Evolutions, such as 1xEV-DO (IS-856) [121],1xEV-DV [67] and HSDPA [75] to the 3G standards provide spectrally efficient data services. In these systems, adaptive coding and modulation is used to accommodate for the variations in the wireless channel conditions. MBMS is an example for this second class of video broadcasting solutions. It can operate over the existing UMTS networks without requiring any change in the infrastructure or hardware.

In the subsequent subsections, we will provide information about these three wireless video broadcasting solutions, DVB-H, MediaFlo and MBMS.

2.2.1 DVB-H

DVB-H is a transmission system proposed by ETSI (European Telecommunications Standards Institute) [26], [27] for carrying multimedia services over terrestrial networks to handheld devices in an efficient manner. It is developed as an extension of DVB-T with additional features enabling efficient delivery of video data to mobile devices. DVB-H offers

many broadcasting services, in which mobile television is the most popular one. Although mobile television is accessible through the readily deployed UMTS networks, it requires many point-to-point connections and generates a large amount of network traffic. Therefore, DVB-H provides an efficient solution for transmitting multimedia to mobile terminals using the classical broadcast channel.

As the DVB-H standard uses OFDM air interface, the mobile terminal can combine different DVB-H transmissions using SFN (Single Frequency Network) functionality. According to SFN, all the transmitters in a specific area use the same frequency for transmission and transmit the same content with identical waveforms. So, signals from different transmitters can be viewed by the receiver as multipaths with different propagation delays from a single transmitter. Also, use of high power transmission at DVB-H stations allow SFN cell sizes of up to ten kilometers. DVB-H includes a variety of features which makes the system efficient for wireless broadcasting. The standard allows mobile terminals to repeatedly turn their power off during the reception in order to reduce the battery usage. Also, soft handover, which is not possible for single antenna DVB-T terminals, is possible in the DVB-H standard by two main features, time slicing and mandatory cell id identifier. At the link layer, time slicing enables a receiver to stay active for only a fraction of the time and creates off times for monitoring neigbouring cells. This feature enables smooth and seamless handover between cells while reducing the average power consumption of the mobile terminal. In addition, at the physical layer, a cell id identifier is carried to assist the handover process.

On the other hand, unlike DVB-T, DVB-H delivers the video and audio data using IP datacasting. Therefore, the data is encapsulated with IP headers and transmitted through the wireless channel. The use of IP in DVB-H enables interactivity as well as immediate access to the vast multimedia content in the Internet. As wireless channel is prone to high error rates, DVB-H makes use of enhanced error protection. This is done at the link layer by using FEC for multiprotocol encapsulated data (MPE-FEC) which provides Doppler effect compensation while improving the SNR at the receiver and the tolerance to impulse interference. MPE-FEC is employed on the IP packet level by combining Reed-Solomon coding with extensive time interleaving. Furthermore, 4K mode, which is an optimized mode for modulation is used at the physical layer for trading off mobility and SFN cell size.

As a result, DVB-H offers flexibility and scability in terms of optimizing system coverage and transmission rates.

A more detailed information on DVB-H can be found in these recent papers [30], [64], [65] and the references therein.

2.2.2 MediaFlo

MediaFlo is a proprietary digital broadcasting technology of Qualcomm designed for high-quality multimedia streaming services. FLO stands for "forward link only", defining uni-directional transmission. As the system was designed from scratch, it has no backward compatibility with the other digital broadcasting solutions. MediaFlo operates on the 700 MHz spectrum, which is a relatively low frequency, with an RF channel bandwidth of 6 MHz and high transmitter power.

In MediaFlo, the physical layer is based on OFDM with 4096 (4K) subcarriers either QPSK or 16QAM modulated and SFN is employed for high spectral efficiency. In addition, active subcarriers are divided into disjoint groups, called interlaces. The interlace structure enables frequency-division multiplexing of MLCs (Multicast Logical Channels) within each OFDM symbol. On the other hand, this structure not only provides fine-grain transmission of MLCs, but also decreases the power consumption at the receiver side by enabling demodulation of only the desired MLCs. [15]. The IP-based MediaFlo network is based on multiple types of encoding schemes, including H.264/AVC and MPEG-4, and layered source coding and modulation types. With layered source coding, video playback can continue at lower bit rates instead of freezing, in case of SNR degradation due to adverse conditions. Besides, Turbo coding (an outer Reed-Solomon code and an inner parallel concatenated convolutional code) based FEC is employed for robust delivery with the use of layered modulation. Similar to DVB-H, the use of different coding and modulation types presents a flexibility for trading off coverage and data rate. For reducing power consumption, MediaFlo supports statistical multiplexing of different services or MLCs by varying time and frequency allocations. Furthermore, this fine-grained multiplexing allows the implementation of a variable rate encoding. Thus, the available bandwidth can be dynamically allocated in contrast to fixed bandwidth allocation for one time slice in DVB-H.

2.2.3 MBMS

MBMS has been standardized in 3GPP (Third Generation Partnership Project) Release 6 [85], [86] to enable point-to-multipoint transmission within UMTS networks. With MBMS, cellular networks can efficiently disseminate the same information to a large number of users at the same time. The concept of multimedia broadcast and multicast service center (BM-SC), which provides scheduling, authentication, charging and security functions, is introduced in MBMS. Although BM-SC is similar to IP datacasting in DVB-H, the functionality of BM-SC is more than an IP encapsulator due to the dynamic bearer management.

The features of MBMS are split into two, the MBMS bearer service [87] and the MBMS user service [88]. The MBMS bearer service defines transmission procedures below the IP layer and it includes two modes of operation, broadcast and multicast. In broadcast mode, all subscribed users receive the same transmitted signal and the movements of users are not tracked. Besides, each user sends a request for joining a group prior to transmission, user activity is monitored and system resources are allocated accordingly in the multicast mode. In this mode, each cell may decide on point-to-point or point-to-multipoint transmission according to the number of subscribed users in the multicast group [95]. On the other hand, the MBMS user service defines service layer protocols and procedures. In MBMS user service, two delivery methods, streaming and download, are offered. While streaming is intended for continous media playback, such as mobile TV, download can be used for small video clips and songs.

In MBMS, point-to-point transmissions are handled by the high-speed downlink shared channel, which is present in WCDMA (Wideband Code Division Multiple Access) systems. Point-to-multipoint transmissions use forward access channel with coding and modulation at a fixed transmission power. Similar to DVB-H, macro diversity can be utilized for enhancing the SNR by combining signals received from different cells as each cell transmits the same information in broadcast mode. In addition, long transmission time intervals are used for time diversity against fading since MBMS services are not delay sensitive. On the other hand, Raptor codes can be used for FEC against transmission errors with the FEC stream bundling concept, introduced in MBMS. The FEC stream bundling includes the data, control and keying stream together and as error correction works on a longer data stream, the FEC efficiency is increased. In addition to FEC, two types of file repair procedures are

offered by MBMS, point-to-point interactive bearers and point-to-multipoint bearers [45].

As a result, MBMS does not require any extra infrastructure if the 3G cellular network is already deployed. In addition, mobile terminals in a UMTS network can use MBMS without changing any hardware, such as antenna and chipset. Therefore, deployment of MBMS to the UMTS network is both easy and cheap compared to building a new infrastructure.

2.3 CDMA2000 1xEV-DO

CDMA2000 1xEV-DO (also known as IS-856 or HDR) is a standard, specifically designed for providing high speed data access to mobile users [8]. It is a third-generation wireless system and high rate packet data air interface standard accepted by 3GPP2 [121]. 1xEV-DO is an abbreviation of Evolution-Data Optimized or Evolution-Data Only. It is proposed to satisfy the increasing demand for data services after the industrialization of wireless telephony. Today, there are more than 105 million CDMA2000 1xEV-DO subscribers and almost 475 million CDMA subscribers including cdmaOne, CDMA2000 1x and CDMA2000 1xEV-DO as of September 2008 [14].

CDMA2000 1xEV-DO includes a number of techniques that make the system spectrally efficient. The system serves one user at a given time-slot using AMC and this is proven to be optimal due to the water-pouring in time [36]. In addition, the forward link is used to transmit data to multiple users in a TDM fashion with time-slots of 1.67 ms. At each time-slot, a user is selected by the scheduling algorithm that exploits multi-user diversity in a fading environment. However, the standard for 1xEV-DO does not specify a certain scheduling algorithm. There are a variety of scheduling algorithms, such as round robin, exponential, proportional fairness, maximum rate and minimum performance in the literature. Among those algorithms, proportional fairness achieves a better compromise between maximizing system capacity and fairness among mobile terminals [10]. The algorithm tries to serve the terminal, which has a much better current channel condition than its average condition over a specific time window. The current channel condition is the data rate requested by the mobile user on the DRC channel and the average condition is the infinite impulse response filtered user throughput. As each mobile terminal tracks two or more base stations simultaneously, the data rate request is transmitted to the base station, with the highest SNR value among all the received SNRs, over the reverse link. By doing so, forward

link is adapted to the varying channel conditions. In addition, as each terminal selects the base station with the highest SNR, interference to other terminals is reduced.

On the other hand, 1xEV-DO system uses incremental redundancy HARQ in addition to AMC for further adaptation to the varying channel conditions. Encoded packets can be transmitted using multiple slots and early termination is possible (via signaling a positive acknowlegment) after partial reception of the encoded packets. Besides, repetition and Turbo coding increases the error resilience capability of the physical layer packets against transmission errors. Under harsh channel conditions, the incremental redundancy HARQ with Turbo coding provides further adaptation to the fading environment and increases the throughput of the system.

On the whole, CDMA2000 1xEV-DO system includes a variety of transmission modes with different system parameters, optimized for a given SNR interval. In each of these modes, system parameters, such as number of slots, packet size, Turbo code rate, modulation type and packet repetition factor, are defined. Therefore, rate adaptation of the forward channel is achieved by switching between these modes according to the channel conditions reported to the system via the uplink channel. A detailed information on the structures of the forward link and the reverse link will be presented in the following two subsections.

2.3.1 Forward Link

The forward link includes the pilot channel, the MAC channel, the forward traffic channel and the control channel as shown in Fig. 2.1, and TDM is used in order to eliminate the unused power margin problem in the older system, IS-95. As base station transmits at full power, no power control mechanism is used in the forward link. The forward channel uses direct-sequence spread at the rate of 1.2288 Mcps and allocates a bandwidth of 1.25 MHz. One time-slot consists of 2048 chips and it has a duration of 1.667 ms. All the four channels in the forward link are time-division multiplexed in one time-slot and transmitted at full power. One time-slot consists of 2048 chips, where 1600 of them are dedicated to data segment, 256 chips are dedicated to the MAC and 192 chips are dedicated to the pilot channel. A detailed description of the forward link structure within a time-slot for both active and idle states can be seen in Fig. 2.2.

The pilot channel is an unmodulated signal used for synchronization and other functions

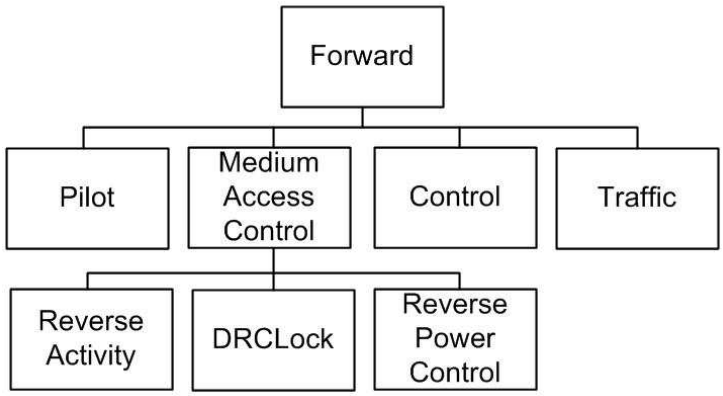


Figure 2.1: Forward Link Hierarchy.

by the mobile station operating within the coverage area. This signal is transmitted at all times by the base station with full power. This signal is composed of all 0's and it is transmitted twice within one-slot of the forward link as seen in Fig. 2.2. As pilot channel provides an estimate of the channel condition for the rate adaptation, it is also used for initial acquisition, phase and timing recovery and symbol combining. Thus, each cell has different PN (Pseudorandom Number) offsets and they are differentiated via the use of pilot channel.

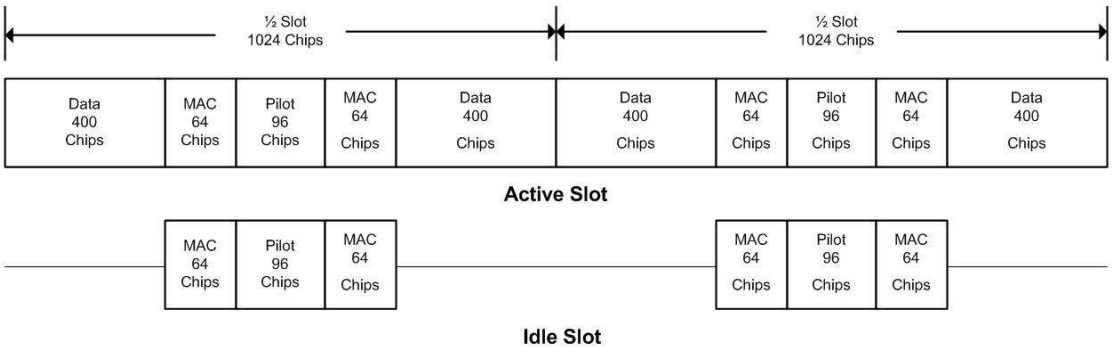


Figure 2.2: Forward Link Slot Structure.

The forward MAC channel is used to transmit reverse power control and reverse activity channel information to the mobile terminals. The reverse activity information is used by

the reverse MAC algorithm to control the interference level. Furthermore, reverse activity channel provides information about the loading condition of the reverse link and whether the terminal should increase or decrease its data rate. On the other hand, forward MAC channel includes the DRCLock channel, which is used to provide information on the condition of the successful decoding of the DRC sent by the terminal to the base station. The reverse power control (RPC) channel includes the RPC bit stream, which is used for reverse link power control by the specified terminal. The RPC channel and the DRCLock channel are transmitted on the same MAC channel using TDM. The forward MAC channel is composed of 16 Walsh channels that are orthogonally covered and BPSK modulated on either in-phase or quadrature phase of the carrier. Each Walsh channel is identified by a MACIndex value and defines a unique 64-ary Walsh cover and associated modulation phase. The MAC symbol Walsh covers are transmitted four times per slot immediately before and after the pilot signal of each slot.

The data channel or the forward traffic channel carries the physical layer data packets of a single user, that is served by the base station at a given time-slot. A preamble sequence is transmitted with each forward traffic channel and control channel using TDM in order to assist the terminal with synchronization. This sequence consists of all 0 symbols that are transmitted on the in-phase component only. The preamble sequence is covered by a 32-chip bi-orthogonal sequence, which reduces the complexity at the demodulator. The packets transmitted by the control channel are used for base station and terminal messaging, while the forward traffic channel packets carry the data payloads for the specific terminal. The cdma2000 1xEV-DO system defines 12 transmission modes with 9 unique data rates for the forward traffic channel as shown in Table 2.1.

The forward traffic channel or control channel physical layer packets are processed as shown in Fig. 2.3 to produce the forward modulated waveform. All the system parameters, such as rate of Turbo encoder, modulation type and repetition, are defined by the selected transmission mode given in Table 2.1. First, the physical layer packets are coded with Turbo encoder for protection against transmission errors. The Turbo encoder uses two systematic, recursive, convolutional encoders in parallel with the Turbo interleaver. The outputs of the convolutional encoders are punctured and repeated in order to match the desired number of output symbols determined by the transmission mode of the system. Then, these bits

DRC Index	Data Rate (kbps)	Slots	Packet Size (bits)	Code Rate	Mod. Type	Packet Repetition	TDM Chips (Preamble, Data, MAC, Pilot)
1	38.4	16	1024	1/5	QPSK	9.6	(1024,24576,4096,3072)
2	76.8	8	1024	1/5	QPSK	4.8	(512,12288,2048,1536)
3	153.6	4	1024	1/5	QPSK	2.4	(256,6144,1024,768)
4	307.2	2	1024	1/5	QPSK	1.2	(128,3072,512,384)
5	614.4	1	1024	1/3	QPSK	1	(64,1536,256,192)
6	307.2	4	2048	1/3	QPSK	2.04	(128,6272,1024,768)
7	614.4	2	2048	1/3	QPSK	1.02	(64,3136,512,384)
8	1228.8	1	2048	1/3	QPSK	0.5	(64,1536,256,192)
9	921.6	2	3072	1/3	8PSK	1.02	(64,3136,512,384)
10	1843.2	1	3072	1/3	8PSK	0.5	(64,1536,256,192)
11	1228.8	2	4096	1/3	16QAM	1.02	(64,3136,512,384)
12	2457.6	1	4096	1/3	16QAM	0.5	(64,1536,256,192)

Table 2.1: Available Data Rates in cdma2000 1xEV-DO Forward Traffic Channel.

are scrambled and reordered by the interleaver. Scrambling is done by XORing the output of the Turbo encoder with a predefined randomized scrambling sequence known by the receiver. Similarly, the interleaving method is known by the receiver for each transmission mode. The goal of the interleaving is to help decoder perform better when dealing with random errors instead of bursty errors. Additionally, interleaving assists HARQ scheme for early terminated, which will be mentioned in the next paragraph.

After scrambling and interleaving, the bits are modulated according to the defined modulation type by the transmission mode. As the transmission mode is selected according to the wireless channel conditions, best modulation method is selected for the channel condition. Thus, cdma2000 1xEV-DO system adjusts the best modulation and coding for the current channel condition of the user in order to maintain the optimal data rate for a specific user. Then, the data rate of the modulated packet is adjusted by repetition or truncation.

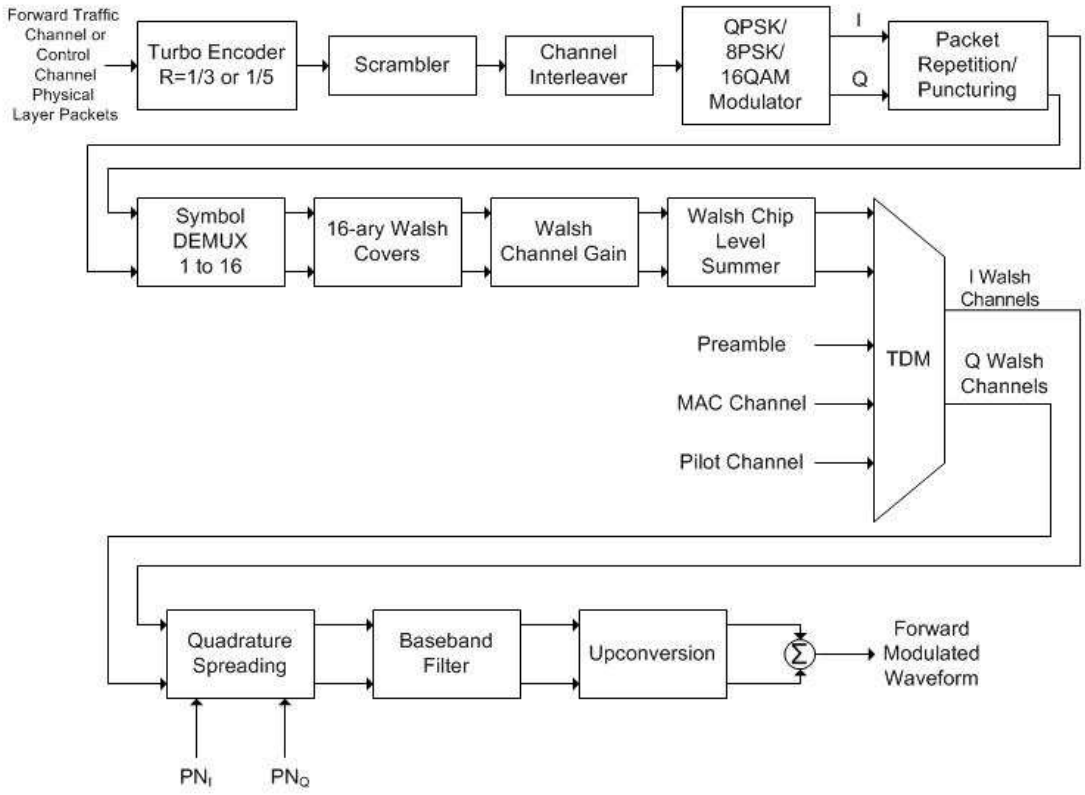


Figure 2.3: Forward Channel Structure.

If the data rate is increased, then modulated packets are truncated and smaller packets are created by removing the last portion of the packet. Conversely, a modulated packet can be repeated. In repetition, the first portion of the packet can be used as an additional partial packet in order to match a lower data rate determined by the system. As the repetition is a form of FEC, it increases the error correction capabilities of the system. Since the original packet is repeated in the transmitted packet, the reception of the entire transmission packet is not necessary for successful decoding. This provides an early termination of the transmission for multiple slot packets, and provides a form of HARQ to the system. This property, incremental redundancy HARQ, is advantageous when accurate channel prediction is not possible, e.g. at high speeds, or interference level is high [10]. The early termination can be achieved by positive acknowledgment, indicating a successful decoding of the packet, after each slot.

After packet repetition or puncturing, each packet is bit-by-bit demultiplexed into 16 blocks. Then, each block is spread using one of the 16 orthogonal Walsh codes and summed together to form the spread transmission packet. On the receiver side, the original information packet is recreated by de-spreading the received packet with the same Walsh codes. It should be noted here that the information bits are spread entirely by means of the error control coding in 1xEV-DO and Walsh code spreading does not cause any further increase in the transmission bandwidth. This is in line with the optimal coding-spreading trade-off discussed in [126] for a system using matched filter receivers since the system transmits to only one user at a given time and thus, does not need multiple accessing in the code domain to separate different transmitted signals. As a result, the data segment of the forward link is created after all of these processes. Then, preamble, MAC and pilot are appended to the data segment using TDM as shown in Fig. 2.2.

2.3.2 Reverse Link

The communication from the mobile terminal to the base station is provided via the uplink channel. In cdma2000 1xEV-DO, uplink channel uses CDMA with a short PN code, which is specific to each mobile user and a long PN code, which is specific to the base station. Mobile terminals are distinguished by the short PN code assignment. Five different transmission modes, ranging from 9.6 kbps to 153.6 kbps, on the reverse link of cdma2000 1xEV-DO system are shown in Table 2.2.

DRC Index	Data Rate (kbps)	Packet Size (bits)	Code Rate	Mod. Type	Packet Repetitions	PN Chips
1	9.6	256	1/4	BPSK	8	128
2	19.2	512	1/4	BPSK	4	64
3	38.4	1024	1/4	BPSK	2	32
4	76.8	2048	1/4	BPSK	1	16
5	153.6	4096	1/2	BPSK	1	8

Table 2.2: Available Data Rates in cdma2000 1xEV-DO Reverse Traffic Channel.

Although we will not go into details of the reverse link, the structure can be viewed in Fig. 2.4. An important point to mention is that the mobile terminals periodically report the achievable data rate, which is dependent on the wireless channel conditions, to the base station using the data rate control (DRC) channel of the uplink. As each data rate is associated with a DRC index, only four bits representing this index are sent. These periodic updates are used to select the data rate that can be supported at a specific time slot, so that the adaptive data rate selection is provided.

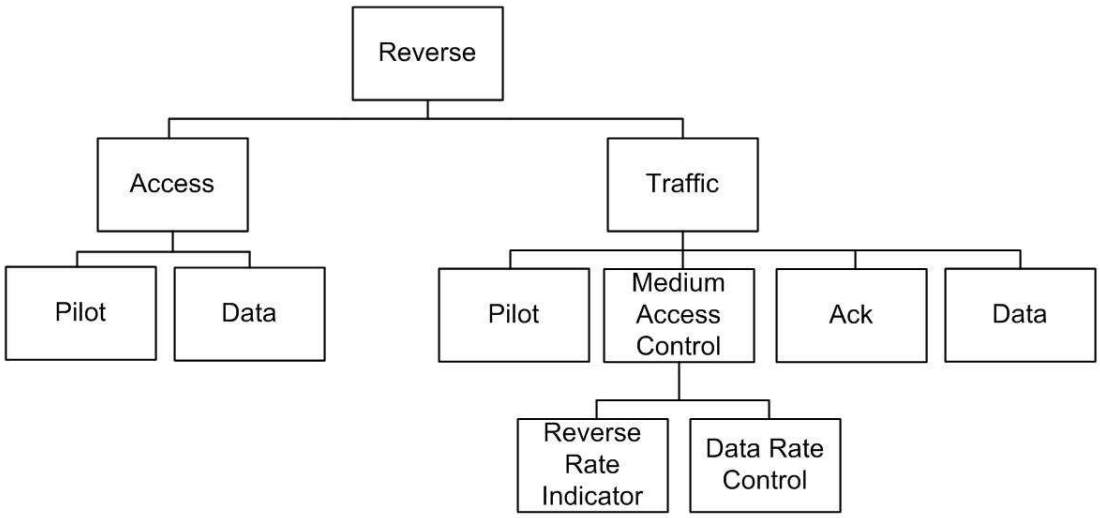


Figure 2.4: Reverse Link Structure.

2.4 Wireless Channel

The wireless channel imposes certain difficulties and it is limited by a variety of factors by its nature. The condition of the wireless channel not only depends on the movement of the receiver but also on the dynamics of the environment. In addition, existence of thermal noise and interference makes the wireless channel condition unpredictable. In general, there are three main causes of the variations in the channel condition, path-loss, shadowing and multipath fading. The variation in the signal power due to path-loss and shadowing are referred to as large-scale variations while change in the power due to constructive and destructive addition of multipath signal components are referred to as small-scale variations

[122]. In the literature, there exist several models, based on analytical results or empirical measurements, proposed for consideration of these effects.

Next, we provide a summary on path-loss, shadowing and multipath fading models used throughout the thesis. As practical considerations are taken into account, we use the models that fit best to our proposed scenario.

2.4.1 Path-Loss

The simplest representation of the path-loss is given by free space path-loss model, where signal is assumed to be transmitted through free space and propagating along a straight line from a transmitter to a receiver with distance d in between. Then, the received signal power to transmitted signal power ratio can be written as:

$$\frac{P_r}{P_t} = \left\{ \frac{G_T G_R \lambda}{4\pi d} \right\}^2 \tag{2.1}$$

where G_T and G_R are transmitter and receiver antenna gains, and $\lambda = c/f_c$ is the signal wavelength.

However, a variety of path-loss models [25], [46], [127], [51], [29] are proposed for different environments to represent the effect of path-loss more accurately. We provide a brief overview on the ITU IMT-2000 and COST231-Walfisch-Ikegami path-loss models, which are used in this thesis. More detailed information on path-loss models can be found in [96], [119], [70].

According to ITU IMT-2000 [54] channel models, there are two types of path-loss models. First model is for outdoor to indoor and pedestrian test environment and is given by the following formula:

$$L = 40\log_{10}(d) + 30\log_{10}(f_c) + 49 \tag{2.2}$$

where d is the separation between base station and mobile station in kilometers and f_c is the carrier frequency.

Second model is for vehicular test environment and is given by the following equation:

$$L = 40(1 - 4.10^{-3}\Delta h_b)\log_{10}(d) - 18\log_{10}(\Delta h_b) + 21\log_{10}(f_c) + 80 \tag{2.3}$$

where Δh_b is the base station antenna height measured from the average rooftop level in meters. This value is taken as 15 m. in our simulations when path-loss is modeled according to ITU IMT-2000 channel models.

According to COST231-Walfisch-Ikegami non-line-of-sight (NLoS) propagation model [119], path-loss is composed of three terms:

$$L_{p(dB)} = \begin{cases} L_o + L_{rts} + L_{msd}, & \text{for } L_{rts} + L_{msd} \geq 0 \\ L_o, & \text{for } L_{rts} + L_{msd} < 0 \end{cases} \tag{2.4}$$

where L_o is the free space path-loss, L_{rts} is the roof-to-street diffraction and scatter loss, and L_{msd} is the multi-screen diffraction loss. These parameters can be computed using the following equations:

$$L_o = 32.4 + 20\log_{10}(d) + 20\log_{10}(f_c) \tag{2.5}$$

$$L_{rts} = -16.9 - 10\log_{10}(w) + 10\log_{10}(f_c) + 20\log_{10}(\Delta h_m) + L_{ori} \tag{2.6}$$

where

$$L_{ori} = \begin{cases} -10 + 0.354(\phi), & 0 \leq \phi \leq 35^\circ \\ 2.5 + 0.075(\phi - 35^\circ), & 35^\circ \leq \phi \leq 55^\circ \\ 4.0 - 0.114(\phi - 55^\circ), & 55^\circ \leq \phi \leq 90^\circ \end{cases} \tag{2.7}$$

is an orientation loss, $\Delta h_m = h_{Roof} - h_m$ is the relative height of mobile station to rooftops, h_m is the height of the mobile station, h_{Roof} is the roof heights of buildings and w is the width of the streets, all in meters.

$$L_{msd} = L_{bsh} + k_a + k_d\log_{10}(d) + k_f\log_{10}(f_c) - 9\log_{10}(b) \tag{2.8}$$

where L_{bsh} is the shadowing gain for cases when the base station antenna is above the rooftops, k_a is the increase in path-loss when base station antenna is below the rooftops of adjacent buildings, k_d and k_f control the dependency of L_{msd} on d and f_c, respectively, and b is the building separation in meters. These parameters can also be computed by:

$$L_{bsh} = \begin{cases} -18\log_{10}(1 + \Delta h_b), & h_b > h_{Roof} \\ 0, & h_b \leq h_{Roof} \end{cases} \tag{2.9}$$

$$k_a = \begin{cases} 54, & h_b > h_{Roof} \\ 54 - 0.8\Delta h_b, & d \geq 0.5 \text{ km and } h_b \leq h_{Roof} \\ 54 - 0.8\Delta h_b d/0.5, & d < 0.5 \text{ km and } h_b \leq h_{Roof} \end{cases} \tag{2.10}$$

$$k_d = \begin{cases} 18, & h_b > h_{Roof} \\ 18 - 15\Delta h_b / h_{Roof}, & h_b \leq h_{Roof} \end{cases} \tag{2.11}$$

$$k_f = -4 + \begin{cases} 0.7(f_c/925 - 1), & \text{medium city and suburban} \\ 1.5(f_c/925 - 1), & \text{metropolitan area} \end{cases} \tag{2.12}$$

where Δh_b is the relative height of base station relative to rooftops.

Throughout the thesis, we assume that $h_b = 13$ m., $h_m = 2$ m., $h_{Roof} = 12$ m., $b = 20$ m., $w = b/2$, and $\phi = 90^\circ$ in the COST231-Walfisch-Ikegami model.

2.4.2 Shadowing

As transmitted signal propagates through a wireless channel, it occasionally loses part of its power when it is obstructed by objects in the environment. This loss in the signal power is a random process since it is caused by changes in the dynamics of the reflection, diffraction and scattering. As a result, a random attenuation occurs in the signal power.

In particular, shadow fading takes into account the large-scale variations in the signal power caused by the interaction between the signal and the large objects in the environment. In general, log-normal distribution is used to model this variation in the signal power. To model this variation, several methods [20], [4], [82], [3], [35], [39] are proposed in the literature. In this thesis, we use the simple model proposed by Gudmundson [41], where log-normal shadowing is modeled as a Gaussian white noise process filtered by a first-order low-pass filter as described in [119]:

$$\Omega_{k+1(dBm)} = \zeta \Omega_{k(dBm)} + (1 - \zeta)\nu_k \tag{2.13}$$

where $\Omega_{k(dBm)}$ is the mean envelope at location k, ζ is a parameter controlling the spatial correlation of the shadow fading, and ν_k is a zero-mean Gaussian random variable with autocorrelation, $\phi_{vv}(n) = \tilde{\sigma}^2\delta(n)$. The spatial autocorrelation function of $\Omega_{k(dBm)}$ can be written as:

$$\phi_{\Omega_{(dBm)}\Omega_{(dBm)}}(n) = \frac{1-\zeta}{1+\zeta}\tilde{\sigma}^2\zeta^{|n|} \tag{2.14}$$

On the other hand, the variance of the log-normal shadowing, which is the value of the autocorrelation evaluated at zero, can be written as:

$$\sigma_\Omega^2 = \phi_{\Omega_{(dBm)}\Omega_{(dBm)}}(0) = \frac{1-\zeta}{1+\zeta}\tilde{\sigma}^2 \tag{2.15}$$

Following these equations, we can model the shadow fading using the Gudmundson's model by relating the decorrelation parameter ζ with the simulation index k. We assume that the envelope is sampled every T seconds. Then, the mobile station moves a distance of vkT in kT seconds. If ζ_D defines the shadow correlation between two points separated by a distance of D in meters, then the autocorrelation function becomes:

$$\phi_{\Omega_{(dBm)}\Omega_{(dBm)}}(kT) = \sigma_\Omega^2\zeta_D^{(vT/D)|k|},\ k \geq 0 \tag{2.16}$$

Comparing eqns. (2.15) and (2.16), we see that $\zeta = \zeta_D^{(vT/D)}$. Using this method, shadow fading is simulated according to parameters given in [41] for the microcellular propagation.

2.4.3 Multipath Fading

Mobile transmissions are exposed to basic propagation phenomena, reflection, diffraction, and scattering. The signal can reach to its destination by being reflected at or diffracted by or scattered from various objects in the environment [84], such as buildings, mountains, trees, etc. As mentioned, large variations in the signal power due to large objects in the environment is modeled by shadow fading. On the other hand, the rapid fluctuations in the signal power is modeled by multipath fading and categorized as small-scale variations.

These fast variations in signal power occurs as either the transmitter or the receiver is moving. In addition, the locations of reflectors in the transmission path changes with time. Thus, multipath channel is time-varying by its nature. The multipath phenomena occurs due to the scattering, when incoming signal hits an object whose size is in the order of the wavelength of the signal or less. On the receiver side, several instances of the original transmitted signal arrive with different amplitudes and time delays, causing a constructive or destructive effect. Furthermore, the movement of mobile station introduces a Doppler shift, a shift in frequency, based on the velocity of the mobile station.

Simulation of multipath fading using a variety of methods has been studied in [100], [31], [6], [21], [55], [137] using several probability density functions. A detailed information on methods for laboratory simulation of multipath fading can be found in [119]. Throughout the thesis, we use the IMT-2000 channel model [54] parameters in the filtered Gaussian noise model and Clarke's scattering model both mentioned in [119] for the simulation of multipath fading.

According to the filtered Gaussian noise model, a fading simulator is generated by filtering two independent white Gaussian noise sources with low-pass filters. In fact, these filtered two noise sources are in-phase and quadrature components of the Rayleigh faded envelope, $g_I(t)$ and $g_Q(t)$. In order to produce a Rayleigh faded envelope, noise sources must have the same power spectral density (psd). If two independent noise sources have the same psd of $\Omega_p/2$ watts/Hz and the low-pass filters have the same transfer function of $H(f)$, then the auto-power spectral density and the cross-power spectral density of the components can be written as:

$$S_{g_I g_I}(f) = S_{g_Q g_Q}(f) = \frac{\Omega_p}{2}|H(f)|^2 \tag{2.17}$$

$$S_{g_I g_Q}(f) = 0 \tag{2.18}$$

In simulations, the low-pass filter $h(t)$ is implemented by a first-order low-pass digital filter which models the fading as a Markov process. If $g_{I,k}$ and $g_{Q,k}$ are Gaussian random variables at the kth simulation time, i.e., $g_{I,k} = g_I(kT),\ g_{Q,k} = g_Q(kT)$, the state equation can be written as:

$$(g_{I,k+1}, g_{Q,k+1}) = \zeta(g_{I,k}, g_{Q,k}) + (1-\zeta)(w_{1,k}, w_{2,k}) \tag{2.19}$$

where $w_{I,k}$ and $w_{Q,k}$ are two independent zero-mean Gaussian random variables with time autocorrelation, $R_{w_{i,k}w_{i,l}} = \sigma^2\delta_{kl},\ i = 1, 2$. The complex output $g_k = g_{I,k} + jg_{Q,k}$ has zero-mean, the envelope $|g_k|^2$ is Rayleigh distributed and the phase $\phi_k = tan^{-1}(g_{Q,k}/g_{I,k})$ is uniformly distributed on the interval $[-\pi, \pi]$. To compute ζ, we can set the 3 dB point of $S_{gIgI}(f)$ to $f_m/4$, where $f_m = v/\lambda$ is the maximum Doppler frequency.

$$\zeta = 2 - cos(2\pi f_m T) - \sqrt{(2 - cos(2\pi f_m T)^2) - 1} \tag{2.20}$$

Then, to normalize the mean square envelope to Ω_p, the variance of two zero-mean Gaussian random variables can be written as:

$$\sigma^2 = \frac{\Omega_p(1+\zeta)}{2(1-\zeta)} \tag{2.21}$$

As long as uncorrelated noise sources with same psd are used, Rayleigh faded envelope can be generated with ease, using a first-order low-pass filter. Although using a higher order filter provides some improvement, this brings complexity to the fading simulator. Thus, a first-order filter is sufficient to simulate multipath fading. In the simulations, we used a first-order digital filter, which has a pole at ζ followed by eqn. (2.20).

A simple model, where radio waves propagate in 2-D plane and arrive at the terminal from all directions with equal probability, was suggested by Clarke [17]. This model is suitable for macrocellular applications, where signal propagates a distance longer than the antenna heights, and it is referred to as Clarke's 2-D isotropic scattering model. In this model, the Doppler power spectral density of the channel is given as:

$$S_{gg}(f) = \frac{\Omega_p}{4\pi f_m} \frac{1}{\sqrt{1-\left(\frac{f-f_c}{f_m}\right)^2}}, \quad |f-f_c| \leq f_m \tag{2.22}$$

Then, the Rayleigh faded envelope is generated similar to the first model by filtering two independent white Gaussian noise sources. The filter is implemented with a 600-tap FIR filter, whose frequency response is:

$$H(f) = \sqrt{\frac{\Omega_p}{4\pi f_m} \frac{1}{\sqrt{1-\left(\frac{f-f_c}{f_m}\right)^2}}}, \quad |f-f_c| \leq f_m \tag{2.23}$$

2.5 Current Video Coding Standards

2.5.1 H.264/AVC

Video coding standards are being developed by two distinct international organizations, VCEG (Video Coding Experts Group) of ITU-T (International Telecommunications Union Telecommunication Sector) and MPEG (Moving Picture Experts Group) of ISO/IEC (International Standards Organization/International Electrotechnical Commission). ITU-T has developed H.261 and H.263 both mainly aimed for video conferencing application. In the

meantime, MPEG has developed MPEG-1 (for video on CD-ROM), MPEG-2 (for regular TV, HDTV and video on DVD) and MPEG-4 standards. MPEG-2 was also adopted by ITU-T as H.262. By 2001, VCEG and MPEG joined forces under the JVT (Joint Video Team) umbrella to develop a standard to provide better compression efficiency. The result was the H.264/AVC video coding standard [1], [18], also referred to as MPEG-4 Part 10.

Different than prior standards, motion-prediction is improved, small block-sizes are available and enhanced entropy coding methods are developed in H.264/AVC. The motion estimation is enhanced by new variable blocksizes ranging from 16 x 16 to 4 x 4, where a 4 x 4 integer-based transform is introduced for the smallest blocksize and motion vector accuracy is improved to one-quarter pixel. By using the state-of-the-art video codec, up to 50% ($\sim$ 2 dB) additional coding efficiency can be obtained over MPEG-2 standard [134]. The H.264/AVC standard consists of two conceptual layers: VCL and NAL. The VCL defines the efficient representation of the video and the NAL converts the VCL representation into a format suitable for specific transport layers or storage media.

The Video Coding Layer

The video to be compressed consists of a sequence of frames, where each frame is a still picture, displayed at a particular frame rate (frames/sec). Video coding technologies exploit temporal and spatial redundancies present in any video to achieve substantial compression. Temporal redundancy exists due to significant similarity between consecutive video frames. Spatial redundancy refers to correlation of intensity patterns within a frame. The video coding technology used in the H.264/AVC VCL is motion-compensated hybrid coding, where temporal redundancy is removed by motion compensation and spatial redundancy is exploited by a block transform followed by quantization and entropy coding of transform coefficients. In the H.264/AVC VCL, quantization of the transform coefficients is performed using a scalar quantizer whose step size is determined by QP. QP can take on one of 52 distinct values. These values are established such that an increase of 6 in the value of the QP results in doubling the quantization step size.

The H.264/AVC video stream is a sequence of GOPs, where each picture is partitioned into macroblocks. One luma (Y) and two chroma (Cr and Cb) components are contained within each macroblock. The macroblocks are the fundamental coding blocks in H.264/AVC

and are coded in either intra or inter mode. In intra mode, a macroblock is coded without using any motion prediction from the macroblocks in other pictures. In inter mode, on the other hand, the macroblock is coded in reference to one or more macroblocks from previous or future pictures. Macroblocks are organized in slices, which represent regions of a picture that can be decoded independent of each other. A picture, on the other hand, comprises a set of slices representing a complete frame or field. Five different slice types have been defined in H.264/AVC. I-slices (where I stands for Intra) contain macroblocks that are intra coded only. A special type of picture containing only I-slices is called an instantaneous decoder refresh (IDR) picture, where any subsequent picture to the IDR picture does not use pictures before the IDR picture as references for motion compensated prediction. P-slices (where P stands for predictive) on the other hand, may contain intra coded macroblocks as well as at least one inter coded macroblock with motion compensated prediction using one reference macroblock from another picture. B-slices (where B stands for bi-predictive) may contain both intra and inter coded macroblocks with at least one inter coded macroblock using two reference macroblocks from other pictures. As multiple references can be used by the motion compensation, a block in a P or B slice can be predicted by one to four reference blocks in earlier frames. The SI- and SP-slices (where SI stands for switching-I and SP stands for switching-P) are defined to allow switching between multiple coded streams, and are useful for random access and switching between different bit-rate encoded versions of the same content. More detailed information on H.264/AVC standard can be found in [130], [93], [78] and [80].

It has been shown that video coding using a GOP structure composed of hierarchical B pictures provides improved coding efficiencies when compared to coding with the classical GOP structure of IBBP.... [107]. A typical hierarchical GOP structure with 4 dyadic hierarchy stages is illustrated in Figure 2.5. The first picture of a video sequence is always encoded as an IDR picture. In the hierarchical GOP structure, each GOP has a key picture that is either intra-coded or inter-coded using previous key pictures as reference for motion compensation. The remaining pictures of a GOP are hierarchically predicted. In this thesis, we assume that the H.264/AVC uses a hierarchical GOP structure for the wireless broadcast service.

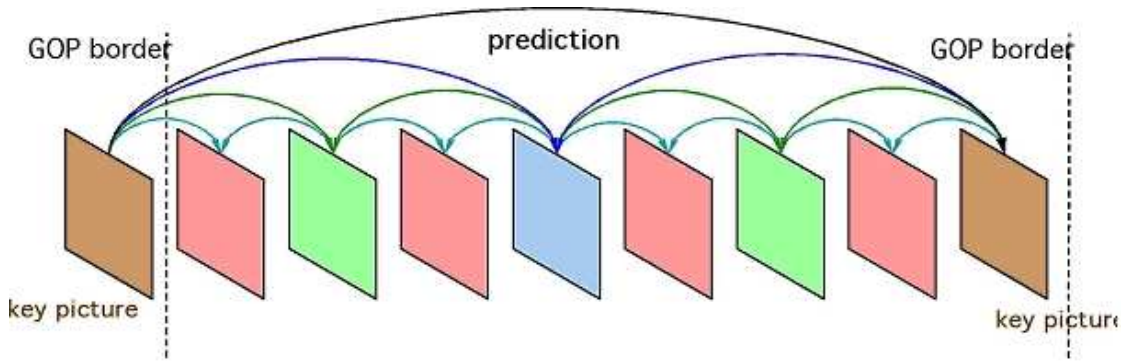


Figure 2.5: GOP Structure with Hierarchical B Pictures.

Network Abstraction Layer

The H.264/AVC NAL converts the VCL representation into a format suitable for specific transport layers or storage media. For circuit-switched transport, the NAL delivers the coded video as an ordered stream of bytes containing start codes such that these transport layers and the decoder can robustly and simply identify the structure of the bit stream. For packet-switched transport, on the other hand, the NAL delivers the coded video in packets without these start codes. The coded video data is organized into NAL units, each of which is effectively a packet that contains an integer number of bytes. The first byte of each NAL unit is a header that contains an indication of the type of data in the NAL unit. The remaining bytes are the payload bytes and are interleaved as necessary with emulation prevention bytes, so that no patterns mimicking start codes are accidentally generated inside the payload.

2.5.2 Scalable Video Coding

Scalable video coding was standardized as an extension of H.264/AVC [18], but the idea of scalable coding existed for older video coding standards [112], such as MPEG-2 [33], H.263 [125] and MPEG-4 [19]. As the rate-distortion efficiency of these codecs for scalable coding (except for temporal scalability) was lower than non-scalable coding, this option was not accepted. However, this efficiency gap is no longer pronounced with the SVC extension of H.264/AVC.

SVC [92] has attracted considerable attention from both video experts and wireless communications community after the introduction of H.264/AVC. SVC enables the capability of encoding a subset of a high-quality video stream with a similar quality that can be achieved by using the H.264/AVC with same quantity of data as in the subset stream [108]. Due to this property, a transmitted video stream can adapt to varying channel conditions and create a granularity at the system throughput instead of a binary case. A scalable video stream consists of substreams, which are valid bit streams with a reconstruction quality less than the original video stream. The substreams of the scalable video coding includes a base layer and one or more enhancement layers. There is a hiererchical dependency between base layer and each enhancement layers. Base layer includes the lowest quality version of the original video stream and it is necessary for successful decoding of each enhancement layers. Similarly, for decoding a higher enhancement layer, successful reception of lower enhancement layers and base layer is required. As expected, decoding a layer increases the video quality and decoding all the layers provides the highest quality.

There are three main modes of scability defined in SVC, temporal, spatial and quality. In temporal scalability, each subset includes the original video stream with reduced temporal resolution, i.e., frame rate. In spatial scalability, a video stream is represented by subsets of bit streams with reduced spatial resolution, i.e., picture size. Furthermore, the substreams represent a lower fidelity version of the original stream in quality scalability, which is commonly referred to as fidelity or SNR scalability. Besides, region-of-interest (ROI) scalability, object-based scalability and combined scability option, where different types of scability can be combined, are also available [108]. In addition, coarse-grain, medium-grain and fine-grain scalability options are defined in SVC extension of H.264/AVC for quality scalable coding. A few SNR points with at least 50 % bit-rate increase from one layer to the next is selected in coarse-grain scalability. In medium-grain scalability, motion-compensated prediction structure is used with combination of coarse-grain scalability for increasing the granularity of quality scalable coding. Although fine-grain scalability, which enabled byte-aligned NAL unit truncation, provided further improvements in the granularity, it was shown that MGS can also provide similar rate-distortion performance with a significantly reduced complexity [131]. Furthermore, feasibility study of using a fine-grain scalability in wireless video transmission can be found in [117].

Chapter 3

HIGH DATA-RATE VIDEO BROADCASTING OF PRE-ENCODED VIDEO CONTENT

3.1 Introduction

The existing 3G systems are originally designed to provide user-specific data to multiple users simultaneously. The wireless broadcast service on the other hand, is aimed to transmit the same data to all (or a sub-group of paid) customers. A conventional approach would be to guarantee every user within the coverage area, the reception and successful decoding of the transmitted information by limiting the transmission data rate to the channel capacity of the worst user. However, such an approach may under-utilize better channels in the system, and thus may be spectrally inefficient. In order to take advantage of the variations in the channel conditions of the users, it may be preferable to divide the data to be transmitted into a number of parallel streams, and subsequently guarantee a certain data rate to all users via a base layer, while sending the rest of the information via a number of enhancement layers in such a way that only users with better channel conditions are able to receive and decode them. This way, different users perceive the broadcast video signal at different quality levels but the basic quality reception is possible for all users. This solution requires the use of a multi-layered (scalable) video source coding scheme. As mentioned in Section 2.5.2, SVC produces a compressed bit stream which is divided into embedded substreams that can be decoded to produce video with improved quality, or larger frame rate, or larger image size depending on the source coding algorithm in use.

When scalable video coding is employed in wireless video broadcasting, clearly one has to ensure that the base layer is successfully decoded by all of the subscribers since the base layer is independently coded and is necessary for basic quality video reception as well as possible subsequent decoding of the enhancement layers. In other words, the base layer decodability is the necessary requirement for the broadcast coverage. At the same time, the system should also be designed such that enhancement layers are successfully decoded

by as many users as possible so that the average perceived service quality is maximized. This is illustrated for a cellular scenario in Fig. 3.1, where users that are closest to the base station are able to decode the base layer as well as all enhancement layers, while the users that are near the cell boundary can only decode the base layer. Then, the limited physical layer resources of the wireless system need to be divided among the multiple layers of the broadcast video in such a way that both of the above stated goals are achieved simultaneously.

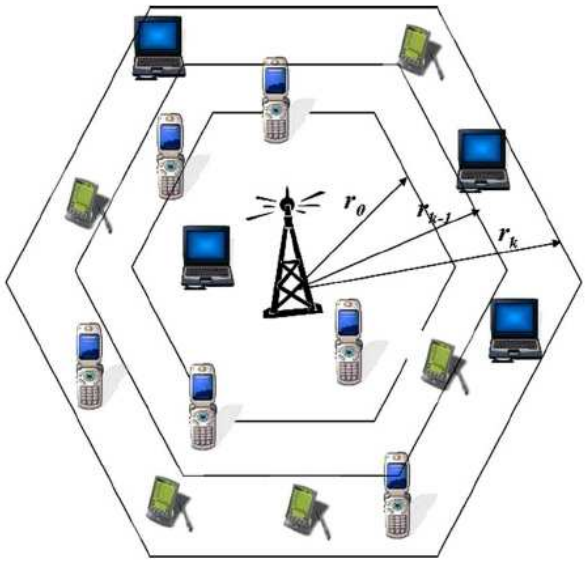


Figure 3.1: Illustration of coverage distances for decodable sub-streams of the video.

However, from an application perspective, scalable coding may not be so desirable in a wireless scenario. This is because, if the base layer data rate is set to be sufficient for an acceptable video quality, the additional benefit from the decoding of enhancement layers may not be significant enough to grant any allocation of system resources for them. If on the other hand, the base layer data rate is not sufficient for acceptable video quality, then stationary users capable of decoding only the base layer will not tolerate this service and thus not use it. The practical benefit of scalable video coding in wireless systems may arise from its potential support for user mobility. Users with high mobility experience rapid variations in their channel conditions. In such instances, unlike non-scalable coding, where the users either receive the full video stream successfully or nothing at all, scalable video

coding may potentially allow for a graceful variation in the observed video quality over time. However, this may be desirable only if significant outage probabilities are not observed as a result of it. In this chapter, we provide a feasibility study of the transmission of broadcast video using multiple layers in wireless cellular systems using 1xEV-DO.

3.2 SVC for Wireless Video Broadcasting

When SVC is employed for broadcast video coding, the wireless system resources of bandwidth, power, time and code need to be efficiently divided among the layers. A number of research results have appeared in the literature on this front. In [50], the authors propose to use unequal error control protection of the SVC layers using rate-compatible punctured convolutional codes. The authors in [63] investigate an on-demand video scenario and the video stream to be transmitted is divided into two distinct layers. Based on the ARQ feedback from the users regarding the success of decoding of video packets, the base layer packets are placed at the front of the transmission queue for retransmission. The enhancement layer packets are retransmitted only if they do not expire during their wait at the transmission queue. In [135], the authors propose to adopt the resource allocation as well as the bit rates of the individual SVC layers based on the changes in the channel conditions for an on-demand video application. The authors propose to use joint source and channel coding so that the base layer and the enhancement layers are channel coded and subsequently multilevel modulated using different parameters in [103] and [13]. The modulation parameters are chosen such that the system provides a significantly stronger error correction capability to the base layer when compared to the enhancement layers. The authors in [103] further propose the use of user feedback to adapt these parameters to changing channel characteristics, user QoS requests and terminal capabilities for on-demand video applications. In [105], the authors characterize the wireless channel using a simple random error model and determine the average broadcast video quality when unequal erasure protection is applied using Reed-Solomon codes. In [58], the transmission power is divided unequally among the layers. Naturally, the base layer is given a larger power allocation than the enhancement layers. [68] on the other hand, proposes to divide the system resources adaptively across multiple video streams based on video content and source coding specifics without paying any attention to the wireless channel.

In all of the papers mentioned above, the division of system resources is done heuristically. In other words, the authors propose to use a pre-determined set of coding, modulation schemes or an unequal power division rule, and compare the performance of their choice relative to a scheme where there is no layeredcoding. The papers do not explicitly specify why such division of resources is proposed and whether it is optimal in any way.

The work conducted in this dissertation differentiates itself from the rest on two fronts. First, here we propose a wireless video broadcasting scheme where the division of resources among the layers of the broadcast video is due to the optimal compromise among the two goals of the system: maximization of the base layer broadcast video coverage and maximization of average total bit-rate decoded by the users, which is proportional to the perceived video quality measured in PSNR [106]. Second, we propose to divide the system resources in the code domain. This way, it is possible to use the same modulation and channel coding schemes to all layers, simplifying the overall hardware design as well as decoding and demodulation complexity. We focus on the problem of providing broadcast video capability to an existing 3G system rather than developing a new end-to-end high power solution like MediaFlo or DVB-H. For a given number of SVC layers, the proposed framework provides the corresponding set of data rates for each layer as well as the code space division for the optimal compromise.

The rest of this chapter is organized as follows: Formulation of the desired objectives used in the resource allocation is given in Section 3.3. The wireless system model and simulation results of the proposed video broadcast scheme are presented in Section 3.4. Finally, conclusions are drawn in Section 3.5.

3.3 Problem Formulation

The proposed low-power broadcast video solution is based on cdma2000 1xEV-DO, which is reviewed in Section 2.3. When 1xEV-DO is used to broadcast a video stream that is scalable coded, the 16 Walsh codes, mentioned in Section 2.3.1, need to be divided optimally among the different layers of the stream. The proposed division of the Walsh codes of the 1xEV-DO system is illustrated in Fig. 3.2. Since multiple streams are code division multiplexed, one needs to make room for CDMA spreading by reducing the repetition coding rate. However, the reduction in the coding should not result in the effective coding rate, described as,

$$\text{Effective Code Rate} = \frac{\text{Turbo Code Rate}}{\text{Repetition Factor}} \tag{3.1}$$

to be greater than 1, since this would mean losing some of the information bits prior to transmission. The reduction of the repetition factor as a function of the number of Walsh codes available per layer is tabulated in Table 3.1 for 1xEV-DO. From the table, it is clear that not all data rates are supportable by all code space allocations. Let us consider, for example the 307.2 kbps transmission rate that uses rate 1/3 Turbo coding. When 2 Walsh codes are available, the repetition factor in the system needs to be 0.25, resulting in an effective coding rate of 4/3 which is not feasible for effective data transmission.

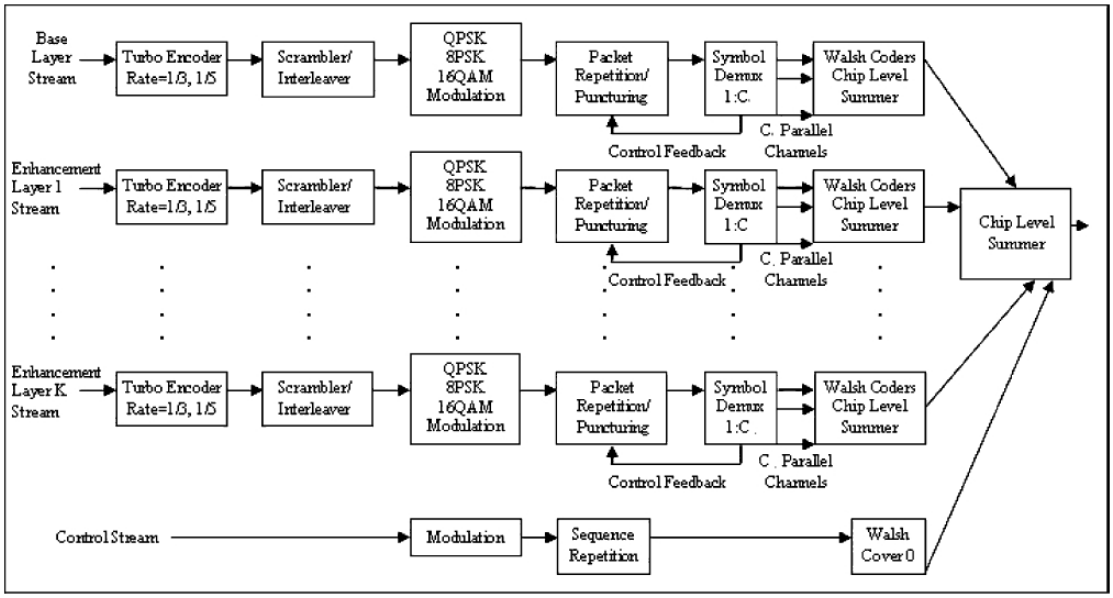


Figure 3.2: Forward Broadcast Traffic Channel Structure using SVC.

In this chapter, we optimally find the set of transmission data rates and the corresponding code space division for a given number of SVC layers to find the best compromise between the maximum broadcast base layer service coverage and average decodable video data rate. There is a direct relationship between the transmission data rate and the PSNR of the received video given a fixed packet error rate. This relationship is dependent on the nature of the source coding algorithm and the error correction and concealment techniques in use. In this part of the thesis, we attempt to optimize the average received video data rate

		Repetition Factor for the Available Number of Walsh Codes														
Data-Rate	Turbo Code Rate	16	14	13	12	11	10	9	8	7	6	5	4	3	2	1
38.4	1/5	9.6	8.4	7.8	7.2	6.6	6	5.4	4.8	4.2	3.6	3	2.4	1.8	1.2	0.6
76.8	1/5	4.8	4.2	3.9	3.6	3.3	3	2.7	2.4	2.1	1.8	1.5	1.2	0.9	0.6	0.3
153.6	1/5	2.4	2.1	1.95	1.8	1.65	1.5	1.35	1.2	1.05	0.9	0.75	0.6	0.45	0.3	0.15
307.2	1/3	2	1.75	1.63	1.5	1.38	1.25	1.13	1	0.88	0.75	0.63	0.5	0.38	0.25	0.13
614.4	1/3	1	0.88	0.81	0.75	0.69	0.63	0.56	0.5	0.44	0.38	0.31	0.25	0.19	0.13	0.06
921.6	1/3	1	0.88	0.81	0.75	0.69	0.63	0.56	0.5	0.44	0.38	0.31	0.25	0.19	0.13	0.06
1228.8	1/3	1	0.88	0.81	0.75	0.69	0.63	0.56	0.5	0.44	0.38	0.31	0.25	0.19	0.13	0.06
1843.2	1/3	0.5	0.44	0.41	0.38	0.34	0.31	0.28	0.25	0.22	0.19	0.16	0.13	0.09	0.06	0.03
2457.6	1/3	0.5	0.44	0.41	0.38	0.34	0.31	0.28	0.25	0.22	0.19	0.16	0.13	0.09	0.06	0.03

Table 3.1: Repetition Factor as a Function of Available Walsh Codes for Transmission.

per user and calculate the corresponding PSNR values for a system that uses PSNR-scalable SVC.

In a wireless broadcast system that uses SVC, the number of users receiving different levels of quality of service depends heavily on the positions of the users over geography as well as their mobility. Since the wireless channel is prone to path loss, shadowing and multi-path fading effects, the system resources need to be carefully managed for efficient service provisioning. For a satisfactory wireless broadcast video service that uses SVC, the number of users capable of decoding at least the base layer and the average observed video quality across users need to be jointly maximized. However, when the system has limited resources these two objectives are contradictory. This is because, when more of the resources are allocated for the transmission of the base layer, fewer resources are left for the transmission of enhancement layers. This results in fewer number of users capable of decoding them successfully, reducing the overall average observed video quality. On the other hand, if fewer resources are used to transmit the base layer, more users will be left with no service at all, but more of the users capable of decoding the base layer will also be able to decode some of the enhancement layers. Thus, an optimal compromise for resource allocation needs to be found.

In the proposed framework, when 1 base layer and K enhancement layers are to be used

for the transmission of the broadcast video, the 16 Walsh codes are to be divided among the $K+1$ layers. We state that the system operates in mode

$$\Upsilon_x = \{(C_0, R_0), (C_1, R_1), ..., (C_K, R_K)\} \tag{3.2}$$

when C_0 code channels are allocated for the transmission of the base layer with a data rate of R_0, and C_i code channels are used for the transmission of the i'th enhancement layer with a data rate of R_i such that

$$\sum_{i=0}^{K} C_i = \begin{cases} 16, & \text{if K} = 0 \\ 15, & \text{if } 0 < \text{K} \leq 14 \end{cases} \tag{3.3}$$

Here, we assume that one code channel out of the 16 available is reserved as the control channel for any mode other than $\{(16, R_0), (0,0), ..., (0,0)\}$.

The control channel is necessary to carry the information that describes which Walsh code channel is reserved for which layer. This information is transmitted continuously so that users joining in at any given time can begin decoding the broadcast video stream. There are a large number of possible modes, Υ_x, $x = 1, 2, ..., X$ for the system operation but not all of these are feasible due to the constraint imposed by the effective coding rate as well as the receiver capabilities. Details on the limit imposed by the receiver is discussed in Section 3.4.1. Through a thorough count, one can easily determine the number of feasible modes of operation for the broadcast system for a given value of K. These are tabulated in Table 3.2 for $K \leq 6$.

Number of Layers	Number of Modes
1	9
2	203
3	1219
4	3313
5	5249
6	5676

Table 3.2: Number of Feasible Modes of Operation versus Number of Layers.

The number of Walsh code channels reserved for each layer and the corresponding data rates directly affect the average number of users that can successfully decode at least the base layer as well as the overall average throughput in the cell. The optimal compromise result for this problem aims to find the appropriate values for $C_0, R_0, ..., C_K, R_K$. Associated with each mode of operation, there exist a set of SNR thresholds, $\gamma_{x,0}, ..., \gamma_{x,K}$, that effectively divide the cell geography into $K+2$ regions as illustrated in Fig. 3.1. When the system operates in mode Υ_x, if a specific user, say user u experiences an SNR value at time slot t, $\gamma(u,t)$, that is less than $\gamma_{x,0}$, then this user is unable to decode even the base layer. If, on the other hand, $\gamma(u,t)$ is between $\gamma_{x,0}$ and $\gamma_{x,1}$, the user can decode the base layer successfully but not any of the enhancement layers. In general, if $\gamma(u,t)$ is between $\gamma_{x,k}$ and $\gamma_{x,k+1}$, user u can decode the base layer and the first k enhancement layers. Let us assume that the broadcast system outage is defined as the percentage number of time slots where a user cannot successfully decode even the base layer video stream averaged across the user population. Assume that M users, uniformly distributed within the cell, are capable of decoding different levels of data according to the data rate matrix $\Omega_{\Upsilon_x}(u,t)$ where Υ_x is the system operation mode, u is the user index and t is the time slot index. The data rate matrix is computed as follows:

$$\Omega_{\Upsilon_x}(u,t) = \begin{cases} 0, & \text{if } \gamma(u,t) < \gamma_{x,0} \\ R_0, & \text{if } \gamma_{x,0} \le \gamma(u,t) < \gamma_{x,1} \\ \vdots & \\ \sum_{i=0}^{K-1} R_i, & \text{if } \gamma_{x,K-1} \le \gamma(u,t) < \gamma_{x,K} \\ \sum_{i=0}^{K} R_i, & \text{if } \gamma(u,t) > \gamma_{x,K} \end{cases} \tag{3.4}$$

Then, the first objective is

$$\min_{\Upsilon_x} \left(\sum_u \sum_t \delta_{\Omega_{\Upsilon_x}(u,t),0} \right) \tag{3.5}$$

where

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \tag{3.6}$$

Similarly, the average decodable video stream data rate for the system can be calculated for a specific mode Υ_x as follows:

$$\overline{R}_{\Upsilon_x} = \frac{1}{MT} \sum_{u=1}^{M} \sum_{t=1}^{T} \Omega_{\Upsilon_x}(u,t) \tag{3.7}$$

where M is the number of users, and T is the duration of the video broadcast.

Thus, the second objective is:

$$\max_{\Upsilon_x} \left(\overline{R}_{\Upsilon_x} \right) \tag{3.8}$$

As stated before, the objectives of (3.5) and (3.8) are contradictory and thus cannot be simultaneously satisfied. We resort to the framework of multi-objective optimization, which is summarized in Appendix B, to find the best compromise operating point in the Pareto-optimal sense for these two objectives. In the proposed broadcast video system, we set the importance weights of the two objective functions at $w_1 = w_2 = 1$. To find the best compromise point for a given number of broadcast video layers with and without a constraint on the base layer data rate, an exhaustive search is computationally feasible. This is because, the number of modes of operation is not large for the number of layers considered in the formulation as observed in Table 3.2.

It is also possible to generate a solution that is better than the actual best compromise solution for one objective function, but worse for the others. This actually corresponds to fine-tuning the optimization decisions in favor of a selected optimization criterion along the multiple Pareto-optimal solutions creating the Pareto-surface. For example, we can come up with a solution that provides better base layer video broadcast coverage with lower average throughput and vice versa. Knowing the client preferences, the server side may prefer to skip the original best compromise optimal solution and offer different solutions by utilizing this property as illustrated in Fig. B.1.

3.4 Simulations

Extensive simulations have been conducted to assess the performance of the proposed broadcast video system. Details of the simulation platform are given in the next section. Results are then presented in Section 3.4.2 comparing non-SVC transmission to SVC transmission with different number of layers. Sensitivity of the system performance when the operating point deviates from the optimal one is discussed in Section 3.4.3.

3.4.1 Simulation Platform

The simulations are composed of three stages:

1. System Level Simulations

2. Physical Layer Simulations

3. SVC Simulations

System level simulations model a 3-tier cellular layout with hexagonal cells having a maximum cell radius of 1000 m. Here, the 3 tiers have 6, 12 and 18 cells around the cell of interest, respectively. For the simulations, a minimum distance between a base station and a mobile is set to be 35 m. The base station transmission level is set to 40 dBm and each base station in the system is assumed to use the same frequency (2 GHz). It is assumed that this transmission level encompasses all other components of a link budget such as antenna gain, etc. The simulation only considers users in the center cell.

In the simulations, we drop 32 mobiles uniformly into the center cell. The simulations are divided into 60 second simulation time blocks. Each mobile is randomly repositioned every time block, simulating movement within a cell. The resulting position of the mobile is used in the path loss calculations of the signals from each of the base stations in the system. Fifty time blocks are used in the simulations resulting in 3000 seconds of broadcast simulation time. The same velocity is used for all the mobiles and is assumed constant for the entire simulation. The system is assumed to operate in an urban environment and path loss, shadow fading, multipath fading and mobility are taken into account using the ITU IMT-2000 channel models [54]. We consider two scenarios from [54], namely, the Pedestrian A and Vehicular B channels, where user velocities are 3 km/hr and 120 km/hr, respectively. Many of the urban environmental factors such as building and base station heights, building separation distances are included in the calculations in these models.

In the simulations, we first calculate the path loss for each terminal and base station pair using the path loss models given in [54] for both of the channel models considered. This gives a set of path loss values for each user with respect to the center cell base station as well as to each of the outer cell base stations. A calculated mean received signal power from

each base station based on the path loss model is used for the duration of the simulation time block. A detailed description of the ITU IMT-2000 path-loss model is given in Section 2.4.1.

Shadow fading is modeled with a Log-Normal distribution with zero mean and standard deviation of 4.3 dB. The distribution has a correlation parameter that is a function of the mobile velocity as well as the geographical distance between the transmitter and the receiver [41]. In the simulations we assume that shadow fading is slow and thus can be modeled to be constant within 0.5 sec. intervals. Details of the shadow fading simulation can be found in Section 2.4.2.

Pedestrian A		
Tap	Relative Delay (ns)	Average Power (dB)
1	0	0
2	110	-9.7
3	190	-19.2
4	410	-22.8
Vehicular B		
Tap	Relative Delay (ns)	Average Power (dB)
1	0	-2.5
2	300	0
3	8900	-12.8
4	12900	-10
5	17100	-25.2
6	20000	-16

Table 3.3: Pedestrian A and Vehicular B Tapped-Delay Line Parameters.

Multipath fading is modeled with a Rayleigh distribution having a Doppler power spectrum modeled using Clarke's scattering model. Unlike shadow fading, multipath fading is faster and is re-computed for each time slot (= 1.67 ms). The effects of the wide-band channel is characterized using a tapped-delay line model for the channel impulse response

[54]. The ITU IMT-2000 channel models provide the number of taps, the time delay relative to the first tap and the average power relative to the strongest tap. These values are tabulated for the Pedestrian A and Vehicular B environments in Table 3.3. Multipath fading simulation details are given in Section 2.4.3.

The system under consideration is for broadcast applications. Therefore, all of the base stations in the cellular layout will be transmitting the same signal at the same time. Then, unlike the traditional 3G systems, macro diversity techniques may be employed to enhance the received SNR of the mobile terminals. A macro diversity system serves a mobile simultaneously using several base stations as illustrated in Fig. 3.3. The mobiles, then, will be able to capture the strongest resolvable paths from the closest base station as well as from the neighboring base stations and combine them using a RAKE receiver. In the simulations, it is assumed that the mobile has a RAKE receiver with ten fingers that lock on to the ten strongest resolvable paths. The receiver is assumed to be able to resolve between two paths that have a delay that is more than a chip time ($\simeq$ 814 ns). The paths that are not resolvable are assumed to contribute with a ratio based on the delay difference. The remaining resolvable paths contribute to the interference at the receiver. In the simulations, MRC is assumed so that the received SNR is the sum of the SNR's of the ten fingers of the RAKE receiver.

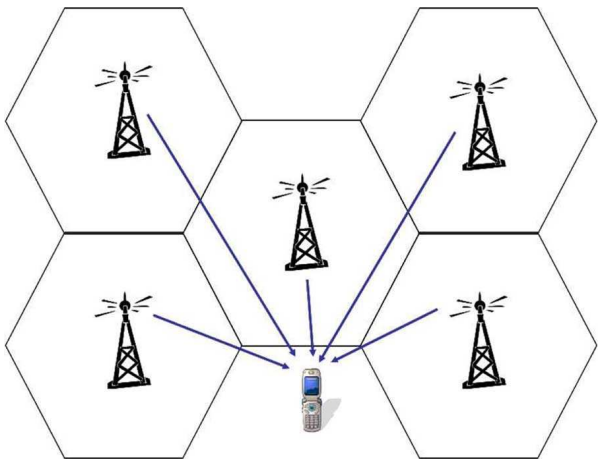


Figure 3.3: Macro diversity with maximal ratio combining for the broadcast system.

The maximum SNR achievable in the mobile receiver is limited by several sources, including inter-chip interference induced by the base-band pulse shaping waveform, radio noise loor, A/D quantization error and adjacent carrier interference. In the system level simulations, the maximum achievable SNR value for the mobiles is assumed to be 13 dB [91]. Then, the effective SNR is given by

$$SNR_{effective} = \frac{1}{\frac{1}{SNR_{combined}} + \frac{1}{10^{13dB/10}}} \tag{3.9}$$

where $SNR_{combined}$ is the instantaneous SNR after MRC.

In summary, system level simulations provide the effective SNR values for each of the 32 mobiles for each 1.67 ms long time-slot in the 3000 sec of simulation time. Once system level simulations are complete, physical layer simulations are conducted to assess what data rates each mobile can accommodate at each time slot.

For a given transmission scheme of the 1xEV-DO system corresponding to a specific transmission data rate, the simulation of the physical layer is performed by processing 150000 randomly generated packets through the transmitter, AWGN (additive white Gaussian noise) channel and the receiver. All of the blocks illustrated in Fig. 3.2 are taken into account in the generation of the transmitter packets. The packets resulting from the receiver are then compared to the transmitted packets for the calculation of the PER. In the simulation, the channel noise power is swept in order to determine the SNR value that results in a 1% PER. The simulation is performed for each combination of data rate and Walsh code division. The results are tabulated in Table 3.4. The "N/A" table entries correspond to modes of operation that are not feasible. The infeasibility stems from either the effective coding requirement discussed in the previous section or the fact that such a mode requires an SNR greater than 13 dB which is not achievable due to eqn. (3.9).

Once both the system level and physical layer simulations are complete, it is possible to find the base layer outage and average decodable video data rate values for each broadcast system operating mode of eqn. (3.2) over the 32 users receiving the broadcast signal for a period of 3000 sec.

	Number of Walsh Codes Available														
Data-Rate	16	14	13	12	11	10	9	8	7	6	5	4	3	2	1
38.4	-11.68	-11.31	-11.11	-10.83	-10.48	-10.26	-9.78	-9.29	-8.72	-7.97	-7.48	-6.11	-4.8	-3	-0.68
76.8	-9.3	-8.73	-8.51	-8.01	-7.64	-7.48	-6.72	-6.2	-5.82	-4.72	-3.68	-2.96	-2.38	-0.63	3.79
153.6	-6.16	-5.8	-5.48	-4.8	-4.15	-3.74	-3.3	-2.95	-2.85	-2.35	-1.69	-0.71	0.88	3.86	N/A
307.2	-3.91	-2.82	-2.31	-1.88	-1.51	-1.25	-1.1	-0.87	-0.16	0.66	1.79	3.59	7.18	N/A	N/A
614.4	-0.91	-0.21	0.27	0.62	1.18	1.8	2.59	3.6	4.99	7.14	N/A	N/A	N/A	N/A	N/A
921.6	1.65	2.7	3.19	3.65	4.29	5.1	6.2	7.79	10.65	N/A	N/A	N/A	N/A	N/A	N/A
1228.8	3.79	4.69	5.28	5.88	6.63	7.56	8.95	11.25	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1843.2	7.8	10.5	12.63	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2457.6	11.3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 3.4: Required C/I Values for the 1XEV-DO Data-Rates for 1% PER as a Function of Number of WALSH Codes Available.

3.4.2 Results

We perform multi-objective optimization to find the best compromise operating points, Υ_x, for the proposed 1xEV-DO based video broadcast system using different number of SVC layers. The optimization jointly considers the two objectives: (3.5) and (3.8), namely, maximization of the base layer video coverage and maximization of the average user decoded video data rate. The limited system resources are divided among the base and enhancement layers optimally to find the best compromise.

An unconstrained optimization naturally results in the best compromise operating point. However, it may yield operating points that are not practical, in that, the base layer data rates may be unacceptably low for these points. To ensure a respectable basic quality for all users, it may be necessary to place a constraint on the minimum base layer data rate. Depending on the aimed application, terminal resolution and the specific SVC coding algorithm in use, this rate may be determined. We investigate scenarios where the minimum base layer rates are chosen to be 307 kbps and 153.6 kbps for the Pedestrian A channel, and 153.6 kbps and 76.8 kbps for the Vehicular B channel. We choose lower data rates for the vehicular channel since it is much more difficult to maintain high rates in wireless systems with high mobility.

Pedestrian A			
Number of Layers	Unconstrained	$BL_{min} \leq 307.2$ kbps	$BL_{min} \leq 153.6$ kbps
1	BL: 614.4 kbps with 16 WC	BL: 614.4 kbps with 16 WC	BL: 614.4 kbps with 16 WC
2	BL: 614.4 kbps with 12 WC	BL: 614.4 kbps with 12 WC	BL: 614.4 kbps with 12 WC
	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC
3	BL: 307.2 kbps with 6 WC	BL: 307.2 kbps with 6 WC	BL: 307.2 kbps with 6 WC
	EL: 307.2 kbps with 6 WC	EL: 307.2 kbps with 6 WC	EL: 307.2 kbps with 6 WC
	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC
4	BL: 38.4 kbps with 1 WC	BL: 307.2 kbps with 6 WC	BL: 307.2 kbps with 6 WC
	EL: 38.4 kbps with 1 WC	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC
	EL: 38.4 kbps with 1 WC	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC
	EL: 614.4 kbps with 12 WC	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC
5	BL: 38.4 kbps with 1 WC	BL: 307.2 kbps with 6 WC	BL: 307.2 kbps with 6 WC
	EL: 38.4 kbps with 1 WC	EL: 307.2 kbps with 6 WC	EL: 307.2 kbps with 6 WC
	EL: 38.4 kbps with 1 WC	EL: 76.8 kbps with 1 WC	EL: 76.8 kbps with 1 WC
	EL: 307.2 kbps with 6 WC	EL: 76.8 kbps with 1 WC	EL: 76.8 kbps with 1 WC
	EL: 307.2 kbps with 6 WC	EL: 76.8 kbps with 1 WC	EL: 76.8 kbps with 1 WC
6	BL: 38.4 kbps with 1 WC	BL: 307.2 kbps with 6 WC	BL: 307.2 kbps with 6 WC
	EL: 38.4 kbps with 1 WC	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC
	EL: 38.4 kbps with 1 WC	EL: 153.6 kbps with 3 WC	EL: 153.6 kbps with 3 WC
	EL: 307.2 kbps with 6 WC	EL: 76.8 kbps with 1 WC	EL: 76.8 kbps with 1 WC
	EL: 153.6 kbps with 3 WC	EL: 76.8 kbps with 1 WC	EL: 76.8 kbps with 1 WC
	EL: 153.6 kbps with 3 WC	EL: 76.8 kbps with 1 WC	EL: 76.8 kbps with 1 WC

Table 3.5: Best Compromise Operating Points as a Function of Number of Layers for the Pedestrian A Channel.

The resulting optimal operating points are tabulated in Table 3.5 and 3.6 for both channels for unconstrained and base layer data rate constrained optimization scenarios. The distances of these operating points from the utopia point and the corresponding base layer video outage and system throughput values are plotted in Figs. 3.4-3.6, respectively.

We have stated that the decoded video data rate and the video quality, measured by PSNR, are related. This relationship depends on the type of video source coding used, as well as the error correction and concealment techniques employed. To assess the results of the multi-objective optimization framework in a practical setting, we conduct PSNR-

Vehicular B			
Number of Layers	Unconstrained	$BL_{min} \leq 153.6$ kbps	$BL_{min} \leq 76.8$ kbps
1	BL: 614.4 kbps with 16 WC	BL: 614.4 kbps with 16 WC	BL: 614.4 kbps with 16 WC
2	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC
	EL: 307.2 kbps with 7 WC	EL: 307.2 kbps with 7 WC	EL: 307.2 kbps with 7 WC
3	BL: 38.4 kbps with 1 WC	BL: 153.6 kbps with 4 WC	BL: 153.6 kbps with 4 WC
	EL: 307.2 kbps with 7 WC	EL: 153.6 kbps with 4 WC	EL: 153.6 kbps with 4 WC
	EL: 307.2 kbps with 7 WC	EL: 307.2 kbps with 7 WC	EL: 307.2 kbps with 7 WC
4	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC
	EL: 153.6 kbps with 4 WC	EL: 153.6 kbps with 4 WC	EL: 153.6 kbps with 4 WC
	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC
	EL: 76.8 kbps with 2 WC	EL: 76.8 kbps with 2 WC	EL: 76.8 kbps with 2 WC
5	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC
	EL: 153.6 kbps with 4 WC	EL: 153.6 kbps with 4 WC	EL: 153.6 kbps with 4 WC
	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC
	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC
	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC
6	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC	BL: 307.2 kbps with 8 WC
	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC
	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC
	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC	EL: 38.4 kbps with 1 WC
	EL: 76.8 kbps with 2 WC	EL: 76.8 kbps with 2 WC	EL: 76.8 kbps with 2 WC
	EL: 76.8 kbps with 2 WC	EL: 76.8 kbps with 2 WC	EL: 76.8 kbps with 2 WC

Table 3.6: Best Compromise Operating Points as a Function of Number of Layers for the Vehicular B Channel.

scalable SVC video coding simulations for the best compromise solutions given in Table 3.5 and 3.6 to find the corresponding PSNR values for the broadcast streams, where WC stands for Walsh codes.We use the standard "Harbour" reference sequence at the CIF resolution of 352x288. This sequence is looped in the simulation so that 150000 physical layer packets constitute the broadcast stream. The SVC software available as the JSVM code at the CVS repository of JVT [61] is used for encoding this sequence with quantization parameters adjusted to yield source data rates as close to the physical layer transmission data rates as possible. The intra period is set to 64 pictures and a GOP size of 16 pictures is chosen.

The video frame rate is chosen to be 15 Hz for the operating points providing a base layer transmission data rate of 614.4 kbps and 7.5 Hz for the operating points providing a base layer transmission data rate of 307.2 kbps.

The video source coder generates video frames whose sizes are not constant and may vary significantly over the course of the video stream. While one frame may fill only a small portion of the physical layer transmission packet, another may occupy multiple packets. In the simulations we assume that video frames are not divided between multiple physical layer packets unless necessary. Therefore some frame fill inefficiencies are unavoidable.

Two types of video frames exist: discardable and non-discardable. The source coder generates base layer frames as nondiscardable and the enhancement layer frames as discardable. The loss of even a single non-discardable frame hurts the video quality, and thus the PSNR value, significantly. Recall that the physical layer of the wireless system operates at 1% packet error rate. Therefore additional protection of the non-discardable frames is necessary to ensure acceptable video quality at the receiver. We employ a simple, rate 1/2 block code to protect the non-discardable frames. No additional protection is provided for the discardable frames. The SVC considered uses a "frame copy" error concealment algorithm [116] where each pixel of the concealed frame is copied from the corresponding pixel of the previous decoded reference frame. Our simulations show that error concealment is very rarely needed when the non-discardable frames are protected by the rate 1/2 block code.

The average PSNR values are calculated in two ways:

1. The received video stream (of 15 Hz frame rate) is compared to the down-sampled version of the original Harbour sequence of 30 Hz frame rate

2. The up-sampled received video stream is compared to the original Harbor sequence of 30 Hz frame rate

These PSNR values correspond to the maximum and minimum achievable levels, respectively. The maximum and minimum PSNR values, as well as the effective video source data rates transmitted are tabulated in Table 3.7 for the best compromise operating points when 1 and 2 layers are employed. When there are 2 layers (1 base layer and 1 enhancement layer), some users will be able to decode both layers, some will only be able to decode

the base layer and the remaining will be able to decode neither. The percentage of users in these categories, their corresponding maximum and minimum PSNR levels as well as their effective video data rates are given in this table as well. The results confirm the direct relationship between the transmission data rate and the video quality observed at the receiver.

Pedestrian A					
Number of Layers	Tx Data-Rate (kbps)	Max PSNR (dB)	Min PSNR (dB)	Video Data-Rate (kbps)	% of Users Decoding
1	614.4	30.2895	27.5481	276.60	99.994 %
2	BL only: 614.4	30.2895	27.5481	276.60	93.155 %
	BL+EL: 768.0	31.1893	28.0926	429.77	88.638 %
Vehicular B					
Number of Layers	Tx Data-Rate (kbps)	Max PSNR (dB)	Min PSNR (dB)	Video Data-Rate (kbps)	% of Users Decoding
1	614.4	30.2895	27.5481	276.60	94.871 %
2	BL only: 307.2	28.9763	25.2487	147.67	94.441 %
	BL+EL: 614.4	31.6747	26.1658	381.09	81.366 %

Table 3.7: PSNR Levels and Effective Video Data-Rates for Best Compromise Operating Points.

From the multi-objective optimization results, we observe that a single layer transmission (which effectively requires a non-SVC video coding algorithm) enables the system to transmit at the 614.4 kbps data rate for both Pedestrian A and Vehicular B channels. The outages observed are 0.006% and 5.129% for these channels, respectively, resulting in average decoded data rates of 614.36 kbps and 582.89 kbps. In other words, when the system broadcasts video using non-scalable coding, the stationary users are almost always capable of successfully decoding the video packets whereas users traveling at 120 km/hr are able to decode them 94.8% of the time. The 614.4 kbps data rate is used to transmit a video data rate of 305.77 kbps since a rate 1/2 block code is used to protect the non-discardable video frames. The corresponding maximum and minimum PSNR levels of the decoded broadcast stream are 30.2895 dB and 27.5481 dB, respectively.

We observe from Fig. 3.4 that the optimal compromise point is reached with two layers

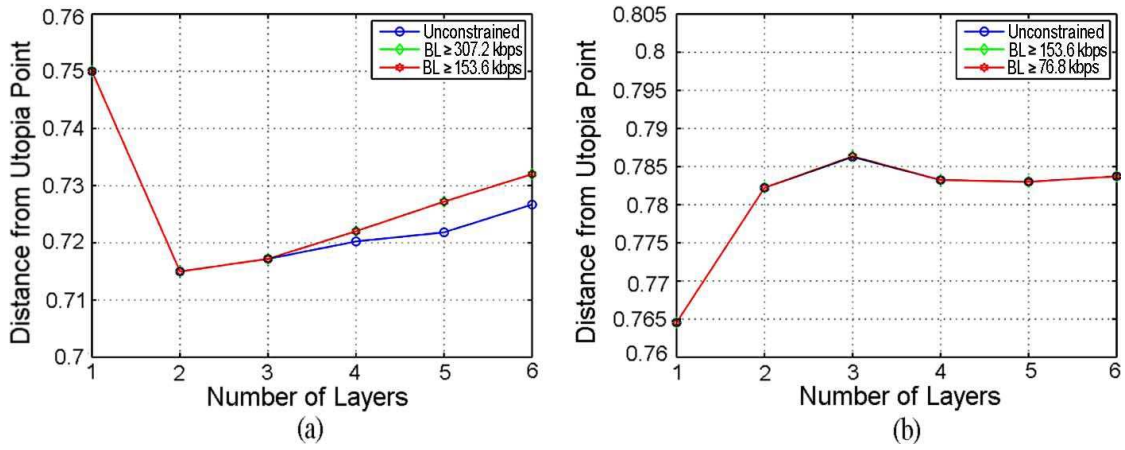


Figure 3.4: Distance from the Utopia Point versus Number of Layers for (a) Pedestrian A & (b) Vehicular B Channels.

for the Pedestrian A channel and with only a single layer for the Vehicular B channel. When we further investigate Figs. 3.5 and 3.6 as well as Table 3.7, we observe for the Pedestrian A channel that while the second layer buys the system 94.13 kbps in the average transmission data rate, and correspondingly a PSNR increase of approximately 0.9 dB for 88.64% of the user population that is capable of decoding both the base and the enhancement layers, this costs an outage increase from 0.006% to 6.845%. In the following section, we investigate whether it is possible to increase the data rate and the PSNR over the single layer transmission while maintaining a reasonable outage level for the Pedestrian A channel. For the Vehicular B channel, on the other hand, we observe that the inclusion of a second layer results in a performance degradation for both objectives. Additionally, the broadcast video stream frame rate needs to be reduced to 7.5 Hz with the inclusion of the second layer since the 307.2 kbps transmission rate for the base layer cannot support the 15 Hz CIF video. As observed in Table 3.7, this results in an additional degradation in the observed PSNR levels relative to the non-layered transmission.

For both types of channels, as the number of layers are increased, the overall performance of the system is reduced since the distance between the best compromise and the utopia points increases. When constraints are placed on the minimum base layer transmission data rate, the deviation from the overall best compromise point becomes more pronounced with increasing number of layers. This is expected because the system resources, code space in

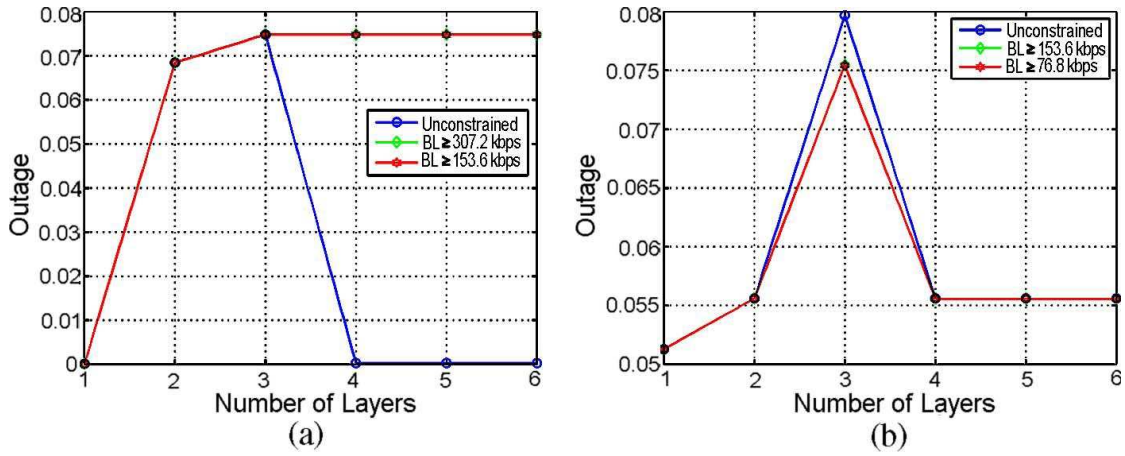


Figure 3.5: Outage versus Number of Layers for Unconstrained and Constrained Optimization for (a) Pedestrian A & (b) Vehicular B Channels.

our case, are limited. When we attempt to divide this limited resource between more and more parallel channels, after a point, the resources allocated for each of the layers become too thin to sustain the system performance over the harsh wireless channel.

Even though the single layer optimal compromise points for both Pedestrian A and Vehicular B channels are the same, we observe that while a second layer increases the system performance in terms of the distance from the utopia point for the Pedestrian A channel, it does the opposite for the Vehicular B channel. The reason for this difference lies in the average values observed by the users in these channels. The 1xEV-DO system provides rate adaptation in the wireless environment through adaptive modulation and coding. However, this adaptation is possible only in discrete steps since only 12 modes with 9 distinct data rates are defined in 1xEV-DO, as mentioned in Section 2.3.1. By intuition, the discrete data rate adaptation incurs a performance loss when compared to continuous data rate adaptation due to the quantization [16].

Then, the system is not always able to distinguish between users that have distinctly different received values. For example, as can be seen from Table 3.4, users with values of -0.91 dB and 1.649 dB are both able to decode a single layer video stream of 614.4 kbps but not the next data rate offered, which is 921.6 kbps. From a resource allocation point of view, the introduction of multiple layers in the transmission allows for finer division of the

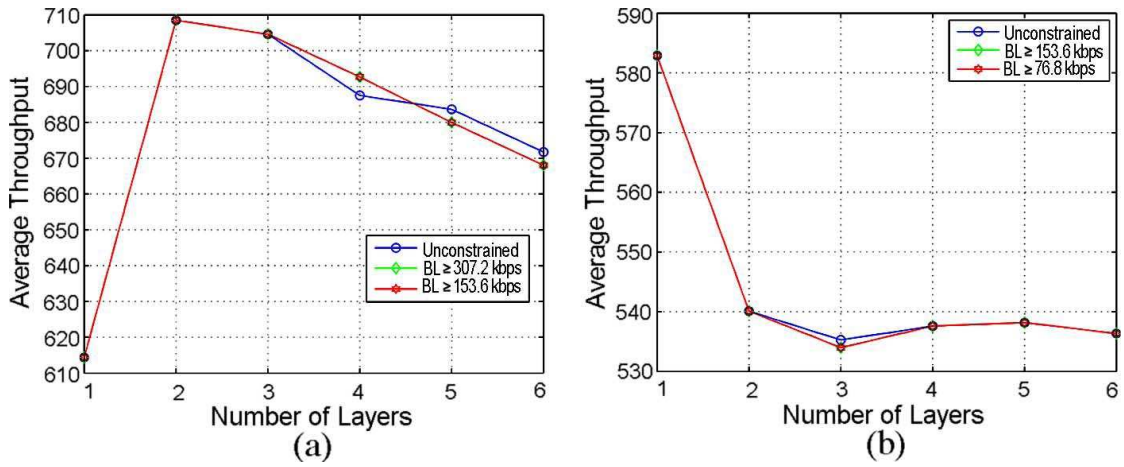


Figure 3.6: Average Throughput versus Number of Layers for Unconstrained and Constrained Optimization for (a) Pedestrian A & (b) Vehicular B Channels.

scale. If, for example, most of the users have values greater than -0.21 dB, from Table 3.4, we observe that we can provide this population the 614.4 kbps video stream using 14 codes, leaving 1 code for transmission of a second layer and still not observe a significant increase in the outage. If the value for most of the users in this population is greater than 0.62 dB, then it is possible to provide the video service at 614.4 kbps using only 12 codes, leaving 3 codes for a second layer. Then, further optimization yields us how best to use these additional codes to increase the average user decoded data rate. While both the Pedestrian A and Vehicular B channels allow for a 614.4 kbps transmission using a non-SVC algorithm, albeit at significantly different outage values, this data rate is supported at drastically different average values of the user population. While the average is close to 0.21 dB for the Vehicular B channel, it is greater than 0.62 dB for the Pedestrian A channel. It is this difference and the 16-step division of the code space in 1xEV-DO that results in the optimal layers of 2 and 1, for the Pedestrian A and Vehicular B channels, respectively.

3.4.3 Sensitivity Analysis

The optimal compromise operating point is reached with two layers for the Pedestrian A channel. However, we observe that this comes at the expense of a significant outage of 6.845%. This may be deemed unacceptably high. Focusing our attention on the scenario

where the base layer minimum transmission data rate is set at 307.2 kbps, we perform a sensitivity analysis. To assess the sensitivity of the objectives of (3.5) and (3.8) to departures from the optimum operating point we first rank all operating points with increasing distances from the utopia point. We conduct the sensitivity analysis for the base layer video outage, and thus we investigate the operating points that are nearest ranked to the optimal point and that have smaller outage values. Obviously if one of these points were to be employed instead of the optimum point, the overall system performance in terms of the distance from the utopia point would be worse. The results, tabulated in Table 3.8, are for two of these points for the Pedestrian A channel operating with two layers. It is observed that if the provider is more interested in reducing the base layer video outage rather than providing a high average video data rate and thus high video quality, it may choose one of these points as the operating point. When the third best operating point is chosen for example, the outage is reduced to 0.329% at the expense of an average transmission data rate drop of 58.95 kbps corresponding to a PSNR drop of 0.2923 dB between the minimum values and 0.4724 dB between the maximum values. However, even with this drop, the third best operating point for this scenario still yields a slight advantage over the single layer operating point which provides an average data rate of 614.36 kbps with maximum and minimum PSNR values of 30.2895 dB and 27.5481 dB, respectively, and an outage of 0.006%.

	Optimal Compromise	Second Best	Third Best
BL	614.4 kbps	614.4 kbps	614.4 kbps
EL	153.6 kbps	307.2 kbps	76.8 kbps
Outage	6.845 %	6.845 %	0.329 %
Throughput	708.49 kbps	658.87 kbps	649.54 kbps
Max PSNR	31.1893 dB	31.9734 dB	30.7169 dB
Min PSNR	28.0926 dB	28.5733 dB	27.8003 dB

Table 3.8: Sensitivity Analysis for 2 Layer Transmission with $BL_{min} \geq 307$ kbps for the Pedestrian A Channel.

In a practical scenario, users traveling at different velocities will be present in the system. While a very modest increase in the system performance is possible with a second layer for stationary users, we observe that the loss this additional layer brings is very pronounced for

the high mobility users.

3.4.4 Generalization of the System Proposal

The simulation results show that optimizing code space among the scalable coded video layers is not beneficial for CDMA 1xEV-DO system. As mentioned, this result depends on the coarse quantization in the physical layer data rates that are supported by 1xEV-DO. However, the idea of optimizing code space at the forward link among the scalable video layers is applicable to any wireless system standard that supports multiple data rates. According to the results, as the number of available physical layer data rates at the wireless system increases the use of SVC may be beneficial. This is because as the number of available data rates increases, the multiple objective optimization may find a better best compromise operating mode where outage is lower and the average throughput is higher.

To illustrate, one can apply the proposed idea to WiMax or DVB-H. In WiMax, various number of physical layer data rates are supported by varying the channel bandwidth, the modulation and the coding scheme, number of subchannels, OFDM guard time and over-sampling rate. A table of supported peak data rates in 5 MHz is given in [34]. Similarly, DVB-H supports a multitude of physical layer data rates. [30] shows the useful net bitrates in 8 MHz channels when FEC for multiprotocol encapsulated data is equal to 3/4, where the supported data rates range from 3.74 Mbps to 23.75 Mpbs. The optimization of the physical layer code space for these two systems among the scalable coded video layers may show that the use of scalable video coding to be beneficial. Therefore, the benefit of using scalable video coding is dependent on the number of supported data rates at the physical layer of the wireless system as well as the separation between them.

3.5 Conclusion

In this chapter, building on the 1xEV-DO system, we propose a novel multi-objective optimization framework for determining the best compromise division of the code space for wireless video broadcasting. The main aim of the proposed algorithm is to provide the best compromise between maximizing the average decoded video data rate and maximizing the geographical coverage area for a basic broadcast video quality. The work differentiates itself from the rest on two fronts. First, here we propose a wireless video broadcasting scheme

where the division of resources is due to the optimal compromise among the two goals of the system. Second, we propose to divide the system resources in the code domain. This way, it is possible to use the same modulation and channel coding schemes to all layers, simplifying the overall hardware design and decoding and demodulation complexity.

The wireless broadcast system benefits from macro diversity across the serving base stations and achieves data rates in the order of 614.4 kbps over the 1.25 MHz channel. While this rate is achieved with almost zero outage for stationary users, high mobility users experience modest outages. The use of scalable video coding is not desirable for this scenario as the use of a second layer provides a very modest benefit for stationary users but causes a significant performance drop for users with high mobility.

Chapter 4

CROSS-LAYER DESIGN FOR WIRELESS VIDEO BROADCASTING

4.1 Introduction

Since existing 3G wireless systems are based on a layered architecture, the traditional approach is to design each layer in isolation from the other layers. Although this layered approach has been accepted and widely used in the past, it presents a suboptimal solution for a wireless multimedia service. It has been recently realized that designing physical (PHY), MAC, transport and application (APP) layers of wireless multimedia systems jointly under a cross-layer framework provides gains not attainable with the isolated layer design [37], [104], [66], [94]. The 3G systems allow for such cross-layer operation thanks to their established feedback mechanisms from PHY to the MAC. It has been shown that it is feasible, and indeed beneficial, to conduct a cross-layer design and operation to provide point-to-point multimedia services, such as on-demand streaming video in 3G systems [94].

In the literature, cross-layer design has been studied in detail to optimize the scarce resources of the wireless environment. The studies focus on approaches where cooperation among different OSI layers provides an improvement to the problem objectives such as QoS guarantees, throughput or video quality maximization, power saving, or delay minimization. According to the works on this front, the main idea is to improve the design objectives in an adaptive manner by exchanging information among various OSI layers. The adaptivity is achieved via the utilization of the feedback information received from the uplink channel. There are numerous articles showing the performance gains of a cross-layer design in different wireless communication systems with feedback mechanisms. Although cross-layer studies include joint adaptation of system parameters among various OSI layers, we want to give brief information about some of these, which can be categorized into three: joint adaptation of PHY & APP layers, PHY & MAC layers and PHY-MAC-APP layers.

In cross-layer designs including the joint adaptation of PHY & APP layers for wireless video transmission, the aim is to improve video quality by adjusting source and channel

coding parameters jointly and according to the channel condition of the mobile users. A general low-complexity analytical channel-distortion estimation model is presented in [47] with a source-coding R-D model to improve the end-to-end video quality by adaptive intra mode selection and joint source-channel rate control. To achieve a certain QoS level, a joint source rate selection and power management for hybrid ARQ/FEC schemes in wireless multimedia multicast is proposed [7]. For example, in the real-time wireless multimedia transmission scenario, [109] proposes a rate-distortion optimized joint ARQ/FEC scheme where APP layer parameters are adjusted according to parameters in PHY layer. Secondly, one can adjust protocols in MAC layer according to the physical layer conditions of mobile users to satisfy a QoS requirement. On this front, [57] combines transport and link layer protocols to guarantee a QoS requirement for voice and data services in wireless cellular networks. Likewise, a cross-layer design between PHY-MAC layers for 802.16e OFDMA system is proposed for efficient resource allocation [69]. Furthermore, adaptation of the APP layer parameters can be incorporated to the cross-layer designs including PHY and MAC layers to improve video quality. To illustrate, [102] proposes a combined APP-MAC-PHY layer design to achieve the optimum rate adaptation from an information theoretical view for real-time video streaming. A similar but more practical approach is taken in [98] for adaptive optimization of cross-layer design parameters. The proposed framework adapts PHY-link-APP layer parameters jointly to achieve QoS requirements for real-time video transmission. In addition, a cross-layer scheduling with joint adaptation of MAC-PHY-APP layers is proposed in [74] to achieve QoS guarantees and efficient bandwidth utilization.

As mentioned, cross-layer studies in the literature are solely based on the wireless communication scenarios, where adaptation of the system parameters are done by exploiting the uplink channel information. Although detailed and extensive works exist in the literature on this front, the applicability of these designs to a wireless video broadcasting scenario poses significant difficulties. The major question is on tuning system parameters via the utilization of the uplink channel information when the communication mode is broadcast (i.e., amount of feedback information is enormous). As a solution, a conventional broadcast system utilizing each of the channel state information, like unicast transmission, may be spectrally inefficient since the broadcast data rate is adapted according to the channel conditions of

the worst user. Additionally, ARQ for retransmission of lost packet to a specific user is not possible in a broadcast scenario as the number of users grows. Therefore, cross-layer studies for cellular wireless broadcasting services are scarce. Since utilization of the feedback information is problematic, cross-layer design for video broadcasting may seem a contradiction to the nature of the cross-layer idea, which utilizes feedback information. Therefore, there has only been a single cross-layer design proposal for cellular wireless broadcast services in the literature until now [56]. In this proposal, the authors propose a novel cross-layer optimization of the forward error correction parameters at the PHY, link, transport and APP layers within the MBMS platform. The authors show that additional FEC on lower layers of MBMS provides performance improvements for real-time applications. In addition, simulations done for H.264 coded video sequences report increase in average received video quality for MBMS over GERAN. Although [56] analyzed the outcome of an optimal FEC selection over different OSI layers, there are no works focusing on the inter-play between the operating SNR region in the PHY layer and the FEC applied in the APP layer for video broadcasting in wireless cellular networks.

The remainder of this chapter is organized as follows. We first discuss the importance of the operating SNR region and the transmitted video content with the contribution of our work in Section 4.2. Then, we formalize the cross-layer problem for wireless broadcast video service and discuss the practical implementations of the proposed framework in Section 4.3. Then, we provide performance evaluations for the proposed system in Section 4.5. Finally, we conclude this chapter by Section 4.6.

4.2 Operating SNR Region & Transmitted Video Content

Since 2G wireless systems are based on providing only voice service to the users, system designs are adapted for improving the quality of this service. After the 3G breakthrough, the multimedia services become available at very high speeds with the channel PER inherited from 2G systems. Although advanced error concealment techniques are developed to obtain a satisfactory video quality at the receiver side, the wireless systems operate at a high PER of 1% [15]. Therefore, the optimality of the PER for a cellular wireless video broadcasting scenario has not been questioned in the literature until now and accepted as 1% since this value is enough to provide an acceptable quality for voice services in 2G and 3G wireless

systems. However, to provide a video broadcasting service through a wireless channel having 1% PER, the transmitter needs to add redundant packets using FEC in order to achieve a satisfactory video quality at the receiver. But, as the applied FEC at the transmitter side increases, it is clear that the effective video data rate, which is directly proportional to the video quality, decreases. To solve this issue, one may change the operating SNR using AMC, so that the PER of the channel is decreased instead of increasing the applied FEC at the transmitter. However, this approach causes a significant coverage drop across the mobile users since users experiencing worse channel conditions may have low SNR than the operating SNR. The optimal compromise for the video broadcasting problem can be found by searching for the optimal operating PER of the system instead of using an optimized PER for vocoders and endeavoring for concealing the data stream received with packet losses. Therefore, the broadcast system has to be designed in a way that service quality is maximized in the average sense. The key challenge in a wireless broadcast problem is to serve as many users while achieving a maximum average video quality at the receiver side simultaneously.

Due to the harsh conditions in the wireless medium, the average video quality at the receiver side is highly dependent on the characteristics of the stream that is being transmitted. Some video content, especially ones including fast motion with high detail, are more vulnerable to packet losses in terms of the perceived video quality at the receiver side. On the other hand, if the video stream that is being broadcasted includes slow motion and/or similarities between consecutive frames, then the error concealment techniques are more likely to correct the part of the frames lost during the transmission. In other words, the video broadcasting system may improve the average video quality by decreasing the operating PER for a video stream with fast motion and/or high detail while it may achieve the similar gain by the use of FEC and error concealment techniques for a video stream with slow motion and/or low detail. According to this observation, the optimal PHY and APP layer parameter set for two video streams having different characteristics will be different from each other. Therefore, the consideration of the content of the video that is being broadcasted may change the optimal solution set for the wireless video broadcasting problem.

The work presented in this chapter differentiates itself from related published work on

three fronts. First, this work questions the optimality of the operating PER for a wireless video broadcasting service. According to our knowledge, the idea of searching for the optimal operating PER in a cellular wireless video broadcasting scenario has not been questioned in the literature until now. Second, it is shown that the optimal APP and PHY layer parameter set is dependent on the characteristics of the transmitted video stream. The questioning of the wireless system operating PER buys additional performance improvements in a content-adaptive video broadcasting optimization. Third, our scheme proposes an initial cross-layer parameter setting for the wireless video broadcasting service prior to the start of the service to achieve the aforementioned goals optimally instead of requiring a necessary change in the existing protocols or system hardware.

In this chapter of the thesis, we propose a novel, content-adaptive, cross-layer optimized framework to improve the performance of video broadcasting using the H.264/AVC codec [81] in a wireless system supporting multiple data rates. In this framework, we set the parameters at APP and PHY layers jointly so that an optimal compromise is reached among the two goals of the system: minimization of the broadcast service outage and maximization of average received video quality, measured in PSNR. In addition, the proposed design can be applied to the existing and future wireless systems since multiple data rate option is available within the existing 3G wireless standards and will be available for the future ones. In order to measure the performance of a wireless video broadcasting scenario, we need to define an outage for the service since a feasible solution to the broadcasting problem would have an outage probability. The outage probability defines the probability of video playback interruption at the receiver in the broadcast service during the transmission. Besides, the received video playback quality at the mobile terminals must ensure satisfiability. However, after some point it is impossible to further reduce outage probability while improving average received video quality, since these goals are conflicting with each other. Since the transmission rate for the broadcasting service is determined according to the channel conditions of all users, the optimality of the system design parameters has to be sought according to the average sense behavior of the whole system.

4.3 Cross-Layer Design

4.3.1 Mathematical Formulation

The cross-layer design problem can be formulated as an optimization problem where the objective is to select an optimal compromise operating point including multiple system parameters. In this chapter, the cross-layer optimization problem includes system parameters from PHY and APP layers. Although the proposed framework includes parameters at PHY and APP layers, extensions can easily be developed to include extra parameters from other layers with the associated constraints. Let n denote the total number of system parameters considered in the cross-layer optimization problem. To clarify, let k and m be the numbers of system parameters at PHY and APP layers, respectively, satisfying $n = k + m$. The optimization problem searches for the optimal n-tuple operating point among possible configurations such as

$$\Delta = \{P_1, P_2, ..., P_k, A_1, A_2, ..., A_m\} \tag{4.1}$$

where P_i and A_i denote the i'th system parameter at PHY and APP layers, respectively. For the sake of simplicity, the subscript Δ refers to the case where the wireless system operates according to a configuration set defined by the n-tuple Δ for the rest of this chapter. Since each of the system parameters is selected among a discrete set, let NPi and NAi denote the cardinality of the discrete sets associated with the i'th system parameter at PHY and APP layers, respectively. Then, there are $N = NP_1\text{x...x}NP_k\text{x}NA_1\text{x...x}NA_m$ possible operating points. Thus, the cross-layer optimization problem at hand seeks to find the optimal operating point among N candidates.

The wireless video broadcasting problem aims to maximize the average received video quality while minimizing the service outage probability at the same time. Firstly, the maximization of the average received video quality is done by averaging the decoded video quality, which is measured in PSNR, across users. Since wireless channel is prone to variations, such as multipath and shadow fading, each user experiences fluctuations in the SNR and some of the packets are lost during the transmission. At the receiver side, the video decoder uses error concealment techniques to retrieve the transmitted video stream. Therefore, the received video quality is different from one user to another according to the channel

conditions of the users.

The first objective, the maximization of the average received video quality, can be represented by the following equation:

$$\max \left\{ \frac{1}{M} \sum_{u} RVQ_{\Delta}(u) \right\} \tag{4.2}$$

where $RVQ_{\Delta}(u)$ defines the video quality experienced by user u and measured in terms of PSNR (dB). To clarify, the average received video quality can be expressed by the concatenated functions below:

$$RVQ_{\Delta}(u) = VD_{\Delta}(WC_{\Delta}(SC_{\Delta}, u)) \tag{4.3}$$

where VD_{Δ} is the video decoder function at the receiver side, which takes the received video packets as the input and returns the video quality as the output. WC_{Δ} is the function that performs packet losses on the transmitted video stream according to the wireless channel conditions of a given user. SC_{Δ} returns the video stream coded by the video coder at the transmitter side. For a feasible video broadcasting session, the transmission rate including the source coding and the FEC is subject to the following constraint:

$$R_{SC_{\Delta}} + R_{FEC_{\Delta}} \leq \overline{R}_{\Delta} \tag{4.4}$$

where $R_{SC_{\Delta}}$ and $R_{FEC_{\Delta}}$ are the data rates required for the encoded video stream by the source coder and the redundant packets created by the FEC technique in use, respectively, and $\overline{R}_{\Delta}$ is the video broadcasting data rate.

Secondly, if user u experiences fading at time slot t, the SNR value, $\gamma_{\Delta}(u,t)$, may drop under a threshold SNR, $\overline{\gamma}_{\Delta}$, for a video broadcasting data rate of $\overline{R}_{\Delta}$. This incident is assumed to be a service outage for user u at time slot t since the effective data rate of user u at time slot t, $R_{\Delta}(u,t)$ is equal to zero. The second objective, the minimization of the service outage probability, can be represented by the equation below:

$$\min \sum_{u} \sum_{t} \delta_{R_{\Delta}(u,t),0} \tag{4.5}$$

where

$$\delta_{i,j} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \tag{4.6}$$

The effective wireless channel data rate is:

$$R_{\Delta}(u,t) = \begin{cases} \overline{R}_{\Delta}, & \text{if } \gamma_{\Delta}(u,t) \geq \overline{\gamma}_{\Delta} \\ 0, & \text{if } \gamma_{\Delta}(u,t) < \overline{\gamma}_{\Delta} \end{cases} \tag{4.7}$$

where $\gamma_{\Delta}(u,t)$ is the SNR experienced by user u at time slot t, and $\overline{\gamma}_{\Delta}$ is the associated threshold SNR value defined by the wireless system to achieve a successful reception at the transmission rate of $\overline{R}_{\Delta}$.

As discussed previously, the cross-layer optimization problem includes two conflicting objectives, (4.2) and (4.5), which cannot be satisfied simultaneously. Therefore, we use the multi-objective optimization framework, explained in Appendix B, to find the best compromise operating point.

4.3.2 Practical Implementations

The proposed cross-layer framework for wireless video broadcasting does not search for the optimal operating mode tuple during the broadcast session dynamically. Since the average video quality is concerned, the optimal configuration set is found prior to the transmission by using the MOO technique. In addition, it is assumed that the video stream is available at the base station before the broadcast session starts. Therefore, this work considers apriori optimization of APP and PHY layer parameters for non real-time video broadcasting as shown in Figure 4.1.

The cross-layer optimization has a broad parameter space and computation of the objective function is required for all candidate parameter tuples. Although the cardinality of the parameter space is large, all of the combinations are not feasible due to the limitations on the system resources. In practice, most of the computation load is occupied by the content adaptation process, which requires a computation time depending on the video characteristics. To decrease the delay incurred by the content adaptation, one may partition the video stream in parts so that future video parts can be analyzed when the current video part is being played. But this solution may result in undesired interruptions during the broadcast

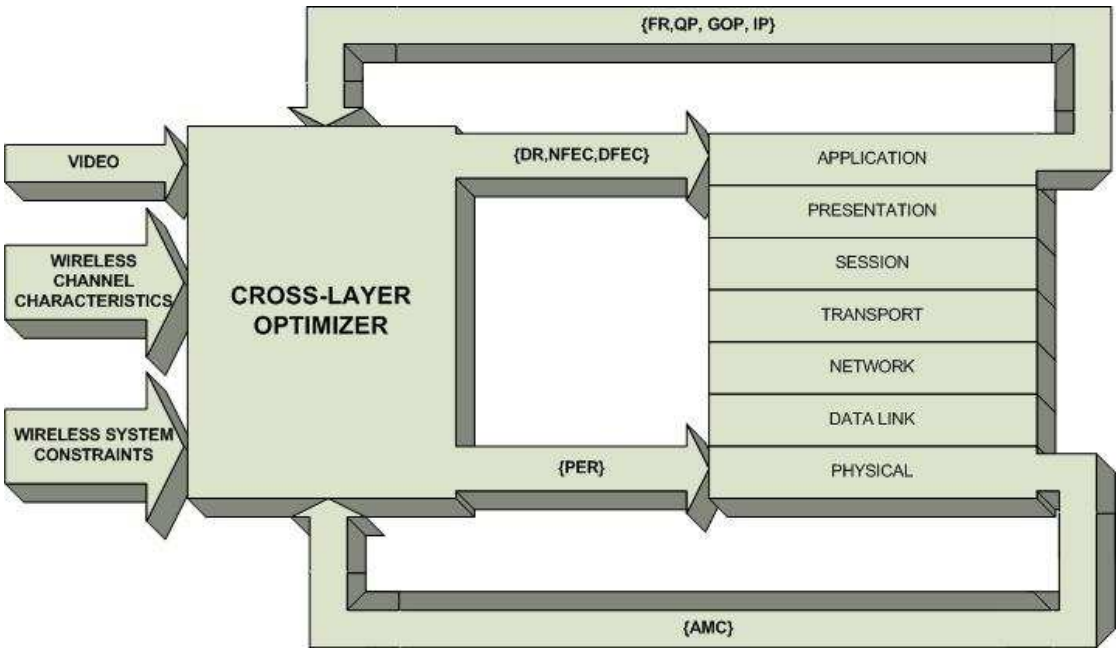


Figure 4.1: An illustration for the proposed cross-layer framework.

session since the processing time for a given video stream is highly dependent on the video content. Therefore, we used video sequences as a whole in our simulations.

The existing 3G wireless standards and future wireless systems not only support various data rate options but also an uplink channel from the mobile user to the base station. The utilization of the feedback information at the base station received via the uplink channel may not be beneficial at all cases. As the number of mobile users receiving the broadcast service grows, the channel information fed back to the base station at each time slot increases as well. In this work, the wireless broadcast service can be seen as a one way communication between the base station and the mobile users since no utilization is performed on the feedback channel information.

4.4 Simulation Platform

Extensive simulations are conducted to assess the performance of the proposed cross- layer multi-objective optimization framework to set the initial system parameters for the wireless video broadcasting service. We use the North American 3G packet data standard, IS-856

[121] in the simulations to provide realistic results. For the H.264/AVC codec, we use the most recent software from JVT available freely [61].

Although selection of the appropriate system parameters for performance optimization makes cross-layer design challenging, quality of the work on a cross-layer design is highly dependent on this criterion. To analyze the performance gains of a detailed cross-layer design for wireless video broadcasting applications, we focused on a variety of system parameters. The PHY and APP layer parameters for the cross-layer design can be listed as follows:

$$\begin{aligned} DR &= \{d_1, d_2, ..., d_n\} \\ PER &= \{p_1, p_2, ..., p_p\} \\ FR &= \{f_1, f_2, ..., f_s\} \\ QP &= \{q_1, q_2, ..., q_t\} \\ GOP &= \{g_1, g_2, ..., g_u\} \\ IP &= \{i_1, i_2, ..., i_v\} \\ NFEC &= \{nf_1, nf_2, ..., nf_y\} \\ DFEC &= \{df_1, df_2, ..., df_z\} \end{aligned} \tag{4.8}$$

where DR defines the data rate supported by the wireless system, PER defines the packet error rate values swept, FR defines the possible frame rate of the broadcasted video, QP, GOP and IP defines the quantization parameter, the group of pictures size and the intra frame period of the H.264/AVC codec, respectively. NFEC and DFEC define the forward error correcting code strength for nondiscardable and discardable video packets, respectively. The details for the nondiscardable/discardable video packets are given in the third stage of the simulations. The simulations are composed of four stages:

1. System Level Simulations provide a detailed wireless channel model, with path loss, shadow fading, and multi-path fading to imitate the mobility characteristics. Path loss is modeled with the COST231-Walfisch-Ikegami [119], shadow fading is modeled as described in [119] using the Gudmundson [41] model with parameters $\zeta_D = 0.3$ for a distance of $D = 10$ m, zero-mean Gaussian variable with $\sigma = 4.3$ dB and a sample rate $T = 0.5$ sec. Multi-path fading is modeled with a Rayleigh distributed as in [119] by filtering two independent white Gaussian noise sources. The effect of the wide-band channel is modeled using two of the ITU-IMT 2000 channel models, namely the

Pedestrian A and Vehicular B models [54]. In system level, we model a 3-tier cellular layout with hexagonal cells. Mobile users are uniformly dropped into the center cell and the resulting SNR values of each terminal are computed for 60 seconds. Then, this procedure is iterated for 50 times to have 3000 seconds of broadcast simulation time. Since the system under consideration is for broadcast service, all of the base stations in the cellular layout will be transmitting the same signal at the same time. Then, macro diversity techniques can be employed to enhance the received signal-to-noise ratio of the mobile terminals [5]. In the simulations, a ten-fingered RAKE receiver with MRC is assumed. The system level parameters are listed in Table 4.1.

Hexagonal cell radius	1000 m
Min. distance between BS and a mobile	35 m
BS transmission level including all components of a link budget	40 dBm
Number of mobile users in the center cell	32
Operating frequency for each BS	2000 MHz
Radiation Pattern	Omnidirectional
Path-Loss Model	Cost231-Walfisch-Ikegami
Log-Normal Shadowing	4.3 dB variance
Multi-Path Fading	Rayleigh distribution by filtering two independent white Gaussian noise sources
Base station height	13 m
Mobile height	2 m
Roof height	12 m
Building separation	20 m
Sampling time	1.67 msec

Table 4.1: System Level Parameters.

2. PHY Layer Simulations determine the minimum SNR levels required to maintain each packet error rate value for the various PHY layer modes corresponding to different

transmission data rates. For this purpose, we used Agilent's Advanced Design System (ADS 2004A) for the PHY layer simulations. The minimum required SNR values to maintain each of the packet error rates are computed for each of the transmission data rates supported by the 1xEV-DO. Once the PHY layer simulations are complete, the results can be incorporated to system level simulations to determine the transmission data rate that can be supported with the corresponding service outage percentage for all users in the cell as a function of time

3. H.264/AVC Video Coding Simulations provide encoded video packets from the NAL for a given video stream. For a given setting of the video resolution, frame rate, GOP size, intra period as well as QP, an encoded video stream is generated by the H.264/AVC codec available in [61]. The H.264/AVC codec tags each encoded video packet as nondiscardable or discardable according to its position in the hierarchical GOP structure. It is clear that losing a nondiscardable video packet instead of a discardable video packet causes more PSNR degradation at the decoded video sequence. To show the content adaptation of the cross-layer framework, we simulated four different reference video sequences, each having different characteristics. First of all, we used "Foreman" sequence since it has a low-motion, low-detail property. Secondly, we used "Soccer" sequence which has a medium-motion in average. Then, "Harbour" sequence is used which has a detailed motion. Finally, "Mobile" sequence is simulated since it includes motion with high-detail in it. Therefore, we can compare the performance of our proposed cross-layer design with traditional approaches for a variety of transmitted video sequence characteristics.

4. Cross-Layer Simulations provide a tool for combining the results of the three simulations described above. It is assumed that the video packets obtained as the output of the video coding simulation are to be broadcasted to the users. The NAL generates video frames whose sizes are not constant and may vary significantly over the course of the video stream. While one frame may fill only a small portion of the PHY layer transmission packet, another may occupy multiple packets. In the simulations, we assume that video frames are not divided between multiple PHY layer packets unless necessary. Therefore some frame fill inefficiencies are unavoidable. During the course

of transmission over the wireless channel we observe that the loss of even a single video frame hurts the video quality, and thus the PSNR value, significantly. Therefore, we applied simple block codes with different strengths to nondiscardable and discardable video packets as FEC. Since the header packets at the start of each video sequence are required for successful decoding at the receiver side, they are protected with a powerful repetition code that almost prevents any erroneous transmission for header files. The encoded video packets are then packed into the PHY layer packets. Transmission over the minimum possible PHY layer transmission-rate is considered. The corresponding average service outage and as well as average received PSNR are calculated after incorporating the results of the system level and PHY layer simulations into the experiment. This procedure is repeated for all settings of the design parameters listed in Table 2. As a result, the values of the two objectives are obtained for all possible values of the design parameters considered. Then, multiple-objective optimization is invoked to determine the best compromise operating point for the proposed cross-layer video broadcast.

	System Parameters	Set
Physical Layer	Data Rate (kbps)	{38.4, 76.8, 153.6, 307.2, 614.4, 921.6, 1228.8, 1843.2, 2457.6}
	Packet Error Rate	$\{10^{-2}, 5.10^{-3}, 2.10^{-3}, 10^{-3}, 5.10^{-4}\}$
Application Layer	Frame Rate (Hz)	{7.5, 15, 30}
	QP	{25, 26,..., 40}
	GOP	{8, 16, 32, 64}
	Intra Frame Period	{8, 16, 32, 64, 128, null}
	FEC List	{1/1, 1/2, 2/3, 3/4, 4/5, 5/6}

Table 4.2: Cross-layer simulation parameters.

4.5 Results

4.5.1 Simulation Results

In order to understand the performance gains achieved by the cross-layer design, we need to compute a traditional system performance for comparison purposes. For this manner, each of four video sequences is encoded with heuristic parameters, i.e., non-cross-layer mode. The GOP size, and the intra frame period are selected to be 16 and 128 respectively, since these values are widely used in the literature due to their good performance in general. We employ a simple, rate 1/2 block code to protect both non-discardable and discardable video packets since the wireless system operates at 1% packet error rate. Then, this configuration is used to encode all the video sequences at each data rate supported by the wireless system.

We perform extensive simulations to determine the cross-layer performance of four standard video sequences. We use each reference video sequence at the CIF resolution of 352x288. A configuration is selected as the optimal mode for a specific PER if it achieves the maximum PSNR among all the others. For the all video sequences, the optimal data rate is found to be 614.4 kbps since it has almost zero outage probability for the Pedestrian A channel while having an acceptable outage probability for the Vehicular B channel among different PER regions. Figure 4.2 compares the PSNR values achieved by the cross-layer design and the traditional approach for all video sequences. We can observe that a significant increase in video quality is achieved with the cross-layer framework as expected. Furthermore, the optimal video configurations for each PER values can be seen in Table 4.3. We observe that as the operating SNR region increases, the optimal video configurations need weaker FEC, meaning that more video information can be sent at the same data rate instead of redundant bits. This can be observed from the QP values since a low QP value at the encoder creates a better quality video as an output. In addition, as PER decreases, intra frames become unnecessary since they are not optimized in the rate-distortion sense, they allocate more space and decrease the performance. The increase in video quality via the use of a cross-layer framework is clearly observed for all video sequences in all PER regions. In "Foreman" sequence, PSNR increases by 1.6 dB at the high PER region, while this increase becomes 2.3 dB at the low PER region. Similarly, in "Harbour" sequence PSNR difference is 1.7 dB and 2.5 dB at high and low PER regions, respectively. In "Mobile" and

"Soccer" sequences these values are 1.5 dB and 2.2 dB at high PER region, 2.6 dB and 3.3 dB at low PER region. It is clear that using a cross-layer design that simultaneously utilizes PHY and APP layer parameters, the average video quality at the receiver side is increased substantially independent of the transmitted video characteristics. As mentioned previously, by lowering the PER of the wireless system, the service outage probability is increased since the operating SNR region is shifted to the right. The change in the outage probability percentage for ITU Pedestrian A and Vehicular B channels at the data rate of 614.4 kbps can be seen in Figure 4.3.

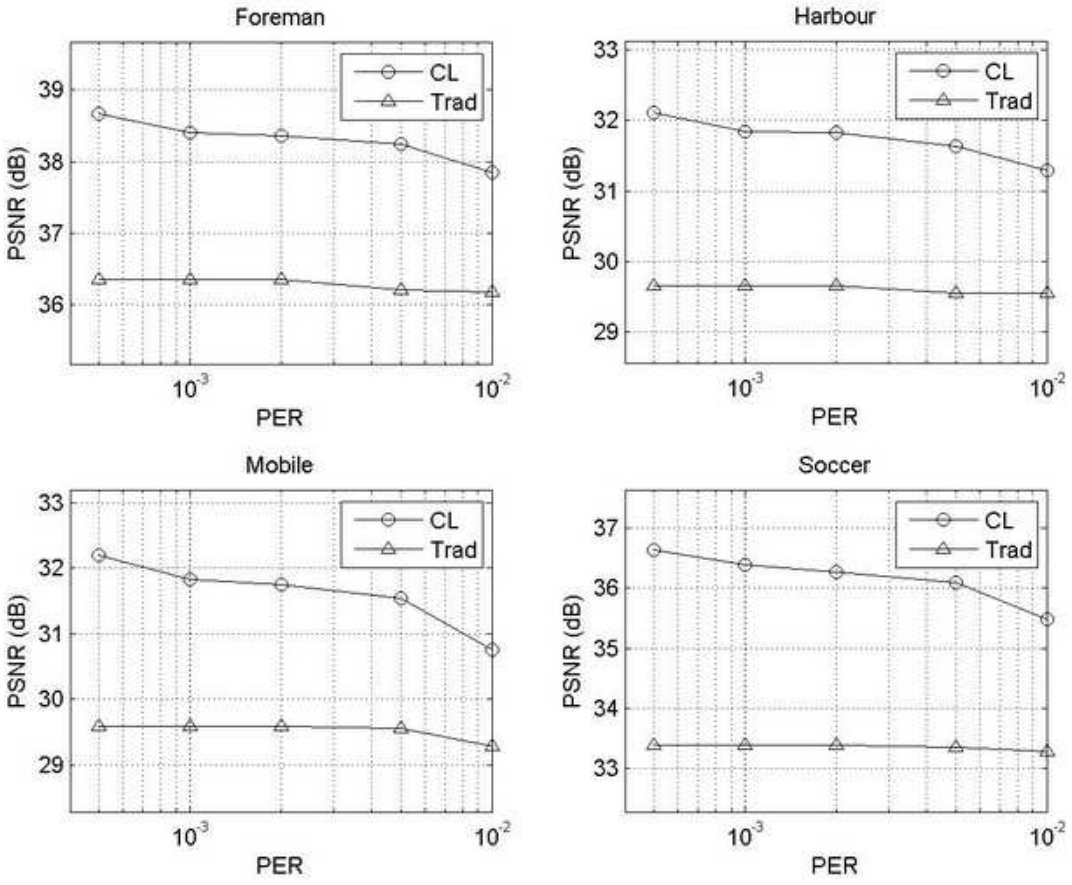


Figure 4.2: PSNR vs PER Plots Between Two Approaches for All Video Sequences.

To observe the actual performance gain achieved via the cross-layer design, we need to invoke MOO framework. For MOO, we first determine the utopia points for each video

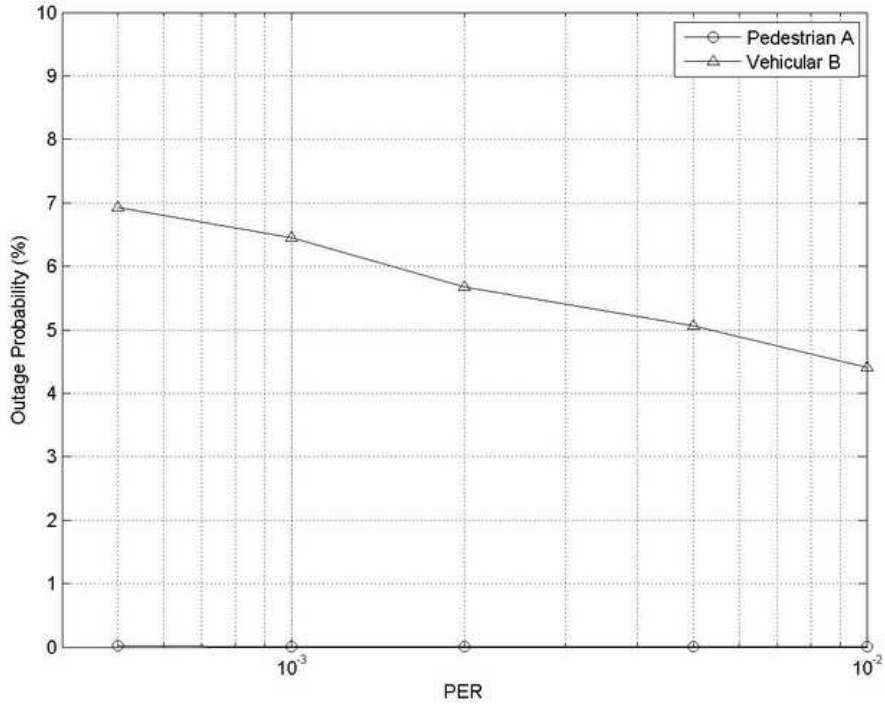


Figure 4.3: Outage Probability vs PER for Pedestrian A & Vehicular B Channels at 614.4 kbps.

sequence used in the simulations. The utopia point has a PSNR value of a video sequence encoded at the maximum data rate supported by the wireless system. In addition, it is assumed that the video sequence is transmitted through an error-free channel so that the outage probability for the utopia point is zero. The maximum transmission data rate for the IS-856 rev. 0 system is 2.4576 Mbps. Since PSNR values above 40 dB do not improve the visual video quality further, 40 db is accepted as an upper bound for the utopia point. Therefore, if one of the resulting PSNR values for the video sequences encoded at 2.4576 Mbps is greater than this upper bound, we set the utopia point PSNR value to be 40 dB. The utopia points for each video sequence are listed in Table 4.4 below.

For MOO, we need to compute distance of each configuration according to the two system goals, maximization of the average received PSNR value and the minimization of the service outage probability. Before using MOO, we need to define upper and lower limits

for each objective function space since this space will be used in the rescaling as mentioned in Appendix B. Therefore, 10% is accepted as the maximum allowable percentage outage probability. In addition, we may define a lower-bound for the PSNR value, i.e., the minimum satisfactory video quality. To analyze the cross-layer performance gains, we conducted MOO techniques for two cases. First, no lower bound is defined for the PSNR and second a lower bound of 29 dB is set. In all cases, MOO is conducted using equal weights for each of the design objectives. The Figures 4.4 and 4.5 show the distances from the utopia point for optimal modes in each PER region for the cross-layer and traditional approaches for "Foreman" and "Soccer" sequences. Similarly, Figures 4.6 and 4.7 draw the same graph for "Harbour" and "Mobile" videos for the above mentioned two cases, respectively. For all video sequences, we observe that the performance difference between the cross-layer and traditional approaches become obvious as we increase the PSNR lower bound. It is expected that the optimal PER region for the pedestrian channel is the lowest possible PER since outage probability is almost the same for all PER values.

For the vehicular case, the results show that the optimal PER region is dependent on the transmitted video characteristics. Although the vehicular channel has significant outage change among all the PER values simulated, as the PSNR lower bound is increased, the optimal PER region changes for the terminals moving at fast speed. This situation can be viewed in the graphs of "Mobile" sequence. The optimal PER region changes from 10^{-2} to 5.10^{-3} for the vehicular case as the PSNR lower bound increases. Therefore, input video content affects the result of the cross-layer optimization according to this observation.

In the "Foreman" and "Soccer" sequences, the optimal PER region for the vehicular case stays at the highest. This is due to the fact that the PSNR lower bound is far away from the achievable PSNR values for these two videos. The graphs of the "Foreman" and "Soccer" show great similarities since these two video sequences have similar characteristics. According to the graphs, at the data rate of 614.4 kbps, "Foreman" sequence can be received with an average PSNR over 38 dB while "Soccer" can be received with an average over 36 dB. Therefore, the specified lower bound on PSNR does not affect the optimal operating PER. This situation is due to the coarse quantization in the supported data rates by the 1xEV-DO system. If fine quantization is available, then optimization may decide on a higher data rate by sacrificing an extra portion on the outage probability. On the other hand, results for the

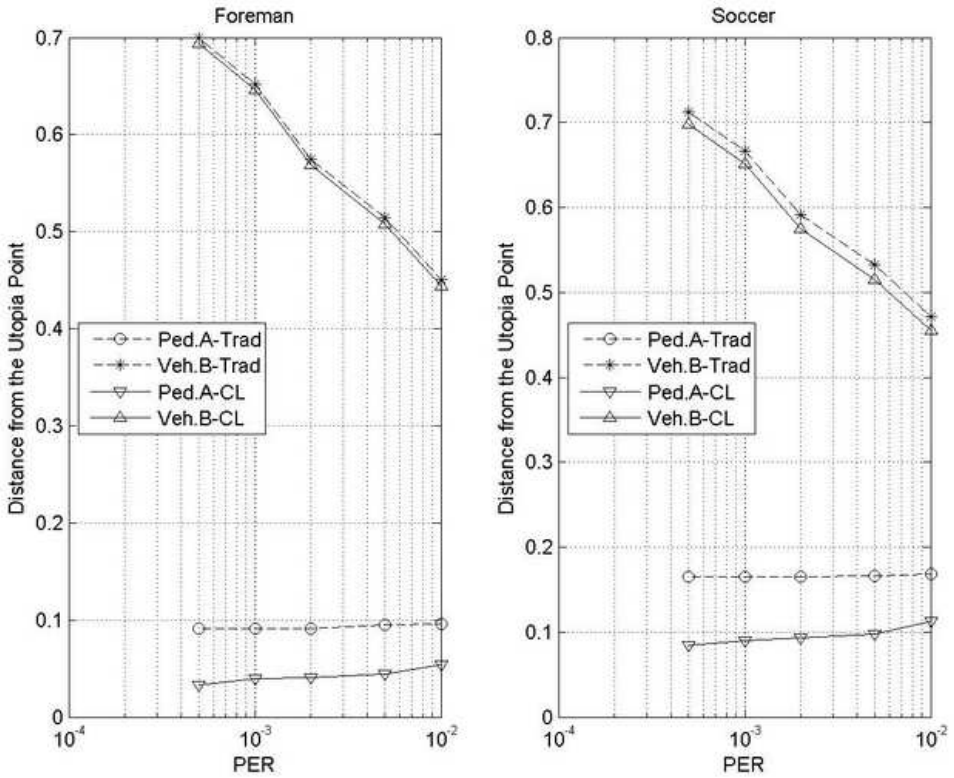


Figure 4.4: Distance from the Utopia Point vs PER plot Between Two Approaches for Foreman and Soccer Sequences (Case 1).

"Harbour" and "Mobile" sequences have similarities due to the same reasoning. Although, optimal PER region for the "Harbour" sequence stayed same, the graph indicates that after an additional increase in the PSNR lower bound, the optimal region will shift towards left. Thus, with the specified lower bound for PSNR, the optimization forces the optimal PER bound to stay at its highest value. These observations indicate the fact that the video content need to be utilized to achieve a further performance gains in cross-layer video broadcasting. The analysis for video sequences with different properties show significant improvements in terms of the average PSNR values at the terminals and service outage probability when cross-layer design is used instead of a traditional approach.

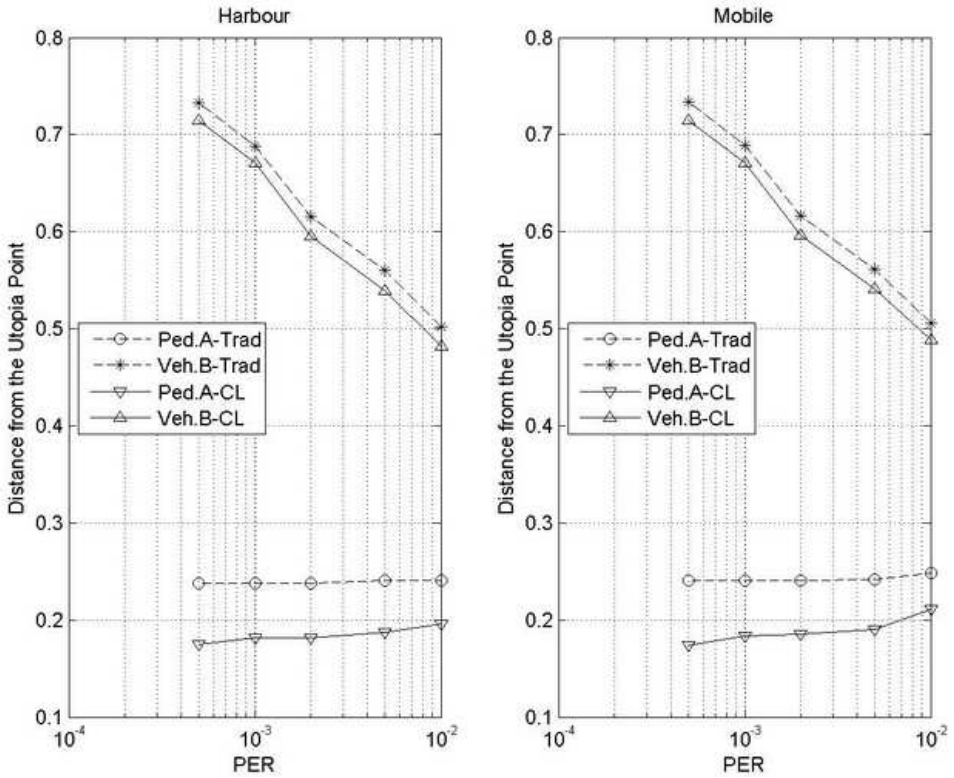


Figure 4.5: Distance from the Utopia Point vs PER plot Between Two Approaches for Harbour and Mobile Sequences (Case 1).

4.5.2 Sensitivity Analysis

We conduct sensitivity analysis for each video sequence and investigate the effect of each parameter to the overall system performance. Specifically, for a given system parameter x, we measure the deviation in distance from the optimal operating mode when all the parameters except x are fixed. In the analysis, we observe that a change in the data rate of the system incurs a significant variation in the distance from the utopia point due to the coarse quantization between the physical layer operating modes of 1xEV-DO. The physical layer data rate can be changed by an increase in the FEC at the application layer or any of the video coding parameters that increase the encoded video data rate. So, the resulting change in the distance may not reflect the sole effect of the selected parameter if the data rate of the system varies. Therefore, physical layer data rate is fixed to its optimal value

when the effect of a system parameter except the data rate is measured.

The analysis show that the most important system parameter is the data rate. As mentioned above, this result is due to the gaps between physical layer data rates defined in the 1xEV-DO standard. For the optimal data rate, 614.4 kbps, the nearest supported data rates are 307.2 kbps and 921.6 kbps. However, switching from 614.4 kbps to 307.2 kbps causes a significant drop in the video quality with a minor improvement in coverage while switching from 614.4 kbps to 921.6 kbps causes a significant drop in coverage with a modest increase in the video quality. Therefore, when the proposed cross-layer design is applied over 1xEV-DO, the system performance is very sensitive to a change in the data rate of the system. The second important system parameter is the operating PER. The results show that users with low mobility are more susceptible to changes in the PER. This is expected because for users with low mobility, the optimal PER is equal to the lowest simulated PER value, 5 times 10^{-4}, while the optimal PER is higher for user with high mobility. Thus, when a system parameter set optimized for low PER forced to operate at high PER, it performs worse than a system configuration, which is optimized for high PER, operating at low PER. Thirdly, the system performance is more sensitive to the application layer FEC applied to nondiscardable video packets than FEC applied to discardable video packets, as expected. Lastly, all the video coding parameters effect the system performance in line with their effect on the video coding performance. However, the effect of FEC and video coding parameters may change according to the video content. In our analysis, their effect on the system performance is similar on the average when all the four video sequences are considered.

4.6 Conclusion

In this chapter, we propose a novel, content-adaptive, multiple objective optimized framework for wireless H.264/AVC video broadcasting using a cross-layer design. The proposed scheme improves the overall performance of the video broadcasting service by finding the best compromise system parameter tuples among PHY and APP layers between maximizing the average received video quality and the geographical coverage. The key contribution of this work is the extensive analysis among system parameters including the wireless system operating SNR region. The simulations based on the North American 3G packet data stan-

dard, 1xEV-DO, draws attention to the performance gains by the cross-layer optimization. Furthermore, the effects of the transmitted video content to the objectives of the wireless system are analyzed.

The simulations show that using a cross-layer design for wireless video broadcasting is necessary and beneficial. Although the performance gain achieved via the cross-layer design varies for videos with different contents, a traditional and heuristic approach for a broadcasting scenario is suboptimal. In addition, the proposed cross-layer design framework can be used to compute a best compromise mode for any wireless system supporting multiple data rates in a broadcasting scenario.

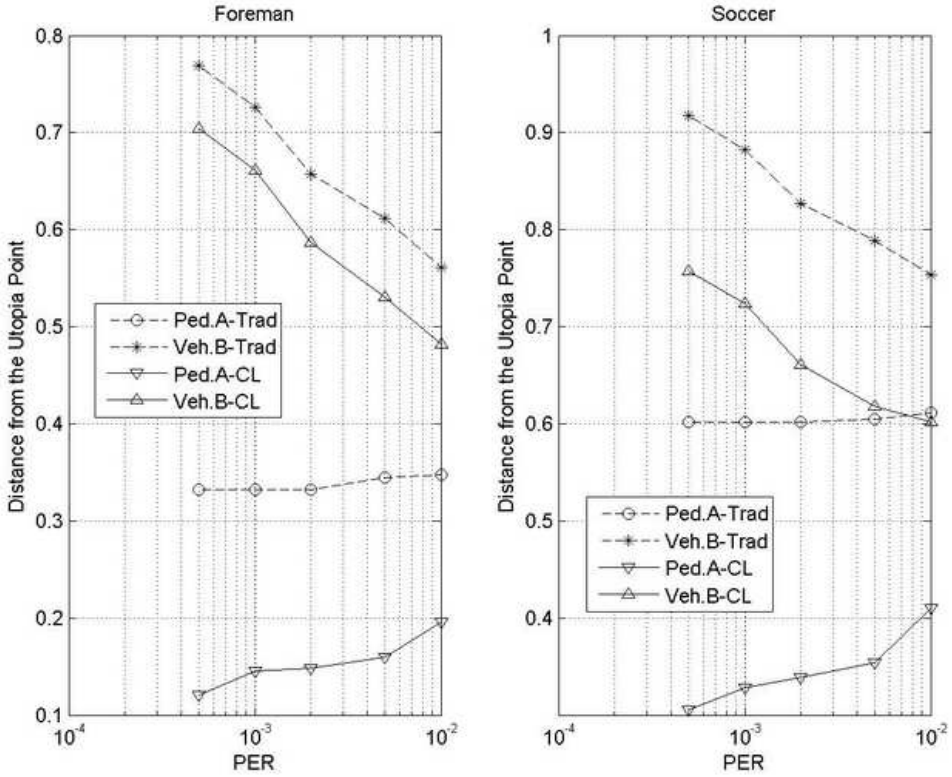


Figure 4.6: Distance from the Utopia Point vs PER plot Between Two Approaches for Foreman and Soccer Sequences (Case 2).

	PER	Frame Rate (Hz)	QP	GOP Size	I Frame Period	NFEC	DFEC
Foreman	10^{-2}	30	28	32	64	2/3	5/6
	5.10^{-3}	30	28	32	128	2/3	5/6
	2.10^{-3}	30	28	32	128	2/3	5/6
	10^{-3}	30	27	16	64	5/6	1/1
	5.10^{-4}	30	27	16	128	5/6	1/1
Harbour	10^{-2}	30	34	32	32	4/5	5/6
	5.10^{-3}	30	34	32	32	4/5	5/6
	2.10^{-3}	30	34	16	Null	4/5	5/6
	10^{-3}	30	34	16	Null	4/5	5/6
	5.10^{-4}	30	33	32	128	5/6	1/1
Mobile	10^{-2}	30	34	32	Null	2/3	5/6
	5.10^{-3}	30	34	32	Null	2/3	5/6
	2.10^{-3}	30	34	32	Null	3/4	4/5
	10^{-3}	30	34	16	128	5/6	5/6
	5.10^{-4}	30	33	16	Null	5/6	1/1
Soccer	10^{-2}	30	31	32	64	5/6	5/6
	5.10^{-3}	30	31	32	128	5/6	5/6
	2.10^{-3}	30	34	32	Null	3/4	4/5
	10^{-3}	30	30	16	128	5/6	1/1
	5.10^{-4}	30	30	16	128	5/6	1/1

Table 4.3: Optimal Configuration Parameters for each PER value.

Sequence Name	Utopia Point PSNR (dB)
Foreman	40
Harbour	38.9023
Mobile	38.9701
Soccer	40

Table 4.4: Utopia Point PSNR Values for each Video Sequence.

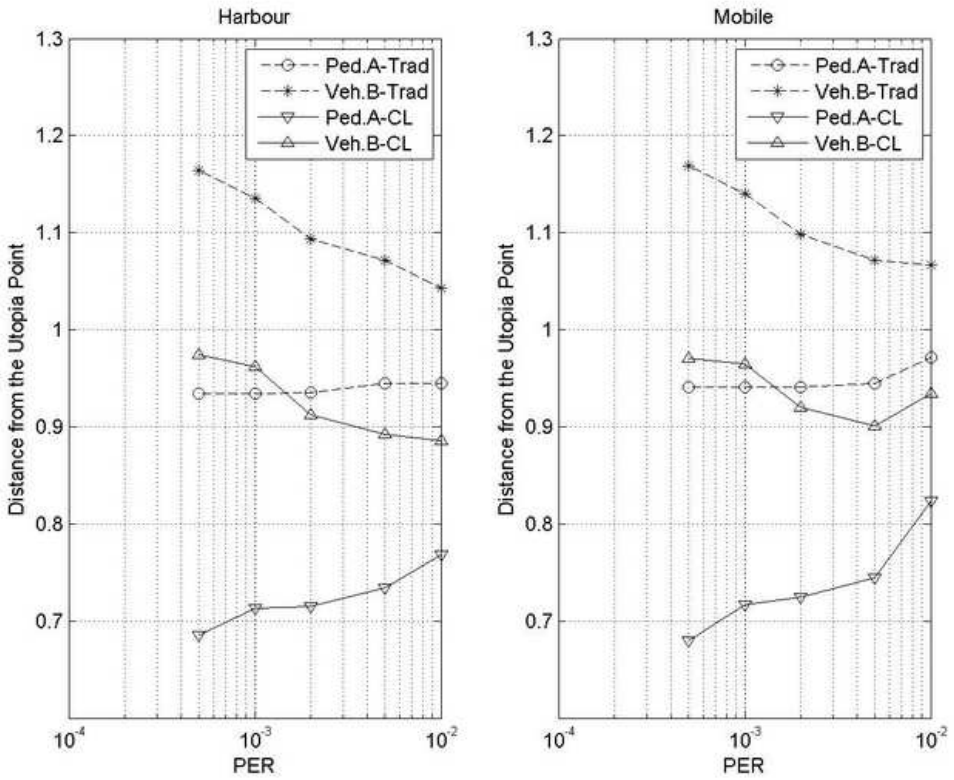


Figure 4.7: Distance from the Utopia Point vs PER plot Between Two Approaches for Harbour and Mobile Sequences (Case 2).

Chapter 5

ACHIEVABLE BROADCAST CAPACITY OF WIRELESS MULTIHOP INTERFERENCE NETWORKS

5.1 Introduction

In recent years, much of the study concentrated on deriving the capacity limits [42], [40] and scaling laws [136], [71] of wireless networks. After the landmark paper of Gupta and Kumar [42], a variety of work [40], [32], [73], [2], [59], [90], [89], [111] has been done on theoretical computation of achievable capacities between source-destination pairs in wireless multihop networks under different assumptions and constraints. Although the capacity limits of unicast connections in a wireless network have attracted much attention, the studies on wireless broadcast capacity limits are scarce. On this front, there are two papers studying the theoretical capacity of wireless networks for broadcasting. The first paper [138] presents a definition of the broadcast capacity and proves that multihop broadcasting is beneficial than single-hop transmission in extended networks while reverse is true for dense networks. The other paper [62] developes bounds for the broadcast capacity of arbitrary connected networks under different channel models and concludes with three power regimes. Although both of these papers focus on broadcast capacity, their results are based on theoretical analysis and asymptotic bounds. However, the computation of optimal broadcast capacity [1] for a given network requires more practical assumptions, such as total energy consumption, interference and the receiver capabilities. Therefore, the optimal broadcast capacity for a given network setting under realistic assumptions needs to be questioned.

In this chapter of the thesis, we present a study questioning how the optimal broadcast capacity of a multihop wireless network behaves with the number of relay nodes under realistic assumptions. The underlying problem is selecting a subset of nodes as relays that achieves the optimal broadcast capacity for a given total energy constraint. However, when

[1]Throughout this chapter, "*optimal broadcast capacity* " stands for the maximum achievable broadcast capacity for a given hop constraint and total energy constraint.

channel is modeled using a simple path-loss model and there are finite users in a network, this problem is shown to be an NP-complete problem [12]. According to our wireless channel model, the computation of the broadcast capacity requires the CSI between each nodes when there are finite users in the network. Therefore, in this work, we study on a wireless network with infinitely many nodes in order to eliminate the need of CSI. Following this analysis, two relay selection methods are proposed with a discussion on their information requirements and broadcasting performances. Throughout the chapter, the broadcast capacity at hand is always associated with a total energy constraint.

The rest of the chapter is organized as follows. In Section 5.2, we summarize existing work on multihop network capacity in the literature. The network model, assumptions and necessary definitions are introduced in Section 5.3. Also, we define the optimal broadcast capacity achieving relay selection problem and our contributions in the same section. Then, capacity analysis is presented under an infinitely many node scenario in Section 5.4. Furthermore, two relay selection methods based on heuristics are proposed in Section 5.5. In order to measure the performances of the theoretical model and the proposed algorithms, simulations are done in Section 5.6. Finally, conclusions are drawn in the last section.

5.2 Existing Work

Wireless multihop networks have gained a lot of interest and have been studied extensively in recent years due to their potential in practical scenarios. Sensor, ad hoc and mesh networks are the examples of wireless networks where a data stream is to be transmitted to a set of nodes. Although there is a wide use of these types of networks in wireless communications, the optimization of system design brings a lot of new challenges. Even a simple wireless channel model including interference presents a lot of difficulties when one needs to find the optimal selection of relay nodes in order to minimize the maximum number of hops and/or total energy consumption while maximizing the minimum data rate over all connections. The number of hops can be minimized because the data can be sensitive to delay between the source and the destination. In addition, battery-life is an important constraint on energy consumption [24], so total energy used for the multihop strategy needs to be minimized. On the other hand, the data rate has to be maximized for QoS, for example in a multimedia broadcasting system data rate is related to the quality of the video.

On this front, there are a variety of works studying the achievable gains of multihop networks and complexity of finding the optimal broadcast capacity achieving relay selection problem.

To start with, the simple idea of cooperative transmission for unicast transmission is defined by [49]. In this work, authors proposed to utilize source-to-destination channel gain together with the relay-to-destination channel gain and thus increase the overall data transmission capacity between the source and the destination. Following this paper, a variety of work has been done in order to utilize spatial diversity gains of relaying and wireless broadcast advantage. Opportunistic large arrays is proposed in [101] where different signals can be exploited by accumulation of their power levels at the receiver in order to reach far distances. Although the diversity gain is not guaranteed by the distributed integrate and fire method, the scheme has low complexity as it eliminates the centralized scheduling of transmissions. [79] focuses on a similar cooperative transmission, accumulative broadcast, where nodes capture more radiated broadcast energy to increase the energy efficiency of the system. Also, the problem of node transmission order is found to be NP-complete and a heuristic algorithm is proposed. Although the analysis in these two works are different, the underlying idea is similar. Both of them tries to utilize the signal powers from different paths in order to increase the coverage or the capacity. The idea proposed in this chapter is also similar since Rake receiver enables utilization of different signal paths via its fingers by using MRC. Besides, a recent work [114] considers the energy efficiency in terms of broadcast energy consumption per bit and models the system as interference free. Closed-form expressions for the energy consumption are derived for both non-cooperative and cooperative transmission models. In addition, physical layer cooperation in a multistage network is studied in [113], where a simple SNR threshold is used to decide whether a node will employ decode-and-forward or not. Although a distributed approach is proposed, the connectivity is achieved when the network density is high. As the analysis is based on a continuum model, closed-form expressions are drawn. Similarly, in this chapter, we analyze the multihop broadcast capacity based on a continuum model and then using the results of this theoretical analysis, we propose heuristics for practical scenarios.

Given a wireless multihop network and a constraint on total energy of the system, the problem of finding the optimal broadcast capacity achieving relays while covering all nodes

is NP-complete. The complexity of this problem for binary transmission level (on/off) was proven to be NP-complete independently in [12] and [97] by reduction from set-cover and dominating-set problems, respectively. Also, [12] proves that the minimum energy broadcast problem is NP-complete in two-dimensional Euclidean metric space. Although no interference analysis is done in these works, a binary transmission level and minimum energy constraint is sufficient to force the problem become NP-complete. Therefore, integrating the interference analysis to this problem increases the computation of the rates between nodes and the overall complexity of the problem.

In the literature, to our knowledge, the selection of optimal broadcast capacity achieving relay problem in an interference included wireless network has not been questioned. However, there are recent works on the scaling laws and bounds for the broadcast capacity under different network models. To start with, [138] defines the broadcast capacity and derives the scaling laws when multipoint relaying is used in extended and dense networks modeled by a Poisson point process. Then, [62] focuses on the broadcast capacity of multihop networks with dynamic power adjustment for different physical layer models. In both of these works, generalized physical model, which is the realistic network model used in this chapter, is considered for deriving bounds of the broadcast capacity. In addition, the protocol model and physical model are also considered in [62] for completeness. Also, broadcast capacity bounds are derived in [83] for spatially correlated and i.i.d. channel models. But in all of these papers, broadcast capacity bounds and scaling laws are based on theoretical assumptions and these works ignore practical considerations, such as CSI availability, the receiver capability and relay selection.

5.3 Multihop Broadcast Capacity

In this section, we present assumptions and necessary elements to define our problem, discuss the important aspects affecting multihop broadcast capacity and summarize the key results.

5.3.1 Assumptions

We assumed all nodes are distributed according to a uniform distribution on a 2-D circular region, $\mathbf{\Gamma}$. Let $\mathbf{Q}$ and r denote the set of all the nodes in the network and the radius of the circular region, respectively. In addition, $\mathbf{S}$, $\mathbf{N}$ and $\mathbf{M}$ denote the source node, the set of

relay nodes, and the set of nodes that are not transmitting, respectively. Then, the relation between these sets can be written as, $\mathbf{Q} = \mathbf{S} \cup \mathbf{N} \cup \mathbf{M}$, where $\mathbf{S} = \{0\}$, $\mathbf{N} = \{1, ..., n\}$, $\mathbf{M} = \{n+1, ..., n+m\}$.

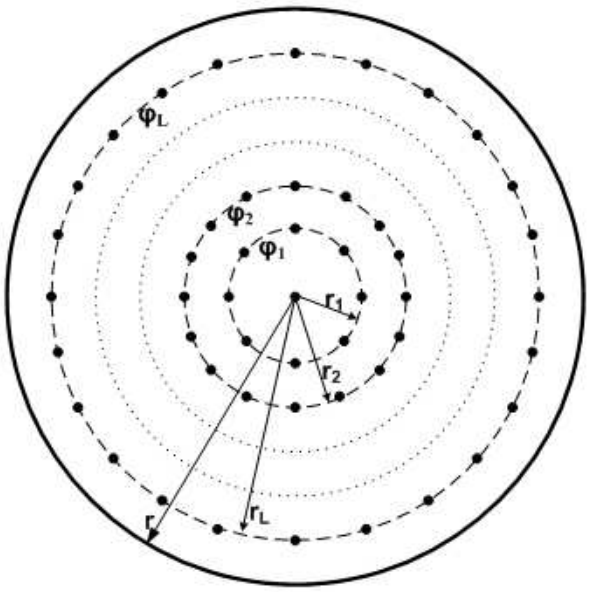


Figure 5.1: Relays Forming a Set of Rings.

Definition 1 (Point Symmetry) *For a geometrical figure in* $\mathbb{R}^2$*, if there exists a point bisecting every line segment passing through it, then the figure is said to have* point symmetry*, and the point is called the* center of symmetry*.*

Following the point symmetry definition, it is obvious that a circle is symmetrical about all of its diameters, i.e., it has a point symmetry about its center. Assume the source node is placed at the center of the region $\mathbf{\Gamma}$. Due to the circular symmetry, the optimal locations of the relays form a set of concentric circles $\Phi = \{\varphi_i \mid i = 1, ..., L\}$ centered around the source node with associated set of radii $\mathbf{D} = \left\{ \overset{\star}{r}_i \mid i = 1, ..., L \right\}$ as shown in Figure 5.1. Although the symmetry of the circular layout simplifies the calculations, the analysis done throughout this chapter can be applied to any form of network topology.

The following assumptions are made:

1. The source node has transmission power P while there is total power constraint, TP, on all cooperating relay nodes in the system. However, no power adaptation is used, i.e., each cooperating relay node transmits at full power, which is equal to TP/n.

2. The network includes $m+n+1$ nodes, which are the source and the intended receivers of the broadcast session. To clarify, the broadcast data has to be delivered successfully not only to the receiver nodes, but also to the relay nodes as well. Therefore, relay nodes can be seen as receiver nodes that are assisting to the source node during the broadcast.

3. As a capacity measure, the famous Shannon capacity formula including the interference from unintended transmitters is used throughout the chapter. Although the capacity analysis does not include the effects of multipath and shadow fading, it includes the effects of path-loss and interference.

4. Since the source node broadcasts data to all nodes, the achievable broadcast capacity is defined by the bottleneck channel capacity among all users.

5. No cooperation is employed at the physical layer, only the wireless broadcast advantage is used.

6. For practical considerations, it is assumed that each of the nodes is equipped with a coherent Rake receiver with R fingers. Consequently, Rake can enable achieving higher SINR by using MRC.

7. Throughout the chapter, broadcast capacity analysis is done for one-ring scenario, i.e., $|\Phi| = 1$ with the optimal broadcast capacity achieving radius $\overset{\star}{r}_1$.

8. To provide full connectivity of the network, all relay nodes lock one of their fingers on to source's signal.

9. All the relay nodes perform DF strategy and wait for a symbol time to decode the received signal. Thus, processing time at relays is assumed to be one symbol time.

10. If there exists at least three points that are not lined up, then Voronoi tessellations can be created. When there are two relays in the network, these relays and the source are positioned along a straight line, thus there are at least three relays in the analysis throughout the chapter.

5.3.2 Broadcast Capacity

Definition 2 (Broadcast Capacity) *The maximum rate at which the source and n relay nodes can broadcast the data to all other nodes such that all $m+n$ nodes (both relays and receiver nodes) can decode the received data correctly is called the* broadcast capacity *(C_B).*

The broadcast capacity can be formulated as the minimum of the channel capacity over a set of transmitters $\mathbf{T}$, i.e., the source and the relay nodes, over a set of receivers $\mathbf{R}$, i.e., all the nodes except the source node, where $\mathbf{T} = \mathbf{S} \cup \mathbf{N}$ and $\mathbf{R} = \mathbf{N} \cup \mathbf{M}$ in the region $\mathbf{\Gamma}$.

$$C_B(\mathbf{\Gamma}) = C_B(S,R) = \min_{i \in R} \left[B \log_2(1+\Psi_i)\right] \tag{5.1}$$

$$C_B(S,R) = \min\left[C_B(S,N), C_B(S,M)\right] \tag{5.2}$$

where Ψ_i denotes the SINR of user i and B is the channel bandwidth.

5.3.3 Voronoi Tessellations

Voronoi tessellations [22] divide the n-dimensional space into a set of nonoverlapping regions according to a given set of points using Euclidean norm. Since we are interested in $\mathbb{R}^2$, two-dimensional Euclidean norm is considered.

Definition 3 (Voronoi Region) *Given a set of n distinct points in $\mathbb{R}^2$, $\mathbf{W} = \{w_1, ..., w_n\}$, $\mathbf{V}(w_i)$ denotes the Voronoi region corresponding to the point w_i iff the following condition is satisfied:*

$$\mathbf{V}(w_i) := \left\{\mathbf{x} \in \mathbb{R}^2 \mid |x - w_i| < |x - w_j|, j = 1, ..., n, j \neq i\right\} \tag{5.3}$$

where the set $\{V(w_i)\}_{i=1}^n$ is called the Voronoi tessellation *of $\mathbb{R}^2$, and the points $\{w_i\}_{i=1}^n$ are called* generators.

Definition 4 (ϵ-neighborhood) *Let $(\mathbf{S}, d)$ be a metric space. Then for $\epsilon > 0$, the ϵ-neighborhood of a point p_0 in $\mathbf{S}$ is the set,*

$$\mathbf{N}_\epsilon(p_0) := \{p \in \mathbf{S} : d(p_0, p) < \epsilon\} \tag{5.4}$$

5.3.4 Rake Receiver

In a wireless multihop broadcasting scenario, a receiving node is subject to a number of signals carrying the same information with different time delays. This situation can be viewed as a unicast transmission with multipath fading. To mitigate multipath fading in a unicast transmission, multipath diversity can be employed using direct sequence code division multiple access (DS-CDMA) with Rake receiver [123]. Similarly, in a multihop broadcast scenario, a user equipped with Rake receiver can improve the total SINR by combining signals from different transmitters.

According to our channel model, the received signal at user j can simply be written as:

$$v_j(t) = \sum_{i \in T} \left[\sqrt{P d_{ij}^{-\alpha}} u(t - \tau_i) + \eta_i(t) \right] \tag{5.5}$$

where ${d_{ij}}^2$ is the distance between node i and j, α is the path-loss exponent, τ_i denote the time delay of path i, and $\eta_i(t)$ is the associated AWGN with power spectral density of $N_o/2$ at path i. As mentioned in [119], using MRC, a Rake receiver with R fingers ($R \leq n+1$) can achieve an SINR value, which is the sum of the SINRs on each finger. At each finger, unresolvable paths contribute to both signal and interference power based on the delay difference while remaining resolvable paths contribute to the interference power. Then, the total SINR achieved at a receiver i, Ψ_i, can be calculated as follows:

$$\Psi_i = \sum_{k=1}^{R} \frac{P d_{t_k i}^{-\alpha} + \sum\limits_{w \in H_k} P d_{wi}^{-\alpha} \left(1 - \frac{|\tau_{t_k} - \tau_w|}{T_c}\right)}{N_0 B + \sum\limits_{w \in H_k} P d_{wi}^{-\alpha} \frac{|\tau_{t_k} - \tau_w|}{T_c} + \sum\limits_{q \in I_k} P d_{qi}^{-\alpha}} \tag{5.6}$$

where k is the finger index, t_k is the index of the transmitter that is locked by kth finger, T_c is the chip time ($T_c \approx 1/B$ as in [38]), and H_k and I_k denote the set of unresolvable and

[2]Throughout the chapter, d_{ij}, $d_{i,j}$, $d(i,j)$ are used to define the Euclidean distance between the points i and j, interchangeably.

resolvable paths with the kth finger, respectively. Alternatively, we can omit sets in eqn. (5.6) by using the unit step function, $u(x)$, as follows:

$$\Psi_i = \sum_{k=1}^{R} \frac{Pd_{t_k i}^{-\alpha} + \sum\limits_{\substack{w=1 \\ w \neq t_k}}^{n} Pd_{wi}^{-\alpha} \left(1 - \frac{|\tau_{t_k} - \tau_w|}{T_c}\right) \widetilde{u}(\tau_w)}{N_0 B + \sum\limits_{\substack{w=1 \\ w \neq t_k}}^{n} Pd_{wi}^{-\alpha} \frac{|\tau_{t_k} - \tau_w|}{T_c} \widetilde{u}(\tau_w)} \tag{5.7}$$

$$\widetilde{u}(\tau_w) = u\left(\frac{|\tau_{t_k} - \tau_w|}{T_c}\right) - u\left(\frac{|\tau_{t_k} - \tau_w|}{T_c} - 1\right) \tag{5.8}$$

5.3.5 Problem Statement & Our Contributions

The problem at hand considers a source node that broadcasts data to all n relay nodes and m receiver nodes continuously. Meanwhile, n relay nodes are assisting to the broadcast by decoding-and-forwarding the received signal from the source and/or other relays to the other relays and receiver nodes. In order to decode a symbol, all relays process received signals for a symbol duration because relays are assumed to be using symbol-by-symbol decoding as in [128]. Therefore, both propagation time and processing time are considered as done in [129] for system modeling. To achieve diversity from signals transmitted from different nodes, each node is assumed to have a Rake receiver with R fingers. Thus, the receiver can utilize signals separated by more than a chip time by locking each finger to the strongest resolvable signals. At the same time, signals that are received in the same chip duration, which are nonresolvable signals, increase signal power constructively or decrease it destructively based on their phase. In this work, only large-scale path loss characteristics are considered in the channel model, which is commonly used in papers on wireless capacity analysis. Furthermore, generalized physical model is used for broadcast capacity analysis since this model includes interference.

In theoretical analysis, nodes selected as relays are assumed to be equally spaced around a ring. For a given total energy constraint, radius of this ring has to be optimized for maximization of the broadcast capacity. This leads to the optimal broadcast capacity for the given total energy and hop constraint due to the point symmetry of the circular network. However, incorporating a second ring to the capacity analysis increases the complexity of the solution exponentially since at each feasible radius of the inner ring, an optimal outer

ring radius has to be found. Therefore, we stick to the one-ring scenario as it improves the broadcast capacity substantially with a reasonable optimization complexity.

Since there are no papers regarding wireless multihop broadcast capacity analysis with realistic assumptions including interference, CSI availability, receiver type and relay selection, this work tries to close the gap between theoretical and practical aspects in wireless multihop broadcasting. To close the gap, first a continuum model is taken into account in order to observe how broadcast capacity behaves when there are infinitely many nodes in the network. Then, we propose two suboptimal but practical solutions for the problem of optimal relay selection [3]. One of the solutions is distributed and requires less CSI while the other one is centralized and requires more CSI. Furthermore, broadcast capacity performances of theoretical and proposed practical scenarios are discussed along with their information requirements. In the performance analysis, a number of system configurations including bandwidth, transmission power, path-loss exponent and receiver type are considered.

According to the assumptions mentioned above, our analysis focuses on the broadcast capacity of a multihop wireless network under a realistic channel model including interference and under a realistic system model incorporating processing time at the relay nodes, the propagation delay of each transmission and total energy consumption. We try to answer the fundamental questions below:

1. How can we find the maximum achievable broadcast capacity when there are infinitely many users in the network?

2. How does the wireless broadcast capacity scale with the number of relay nodes for a continuum model under different system settings (bandwidth, tx power, path-loss exponent, number of fingers at the receiver and network radius)?

3. How well does a simple randomized and distributed relay selection strategy perform in improving broadcast capacity?

[3]By "*optimal relay selection*", we mean selection of relays that achieve optimal broadcast capacity.

5.3.6 Complexity Issues

Cagalj et al. [12] proved that the minimum energy broadcast problem, where wireless channel is described by a simple path-loss model, in two-dimensional Euclidean metric space is NP-complete. In this problem, interference is not taken into account. The problem is to find a node power assignment vector such that all the nodes are covered and connected while satisfying a total power constraint. The link cost between two nodes depends on the path-loss exponent and the distance between them. However, for a complete realization of a wireless ad hoc network, interference, transmission time of cooperating nodes and the diversity at the receiver side has to be included in the equation of SINR. Therefore, in this paper, we model the wireless channel using the eqn. 5.6.

When wireless channel is modeled by eqn. 5.6, an achievable rate between two nodes can only be computed by knowing the node power assignment vector, $A = [p_1^v p_2^v ... p_{|V|}^v]$. But, the problem is whether there exists a node power assignment vector such that all the nodes in the network are covered and connected while the sum of assigned powers is less than or equal to a total power constraint. Although the question is similar to one in [12], the solution of our problem requires the node power assignment vector in order to check whether the given network is covered and connected and total power constraint is satisfied. In addition, all the receiving nodes locks their fingers on the strongest resolvable signal paths and this situation changes the structure of the problem. Although when there is only one finger at the receiver, i.e. $R = 1$ case, the problem is similar to the problem in [12], our model considers the constructive and destructive effect of nonresolvable signal paths at the receiver. Therefore, finding a cooperating relay set that achieves the optimal broadcast capacity requires the computation of rates between each node in the network for each and every possible node power assignment vector. Although we have assumed full power transmission or no transmission at all, the number of feasible node power assignment vectors increases exponentially with the number of nodes in the network. As a result, for a given network, finding a node power assignment that makes the network connected and satisfies a given data rate and total power constraint can not solved practically.

5.4 Optimal Broadcast Capacity

Since the broadcast capacity is defined to be the minimum channel capacity among each transmitter-receiver pair, we can find the points, which have the minimum capacity, in a network with infinitely many nodes. According to point symmetry, the optimal broadcast capacity of a circular region is equal to the optimal broadcast capacity of π/n-degree pie slice. Therefore, we can simply analyze the broadcast capacity of a pie slice between 0 and π/n-degree, let us call this region as $\mathbf{\Pi}$, to find the optimal broadcast capacity. Although this property simplifies the analysis, we need to consider the chip time and the number of fingers at the receiver.

To start with, we analyzed the case where there is only one finger at the receiver. As MRC is used to combine signals from each finger of Rake receiver, the capacity analysis for more than one-finger case becomes complex. Therefore,we focus on the limiting case $T_c \to 0$ and continue with $T_c > 0$ case for a one-fingered Rake receiver, and then continue with more than one-finger scenario.

5.4.1 One-Finger Scenario

$T_c \to 0$ *Case*

According to our wireless channel model, signal power decays exponentially as the transmitted wave propagates. Under this channel model, Voronoi tessellations can be used to determine the points[4] having minimum broadcast capacity over a region. Let us call the intersection of Voronoi tessellations as Voronoi intersection points. Then, we can write the following theorems for the region $\mathbf{\Pi}$ as $T_c \to 0$.

Theorem 1 *Assume that receiver can not utilize more than one transmission (even they have the same phase), thus all the transmitters except the intended one are interferers. For any interior point, $p \in \mathbf{\Pi}$, $\exists$ a direction d with $|d| = 1$ and a scalar $\epsilon > 0$ such that $\widetilde{p} = p + \epsilon d$ is in $\mathbf{\Pi}$ and $\Psi_p > \Psi_{\widetilde{p}}$.*

Proof 1 *As signal power decays exponentially with the distance, we can take into account the effects of two nearest interferers only and write the SINR function for the interior point*

[4]Let us call these points as "*critical points* ".

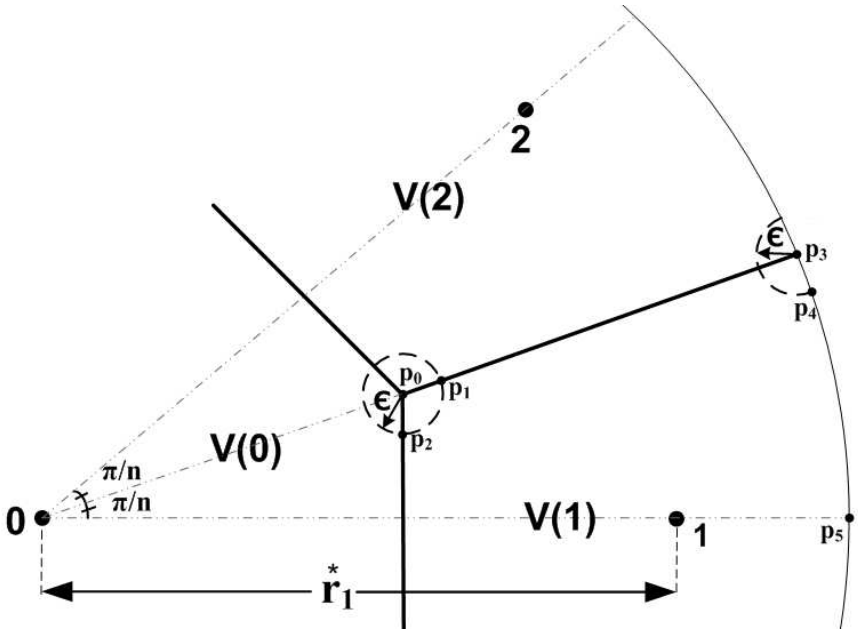


Figure 5.2: Voronoi tessellation of π/n-degree pie slice for theorems 1 and 2.

as $\Psi_p = \frac{Pd_{t,p}^{-\alpha}}{N_0B+Pd_{i_1,p}^{-\alpha}+Pd_{i_2,p}^{-\alpha}}$ *where* t *is the transmitting node,* i_1 *and* i_2 *are the two nearest interferers, w.l.o.g. assume* $d_{i_1,p} < d_{i_2,p}$. *Then we can find a direction* d *and a scalar* ϵ *such that* $\widetilde{p} = p + \epsilon d$ *still resides in* $\mathbf{\Pi}$ *with* $d_{t,p} < d_{t,\widetilde{p}}$ *and* $d_{i_1,p} > d_{i_1,\widetilde{p}}$ *since* p *is an interior point. Further assume that* $d_{i_2,p} < d_{i_2,\widetilde{p}}$. *Then we have the following relations,* $d_{i_2,\widetilde{p}} > d_{i_2,p} > d_{i_1,p} > d_{i_1,\widetilde{p}}$. *As* $\alpha \geq 2$, *using the convexity of the exponential function we can observe that* $d_{i_2,\widetilde{p}}^{-\alpha} + d_{i_1,\widetilde{p}}^{-\alpha} > d_{i_2,p}^{-\alpha} + d_{i_1,p}^{-\alpha}$ *holds. Since* $d_{t,\widetilde{p}}^{-\alpha} < d_{t,p}^{-\alpha}$, *we can conclude that* $\Psi_p > \Psi_{\widetilde{p}}$. ■

Conclusion 1 *If the receiver can not utilize more than one transmission, then the critical point lies on the boundary of the region according to theorem 1. Then, the broadcast capacity is determined by* p_0, p_3 *or relay* 1.

Theorem 2 [5]*For* $|N| \geq 6$, *assume* $p_0 \in \mathbf{\Pi}$ *is a Voronoi intersection point and* $p_3 \in \mathbf{\Pi}$ *is a point at the intersection of Voronoi tessellation and the boundary of* $\mathbf{\Gamma}$. $\exists\ \epsilon$ *such that* $\epsilon > 0$ *with* $\|p_0\|_2 \gg \epsilon$. *Further assume,* $p_1 \in \mathbf{\Pi}$ *and* $p_2 \in \mathbf{\Pi}$ *are in* ϵ*-neighborhood of* p_0, *and* p_1 *is on Voronoi tessellation of* 1 *and* 2 *while* $p_2 \neq p_1$. *On the other hand,* $p_4 \in \mathbf{\Pi}$ *is in*

[5]Throughout the paper, $|N| \geq 6$ is assumed, i.e., there are at least six relays, in the continuum model.

ϵ-neighborhood of p_3 and p_4 is on the boundary of $\mathbf{\Gamma}$, while $p_5 = [r, 0]$. Then, the following inequalities hold:

$$\Psi_{p_1} > \Psi_{p_0} > \Psi_{p_2} > \Psi_1 \tag{5.9}$$

$$\Psi_{p_3} > \Psi_{p_4} \tag{5.10}$$

$$\Psi_{p_5} > \Psi_{p_4} \tag{5.11}$$

Proof 2 *The situation is depicted in Fig. 5.2 above, where only p_2 is placed arbitrarily. Following the result of theorem 1, the candidates for critical points are p_0 and p_3, however the receiver can utilize more than one transmissions with the same phase in this scenario. Since the received power decays exponentially with the distance, all transmitters except 0, 1, 2 and n (which is horizontally symmetric with 2) can be ignored for simplicity. Also, the transmission time of the transmitters can be assumed as $\tau_0 = 0$, and $\tau_1 = \tau_2 = \tau_n = \tau$, as $T_c \to 0$ all the signal components that are in delay with respect to the intended transmission would be treated as interference. At p_1 and p_0, the receiver locks on 1 & 2's transmission since signals are synchronized and adds up constructively, thus they create a stronger signal component. $\Psi_{p_0} = \frac{2Pd_{p_0}^{-\alpha}}{N_0B+Pd_{0,p_0}^{-\alpha}+Pd_{n,p_0}^{-\alpha}}$, $\Psi_{p_1} = \frac{2Pd_{p_1}^{-\alpha}}{N_0B+Pd_{0,p_1}^{-\alpha}+Pd_{n,p_1}^{-\alpha}}$ where $d_{1,p_0} = d_{2,p_0} = d_{p_0}$ and $d_{1,p_1} = d_{2,p_1} = d_{p_1}$. From p_0 to p_1, signal power is increased since $d_{p_1} < d_{p_0}$ while interference power is decreased as $d_{0,p_1} > d_{0,p_0} \implies \Psi_{p_1} > \Psi_{p_0}$. For p_2 shown in Fig. 5.2 SINR function can be written as, $\Psi_{p_2} = \frac{Pd_{0,p_2}^{-\alpha}}{N_0B+Pd_{1,p_2}^{-\alpha}+Pd_{2,p_2}^{-\alpha}+Pd_{n,p_2}^{-\alpha}}$ where $d_{i,p_2} \approx d_{p_0}$ for $i = 0, 1, 2 \implies \Psi_{p_0} > \Psi_{p_2}$. It is important to mention that this function is not dependent on the arbitrary position of p_2. When $|N| = 6$, $\Psi_{p_2} = \frac{Pd_{0,p_2}^{-\alpha}}{N_0B+Pd_{1,p_2}^{-\alpha}+Pd_{2,p_2}^{-\alpha}+Pd_{n,p_2}^{-\alpha}}$ where $d_{i,p_2} \approx \frac{\overset{*}{r}_1}{\sqrt{3}}$ for $i = 0, 1, 2$ and $\Psi_1 = \frac{Pd_1^{-\alpha}}{N_0B+2Pd_1^{-\alpha}}$ where $d_{j,1} = \overset{*}{r}_1$ for $j = 0, 1, 2, n$. Since $N_0B > 0 \implies \Psi_{p_2} > \Psi_1$. As 1 locks on 0 in order to provide full connectivity of the network, Ψ_1 decreases due to the increase in interference from neighboring relays for $|N| > 6$. So, $\Psi_{p_2} > \Psi_1$ holds for $|N| \geq 6$. Thus, $\Psi_{p_1} > \Psi_{p_0} > \Psi_{p_2} > \Psi_1$ and this proves eqn.(5.9). For eqn. (5.10), we can follow the same analysis used for eqn. (5.9) but since node 0 is far away, we can neglect all transmitters except 1, 2 and n. At p_3, the receiver locks on 1 & 2's transmission because they create a stronger component by adding up constructively. $\Psi_{p_3} = \frac{2Pd_{p_3}^{-\alpha}}{N_0B+Pd_{n,p_3}^{-\alpha}}$ where $d_{1,p_3} = d_{2,p_3} = d_{p_3}$. At p_4, the receiver locks on 1, $\Psi_{p_4} = \frac{Pd_{1,p_4}^{-\alpha}}{N_0B+Pd_{0,p_4}^{-\alpha}+Pd_{2,p_4}^{-\alpha}}$. Since $d_{p_3} - d_{1,p_4} < \epsilon$ the signal power is decreased while the interference power is increased, $\Psi_{p_3} > \Psi_{p_4}$. According to the result of WIS (see Appendix),*

$\Psi_{p_5} > \Psi_{p_4}$, *which proves eqn. (5.11). This completes the proof.* ■

Theorem 3 *According to theorems 1 and 2, the maximum broadcast capacity of a circular region,* $\mathbf{\Gamma}$*, is equal to the maximum broadcast capacity of* $\mathbf{\Pi}$ *, which is equal to:*

$$C_B(\mathbf{\Gamma}) = C_B(\mathbf{\Pi}) = \min\left[C_B(S,1), C_B(1,p_4)\right] \tag{5.12}$$

Proof 3 *The broadcast capacity in a multihop network is determined by the minimum of source-to-relay and relay-to-destination channels as relays have to decode the data successfully. Therefore, for full connectivity, all relays must lock on the source's transmission. Since relays are placed on a concentric circle with radius* $\overset{\star}{r}$*, the SINR on the channel between* 0 *and* 1 *can be formulated as* $\Psi_1 = \frac{P\overset{\star}{r}^{-\alpha}}{N_0 B + 2P\left[2\overset{\star}{r}\sin(\frac{\pi}{n})\right]^{-\alpha}}$*, excluding the effects of all transmitters except* 0, 1,2 *and n for simplicity. On the other hand, among all the infinitely many nontransmitting nodes in* $\mathbf{\Pi}$*, the capacity is determined by the point having the minimum capacity. We concluded that this capacity is achieved on* p_4 *since all the other points in* $\mathbf{\Pi}$ *have a better SINR as shown in theorem 2. As a result, the maximum broadcast capacity of* $\mathbf{\Pi}$ *is the minimum of* $B\log_2(1+\Psi_1)$ *and* $B\log_2(1+\Psi_{p_4})$*.* ■

Conclusion 2 *The maximum broadcast capacity of a circular region,* $\mathbf{\Gamma}$*, is bottlenecked by either a relay or a point that is in* ϵ*-neighborhood of an intersection of a Voronoi tessellation and the boundary of the region. Thus, we can easily compute the achievable broadcast capacity of a given network with 1-fingered Rake receiver for asymptotic case* $T_c \to 0$*.*

$T_c > 0$ Case

Following the analysis on the limiting case, $T_c \to 0$, we can examine the realistic scenario, $T_c > 0$. With the introduction of the chip time, the Rake receiver can utilize signals delayed by less than T_c other than the intended transmission according to eqn. (5.6). The region of the utilization, $\mathbf{\Lambda}_\Pi$, is a set of points that satisfy the following inequality:

$$\mathbf{\Lambda}_\Pi = \{p \in \mathbf{\Pi} \mid d_{p_2 p} - d_{p_1 p} < c.T_c\} \tag{5.13}$$

where c is the speed of light. A simple demonstration for this utilization region, depicted by dashed lines, is shown in Fig. 5.3 below.

Theorem 4 *Assume the utilization region intersects the boundary of* $\mathbf{\Gamma}$ *at* $\widetilde{p_4}$*. Then,* $\exists\ \overset{*}{p}_4 \in arc(p_4\widetilde{p_4})$*, where* $arc(p_4\widetilde{p_4})$ *denotes the arc joining points* p_4 *and* $\widetilde{p_4}$*, such that the maximum broadcast capacity of a circular region,* $\mathbf{\Gamma}$*, is equal to:*

$$C_B(\mathbf{\Gamma}) = \min\left[C_B(S,1), C_B(1, arc(p_4\widetilde{p_4}))\right] \tag{5.14}$$

Proof 4 *Following the analysis in theorem 1, we claim that for any interior point,* $p \in \mathbf{\Pi}$ *in the* ϵ*-neighborhood of the network boundary,* $\exists$ *a direction* d *with* $|d| = 1$ *and a scalar* $\epsilon > 0$ *such that* $\widetilde{p} = p + \epsilon d$ *is in* $\mathbf{\Pi}$ *and* $\Psi_p > \Psi_{\widetilde{p}}$*. As signal power decays fast with the distance, if the interior point moves in a direction that will make it far from the transmitter, the change in the signal power causes a decrease in SINR regardless of the change in the interference power since the signal power dominates it. This situation can be observed over the x-axis, from* $(\overset{*}{r}, 0)$ *to* $(r, 0)$*, which is* p_5 *in Fig. 5.2. For* $|N| \geq 6$ *and* $r > \overset{*}{r}_1 > r_l$ *where* r_l *is a lower bound for wireless channel model to hold,* Ψ_p *always decreases as* p *moves from* $(\overset{*}{r}_1, 0)$ *to* p_5 *regardless of other system parameters. Therefore, the utilization of different signal components by a Rake receiver with one-finger does not change the solution space for the optimization problem. Thus, the critical point of the region* Π *still resides on the arc joining* p_4 *and* p_5*, i.e. on the network boundary. As a result, the maximum broadcast capacity is the minimum of the capacities of a relay and a set of points on the arc joining points* p_4 *and* $\widetilde{p_4}$*.* ■

Theorem 5 *For* $T_c > 0$ *case, similar to the theorem 3 in* $T_c \to 0$ *case, the maximum broadcast capacity of a region is equal to:*

$$C_B(\mathbf{\Gamma}) = \min\left[C_B(S,1), C_B(1, \overset{*}{p}_4)\right] \tag{5.15}$$

Proof 5 *Since the maximum broadcast capacity is bottlenecked by a relay and an arc, this condition can be simplified to a case where the capacity is minimum over relay* 1 *and a point in the arc. The point* $\overset{*}{p}_4$ *is the solution to the equation;* $\min_{p \in \mathbf{\Upsilon}} (\Psi_p)$ *where* $\mathbf{\Upsilon} = arc(p_4\widetilde{p_4})$ *and* Ψ_p *is defined by eqn. (5.6).* ■

Conclusion 3 *The maximum broadcast capacity of a circular region,* $\mathbf{\Gamma}$*, is bottlenecked by either a relay or a point on the boundary of the region. Thus, we can compute the achievable broadcast capacity of a given network with one-fingered Rake receiver for* $T_c > 0$ *case.*

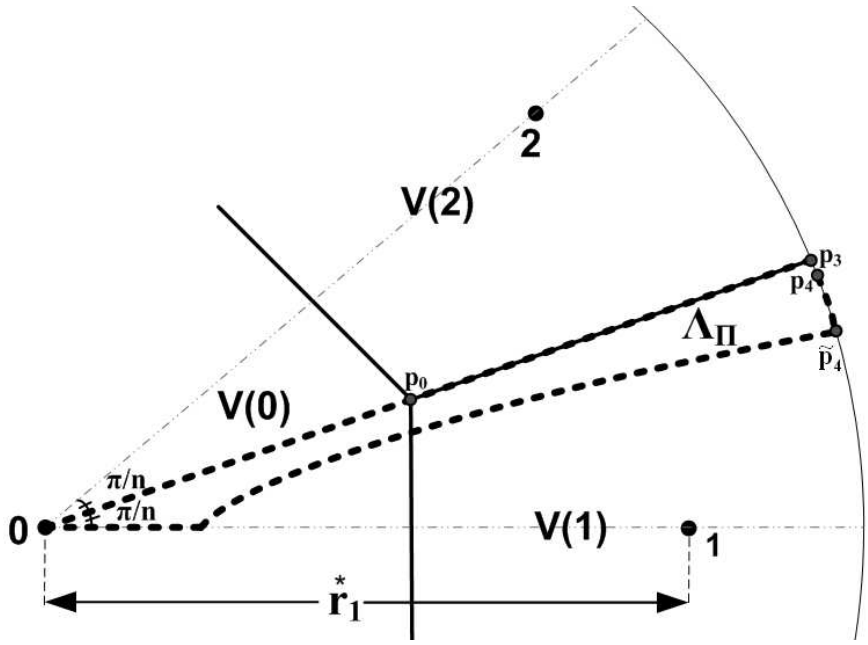


Figure 5.3: Voronoi tessellation of π/n-degree pie slice for theorem 4 and 5.

5.4.2 More than One-Fingers Scenario

When there are more than one-fingers at the receiver, for any interior point, $p \in \mathbf{\Pi}$, a direction d with $|d| = 1$ and a scalar $\epsilon > 0$ such that $\widetilde{p} = p + \epsilon d$ is in $\mathbf{\Pi}$ and $\Psi_p > \Psi_{\widetilde{p}}$ may not exist. This is because the SINR function of a receiver is a sum of SINR functions, such as for two fingers (ignoring far transmitters) $\Psi_p = \frac{Pd_{t,p}^{-\alpha}}{N_0 B + Pd_{i1,p}^{-\alpha} + Pd_{i2,p}^{-\alpha}} + \frac{Pd_{i1,p}^{-\alpha}}{N_0 B + Pd_{t,p}^{-\alpha} + Pd_{i2,p}^{-\alpha}}$, for $d_{t,p} < d_{i_1,p} < d_{i_2,p}$, which may not always decrease as the point moves away from the its strongest transmitter to its strongest interferer. Therefore, when there are more than one finger at the receiver, the critical points for the broadcast capacity may not be on the network boundary. However, SINR is a smooth, continuous, and well-behaving, i.e. no rapid fluctuations and abrupt changes, function except at points that are very close to any transmitter. So, the candidates for the critical point can be found using a method, which we call as feasible direction method (FDM), that converges to a local critical point (nearest critical point to the initial guess) of the region.

Feasible Direction Method

Similar to the theorem 1, we can compute SINR over a ball, defined by ϵ and a set of direction vectors d_i with $\|d_i\| = 1$ for $i = 1, ..., k$, where k denotes the number of intervals on the ball, around an interior point. The selected interior point is only an initial guess, and it is updated if there is a point on the ball with a worse SINR. If there is none, ϵ can be decreased (e.g. halved) for a finer-grained search. Ultimately, a critical point is found when $\epsilon < \epsilon_l$, where ϵ_l is a lower bound on ϵ.

The feasible direction method finds the local critical point for a given initial guess. In order to select an initial guess, the region at hand has to be divided into subregions. We used both Voronoi tessellations and Delaunay triangulation, which corresponds to the dual graph for the Voronoi diagram, for determining these subregions. Then, we can find the local critical point of a subregion for a given system setting as shown in Fig. 5.4. The convergence of FDM from an initial guess (IG) to a local critical point (LCP) for different guesses is shown in the figure. Yellow lines denote the Voronoi tessellations, red lines denote the Delaunay triangulation while grey area is the utilization region. As we can see, different initial guesses in the same subregion converges to the same local critical point.

As a result, we can use FDM to locate the local critical points of a given region. Then, the maximum achievable broadcast capacity is the minimum of those local critical points and the capacity between source and one of the relays.

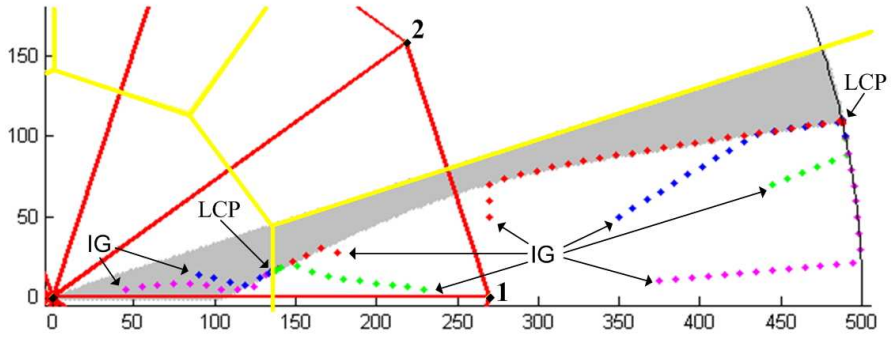


Figure 5.4: Convergence of FDM for different initial guesses.

5.5 Proposed Heuristics

As mentioned previously, for a given total energy constraint, selecting optimal broadcast capacity achieving subset of nodes as relays is a complicated problem in a practical broadcasting scenario, where there are finitely many users in the network. To find a good suboptimal solution to this problem, we performed an analysis in the previous section, considering the case where there are infinitely many nodes in the network. This analysis constitutes a basis for our proposed methods that can be used in practical scenarios, because the optimized parameters of this analysis are used in our methods heuristically. The proposed relay selection procedures can be applied to the existing ad hoc networks in order to increase the broadcast capacity since our proposition does not require any change in the hardware. Here, we present two simple algorithms, based on the parameters optimized under the continuum model, as suboptimal but practical solutions to the relay selection problem. First one, random selection (RS), is a distributed algorithm based on random selection and requires node-to-source CSI only in the selection phase, while second one, selection by triangularization (ST), is a centralized algorithm based on not only node-to-source CSI but also node-to-node CSI.

According to RS, the source node defines two threshold SINR values per rings for a given total energy constraint using the optimal radii found under the continuum model as heuristics. These threshold SINR values determine the number of relay candidates as users having an SINR in between these thresholds are candidates for relay selection. As source node knows each node-to-source CSI, it computes a number in between 0 and 1 for selection and sends this number to candidate nodes. Then, each node generates a random number between 0 and 1 independent of each other. If the number generated at a node is less than the number sent by the node, then the node is selected as a relay node. As seen, RS is a randomized and distributed relay selection method and it requires only the source-to-node CSI. Therefore, in this method, $(m+n)$, which is $O(m+n)$, channel information is necessary to initiate the broadcast session.

In ST, the goal is to imitate the ring structure created under the continuum model. First, source node selects a node, which has an SINR value closest to the SINR value of the optimal ring found in the continuum model, as a relay. Then a second node is searched using triangularization of the source node and the first selected node. The second selected node is approximately $2\pi/n$ degree apart from the first selected node while it has the closest

SINR to the SINR of the first selected node. Following relays are selected by the source node and two neighboring relay nodes using the same strategy. As source node plays an active role in deciding each relay node, this method is centralized. On the other hand, as triangularization is used, not only node-to-source CSI but also node-to-node CSI is required for this algorithm. As a result, in ST, $\left[\frac{3}{2}(n^2-n)+3m(n-1)+1\right]$, which is $O\left(n\left(m+n\right)\right)$, number of channel information is required prior to broadcasting.

5.6 Simulations

In this section, we conducted comprehensive simulations to observe the behavior of the broadcast capacity under different system configurations. First, we computed the broadcast capacity under the continuum model and then performed Monte-Carlo simulations for the proposed algorithms in a finitely many user network. Bandwidth, power, path-loss exponent and radius of the network are swept in order to understand how broadcast capacity changes according to the system parameters.

First of all, in order to observe the gains of the multihop strategy, we consider capacity ratio between multihop and direct transmission as a measure for the simulations. In a direct transmission, i.e., no-hop, there are no relays assisting to the source, therefore source is the only transmitter. The broadcast capacity for direct transmission can be computed with the following formula:

$$C_B(\mathbf{\Gamma}) = B\log_2\left(1+\frac{Pr_{\max}^{-\alpha}}{N_0B}\right) \tag{5.16}$$

where $r_{\max} = \max_i(r)$, $i = 1,...,n+m$, i.e., $r_{\max}$ is the maximum distance between the source and all the nodes in the network.

5.6.1 Simulation Parameters

As mentioned previously, all relay nodes perform DF and symbol-by-symbol decoding, thus each relay waits for a symbol duration, T_s, before transmission. In simulations, the ratio of symbol duration to chip duration, i.e., processing gain $N = T_s/T_c$, is taken to be 100, where $N >> 1$ is common. When computing the optimal ring radius, a lower bound, $r_{1_{\min}} = 50$ m., is set in order to avoid small values of $\overset{\star}{r}_1$ as the channel model is not accurate for small distances. Threshold SINRs for RS are selected such that the number of candidate nodes

is two times the relay nodes, where this value is set heuristically. The noise power spectral density is taken to be 2.10^{-9} W/Hz.

Table 5.1: Simulation Parameters

Simulation Parameter	Set of values
B	{1,10,50,100} (MHz)
P	{125,250,500} (mW)
α	{2,2.8,3.6}
r	{500,1000,2000} (m)
R	{1,2,3}

As this work aims to consider practical aspects of wireless broadcasting, system parameters are selected realistically. All sets of parameter values are listed in Table 5.1. For the ease of understanding, we define a system configuration, which is a set of parameters; bandwidth, power, path-loss exponent, network radius and number of fingers at the receiver. Let Υ be a system configuration, then according to Table 5.1, it can be written as $\Upsilon = \{B_\Upsilon, P_\Upsilon, \alpha_\Upsilon, r_\Upsilon, R_\Upsilon\}$. Then, the optimization problem finds the optimal ring radius that achieves the maximum broadcast capacity for a given number of cooperating peers, $\overset{*}{\Upsilon} = \max_{r_{1_\Upsilon}} \Upsilon$, where $\overset{*}{\Upsilon} = \{B_\Upsilon, P_\Upsilon, \alpha_\Upsilon, r_\Upsilon, R_\Upsilon, \overset{*}{r}_{1_\Upsilon}\}$ and $r_{1_{\min}} \leq r_{1_\Upsilon} \leq r_\Upsilon$. Before computing the optimal ring radius for a given system configuration, utilization region is determined. If utilization region does not cover any boundary of region $\mathbf{\Pi}$ entirely, then there exists points both out of the utilization region and on the boundary. If this is the case, it is proven that there exists a critical point over the boundary and out of the utilization region.

However, if utilization region covers any boundary of region $\mathbf{\Pi}$ entirely, then FDM is used to find the critical points, since the conclusion that the critical point exists on the boundary may not hold. For the optimization of ring radius, we used line search although faster algorithms can be implemented, but this is out of the scope of this chapter. First, for a given system configuration, broadcast capacity is computed from $r_{1_{\min}}$ to r in a discrete manner with a large step size at the first iteration. Since we want to find $\overset{*}{r}_1$ that maximizes broadcast capacity, at each iteration we find two points as new boundaries for the next iteration. These

boundaries are the discrete points where broadcast capacity can be maximum in between. Finding these boundary points is easy since the broadcast capacity function of ring radius is smooth and well-behaving function as mentioned previously. Then, step size is halved at each iteration to increase the accuracy of the optimal ring radius and also the corresponding optimal broadcast capacity. Finally, the computation of optimal radius is terminated when step size is 0.5 m. In the continuum model, when finding critical points, feasible direction method iterations are terminated when step size is 0.1 m.

At last, the optimal ring radius computed for a given system configuration and number of cooperating peers is used for measuring the capacity performances of the proposed heuristics.

5.6.2 Simulation Results

To observe how broadcast capacity behaves, we have to investigate three dimensional plots since there are more than two system parameters affecting this behavior. To start with, we focus on small r with high P in order to see the interference effect when there is only one finger at the receiver. Fig. 5.5 shows that at smaller bandwidth and in a radially small network, increase in transmission power creates an interference, thus decreasing the broadcast capacity as number of relays increases. Therefore, for a given total energy constraint, it may be beneficial not to use all of the energy but use it to some extent where broadcast capacity ratio is maximized. This situation is depicted in the upper left graph of Fig. 5.5.

On the other hand, if the network radius, r, increases, the distance between neighboring relays increases and the effect of interference is less pronounced. As we can see from Fig. 5.6, the interference is no longer decreasing the broadcast capacity. However, as we move to lower bandwidth values or increase the number of relays furthermore (so that the distance between each relay is decreased), same effect can be observed. From these figures, we can also infer that the network radius has negligible effect on the broadcast capacity ratio unless there is interference effect. Above all, in a practical scenario there are more than one finger at the receiver. Then, the interference is mitigated by the diversity achieved from these fingers as shown in Fig. 5.7. However, there is a negligible increase in broadcast capacity ratio as we increase the number of fingers at the receiver from 2 to 3 in lower bandwidths (longer chip durations), thus we do not show the graphs for three-fingered scenario due to space constraints. Another observation is the decreasing benefit of multihop-multifinger

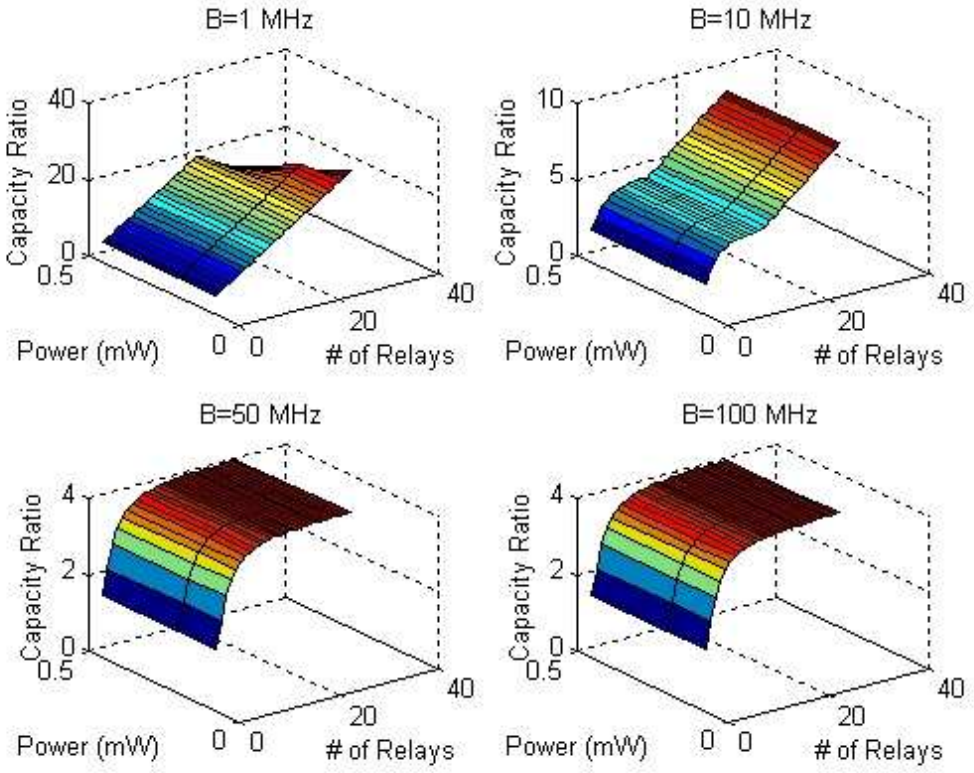


Figure 5.5: Broadcast capacity ratio versus bandwidth and power when $r = 500$, $\alpha = 2$, $R = 1$.

structure as bandwidth increases. This is due to the fact that as bandwidth increases, chip time decreases and the number of nonresolvable signal paths decreases. But even in the worse scenario, there is capacity ratio of 4 as we can see in the lower right graph of Fig 5.5.

The path-loss exponent as another system parameter has an expected effect on broadcast capacity ratio. As channel model is distance based, broadcast capacity diminishes fast with distance at a higher exponent. Therefore, the multihop broadcast capacity increases tremendeously as path-loss exponent increases. This situation can be observed in Fig.s 5.7 and 5.8, where increase in α from 2 to 3.6 brings more than 4-fold increase in each case.

Following the performance of the theoretical case, we performed Monte Carlo simulations in order to measure achievable broadcast capacity values for the proposed methods. As proposed heuristics reflect realistic network scenario where there are finitely many users in

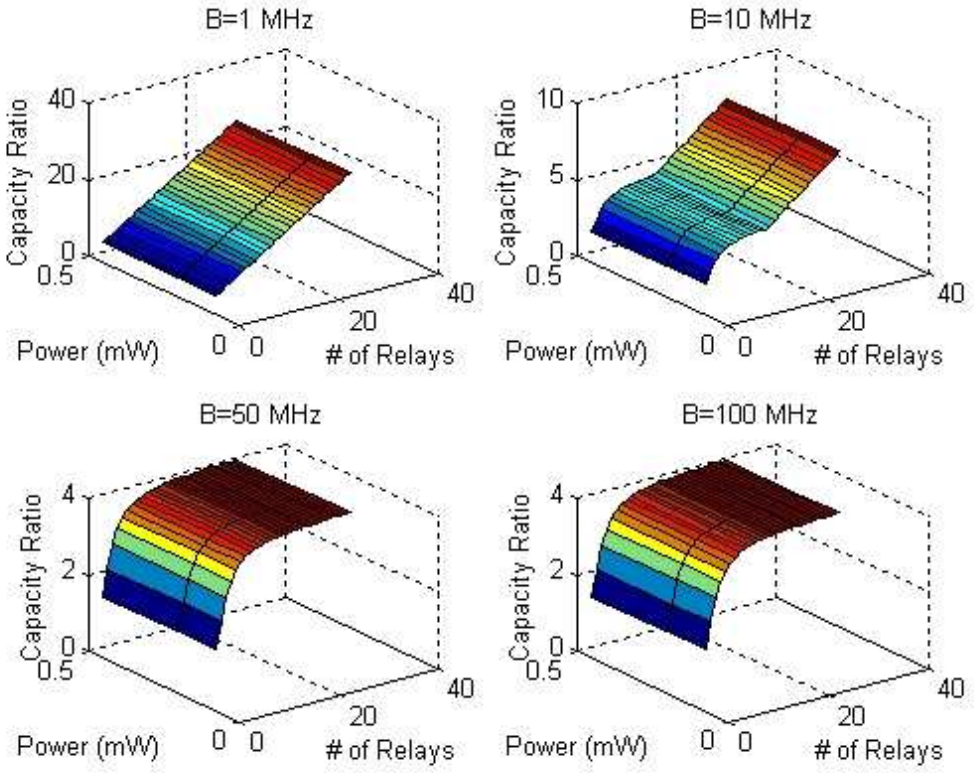


Figure 5.6: Broadcast capacity ratio versus bandwidth and power when $r = 1000$, $\alpha = 2$, $R = 1$.

the network, we measured broadcast capacity under two different network densities, scarcely populated (100 users), and highly crowded (2500 users). Since there are a variety of system parameters at hand, we compared the performance of the proposed methods under four different system configurations. This comparison can be observed for $\overset{*}{\Upsilon}_1 = \{1, 500, 2, 500, 3\}$, $\overset{*}{\Upsilon}_2 = \{1, 125, 2, 2000, 3\}$, $\overset{*}{\Upsilon}_3 = \{100, 500, 2, 500, 3\}$, and for $\overset{*}{\Upsilon}_4 = \{100, 125, 2, 2000, 3\}$ in subplots of Fig.s 5.9-5.16 where $TP = k$ denotes that total power constraint on all the cooperating nodes is k times the source power and each node is transmitting at kP/n. In the first and second configurations, system operates at a narrow bandwidth, while in the third and fourth, it operates at wideband.

When the system is operating at narrowband, Rake receiver is much more beneficial as it can utilize some part of the energy of the nonresolvable signals from other transmitters and

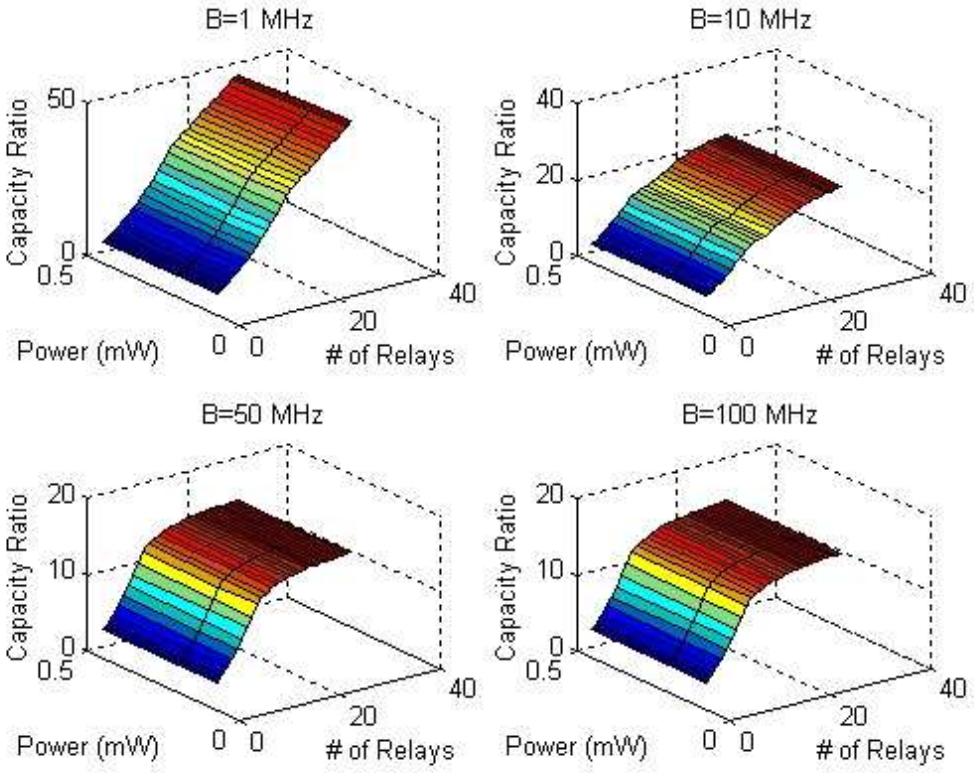


Figure 5.7: Broadcast capacity ratio versus bandwidth and power when $r = 500$, $\alpha = 2$, $R = 2$.

decrease the interference power. Although ST outperforms RS when number of cooperating nodes is less, the opposite happens when number of relays increases as ST forces to select relays that form a ring. The advantage of mimicing the ring structure is observed in Fig. 5.11 while in all the other cases the maximum capacity is achieved when all the users are transmitting. It is important to observe that the maximum broadcast capacity is achieved when all the users are cooperating the broadcast session regardless of the network radius, source transmission power, total power constraint and the density of the network except the scarcely populated network which is using the second configuration and selecting relays via ST. This indicates that when total power is equally divided among all the users, letting each user to act as a relay node achieves the maximum broadcast capacity. As a result, the situation where each node transmits what it decodes provides a better broadcast capacity

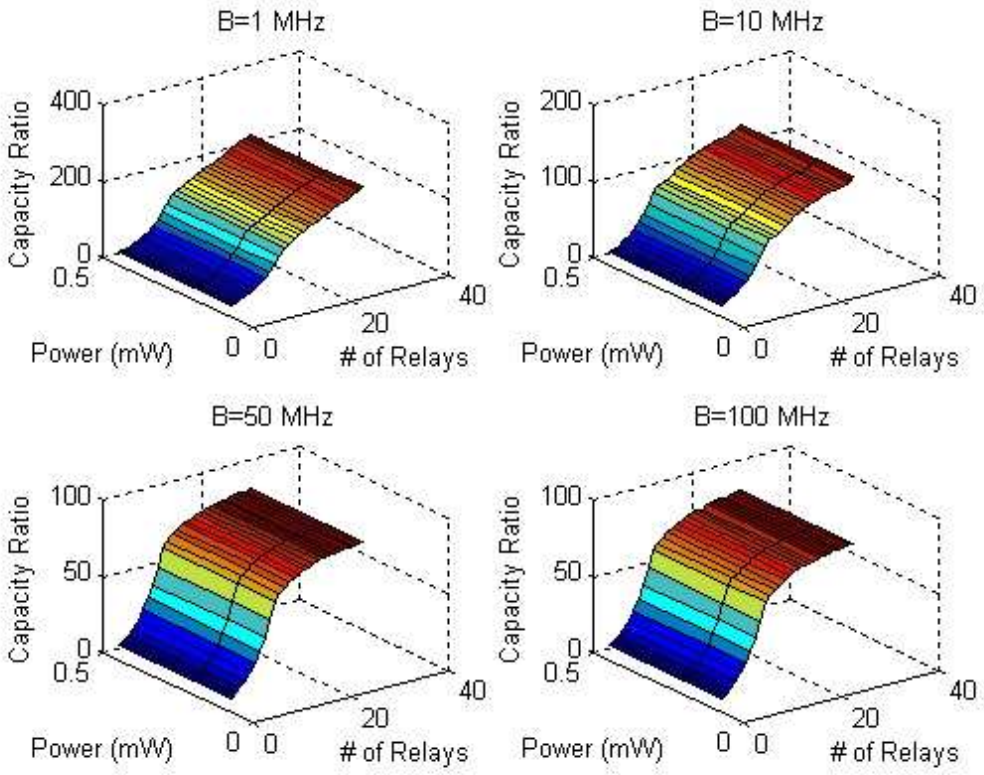


Figure 5.8: Broadcast capacity ratio versus bandwidth and power when $r = 500$, $\alpha = 3.6$, $R = 2$.

instead of increasing the interference power when the system is operating at narrowband regime.

When the system is operating at a wider bandwidth, the capacity gains in both methods decrease due to the property of the Rake receiver. As bandwidth is increased, chip time decreases and transmitters, whose signal arrive in delay more than the chip time, increase the interference power. The optimal number of relays is no more than 10% when the network is highly populated and no more than 30% when the network is scarcely populated. This indicates that interference starts to hurt the broadcast capacity if more number of nodes are selected as relays after these peak points. On the other hand, using a centralized approach rather than a distributed and randomized one brings at most 2-fold increase in broadcasting capacity as shown in Fig.s 5.13-5.16. However, the realization of the centralized

relay selection procedure is unpractical when there 2500 users in the network due to the high number of required CSI. The number of required CSI for proposed heuristics is shown in Fig. 5.17. Also, in networks with high mobility, the relay selection process can be repeated in periodic intervals in order to adapt varying channel conditions. To repeat the relay selection process in a reasonable amount of time, RS can be used instead of ST. Therefore, the complexity of relay selection procedure can be one of the system design objectives that needs to be minimized. In addition, the total power consumed by the whole system may need to be minimized in order to increase battery life for mobile handsets. Therefore, the system may be adjusted to operate at a smaller broadcast capacity while requiring negligible amount of CSI and consuming less power. One can find an optimal compromise operating point between maximizing broadcast capacity, minimizing total energy and minimizing complexity of the relay selection process (number of required CSI) by using multiple-objective optimization, which is out of the scope of this chapter.

5.7 Conclusion

In this work, wireless multihop broadcast capacity is analyzed from a practical perspective. The goal is to close the existing gap between scarce theoretical and practical studies on wireless multihop broadcast capacity in the literature. As the optimal broadcast capacity achieving relay selection problem in a finitely many user network is complicated, we focus on a continuum model, where there are infinitely many nodes in the network. After finding critical points determining the broadcast capacity in the continuum model, we proposed a centralized algorithm and a distributed algorithm for relay selection. Both of these methods use the optimal ring radius, computed by the theoretical model, heuristically.

The performance of the proposed methods are compared in different system configurations. The results show that using a simple, randomized and distributed relay selection method based on the results of a comprehensive theoretical model brings tremendous capacitys gain on the average even in the worse case scenario. As a result, in a broadcasting scenario, signals in the network can be utilized by multiple fingers at the receiver via diversity instead of destructing the intended transmission.

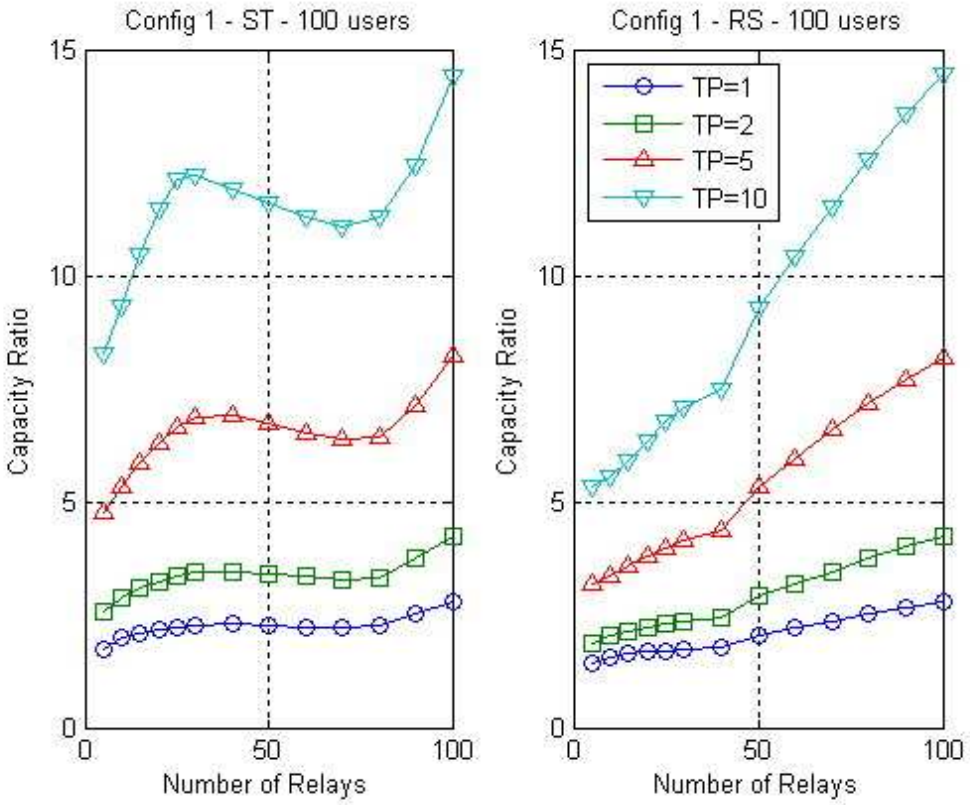


Figure 5.9: Performance comparison of two methods for $\overset{*}{\Upsilon}_1$ when the network is scarcely populated.

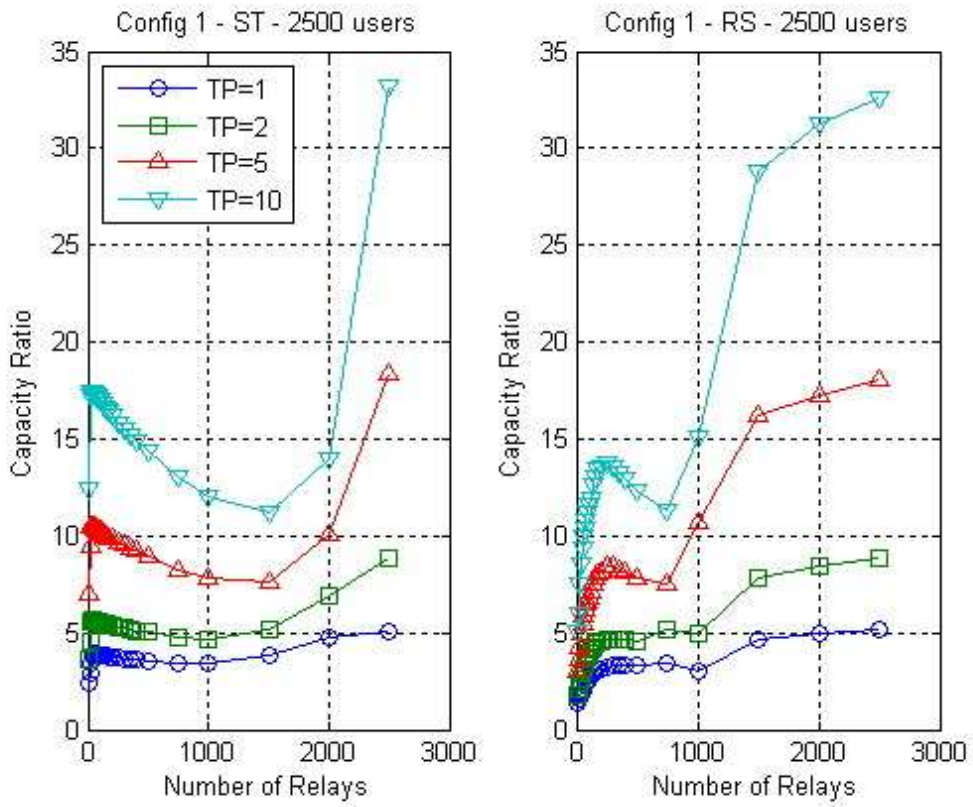
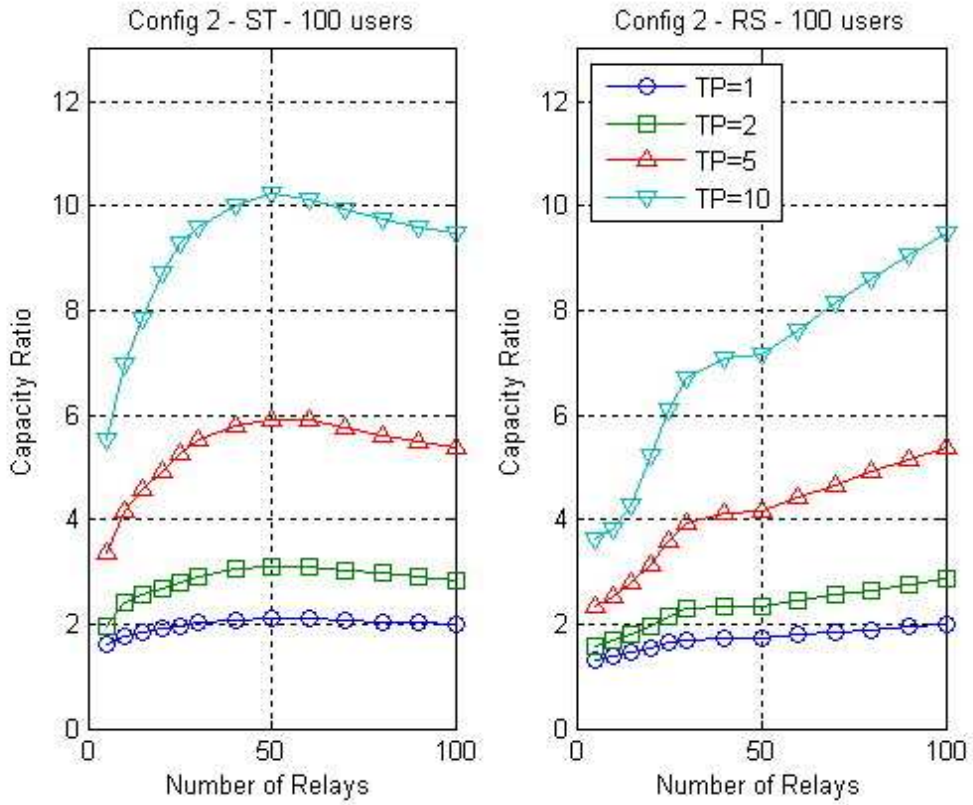


Figure 5.10: Performance comparison of two methods for $\overset{*}{\Upsilon}_1$ when the network is highly populated.



q

Figure 5.11: Performance comparison of two methods for $\overset{*}{\Upsilon}_2$ when the network is scarcely populated.

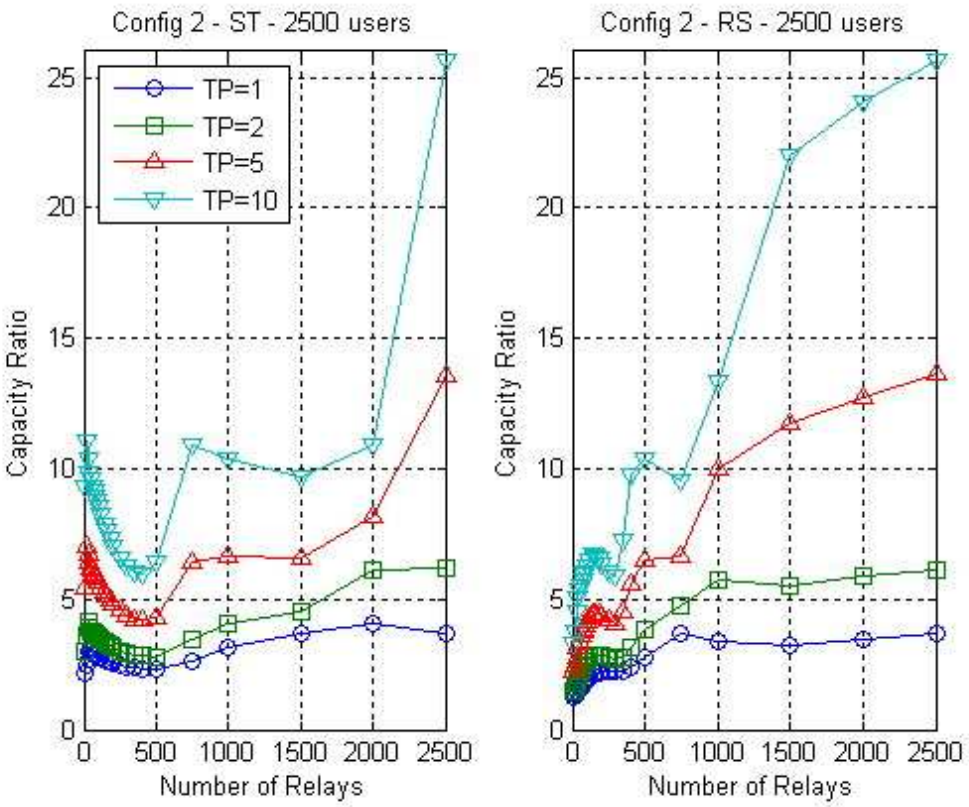
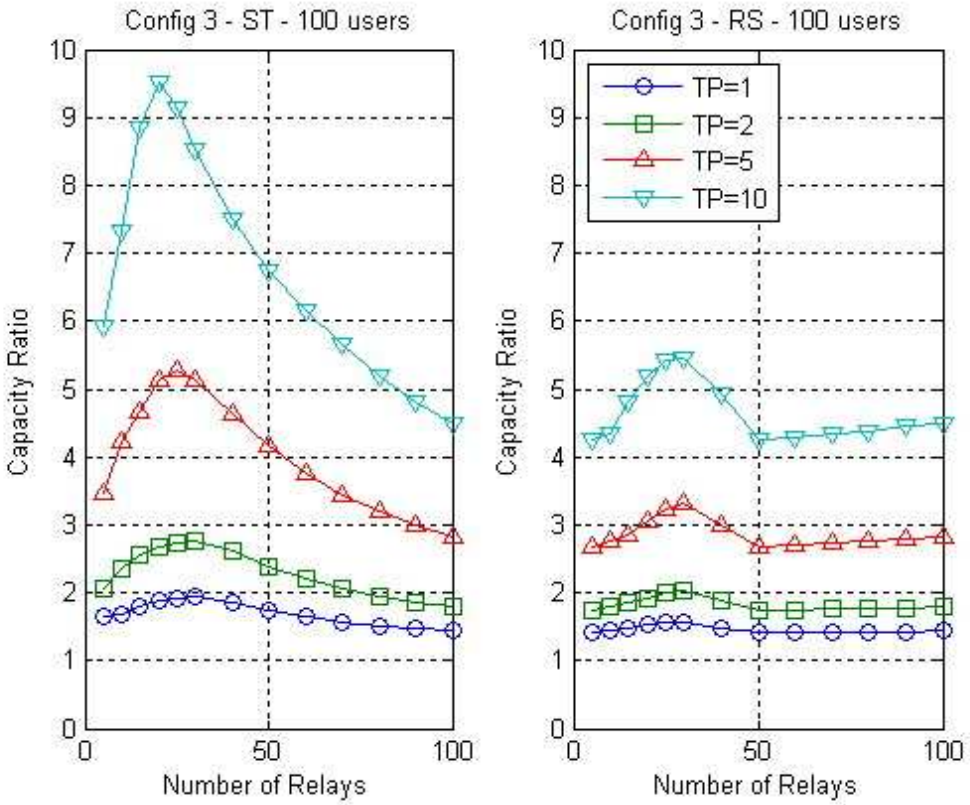


Figure 5.12: Performance comparison of two methods for $\overset{*}{\Upsilon}_2$ when the network is highly populated.



q

Figure 5.13: Performance comparison of two methods for $\overset{*}{\Upsilon}_3$ when the network is scarcely populated.

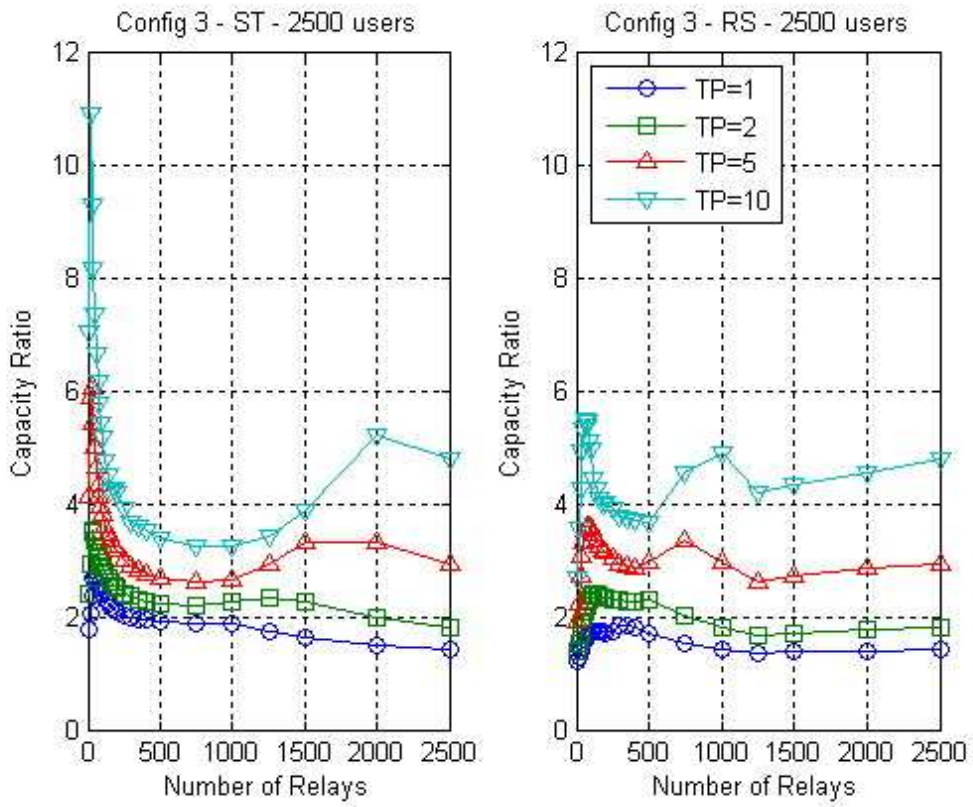
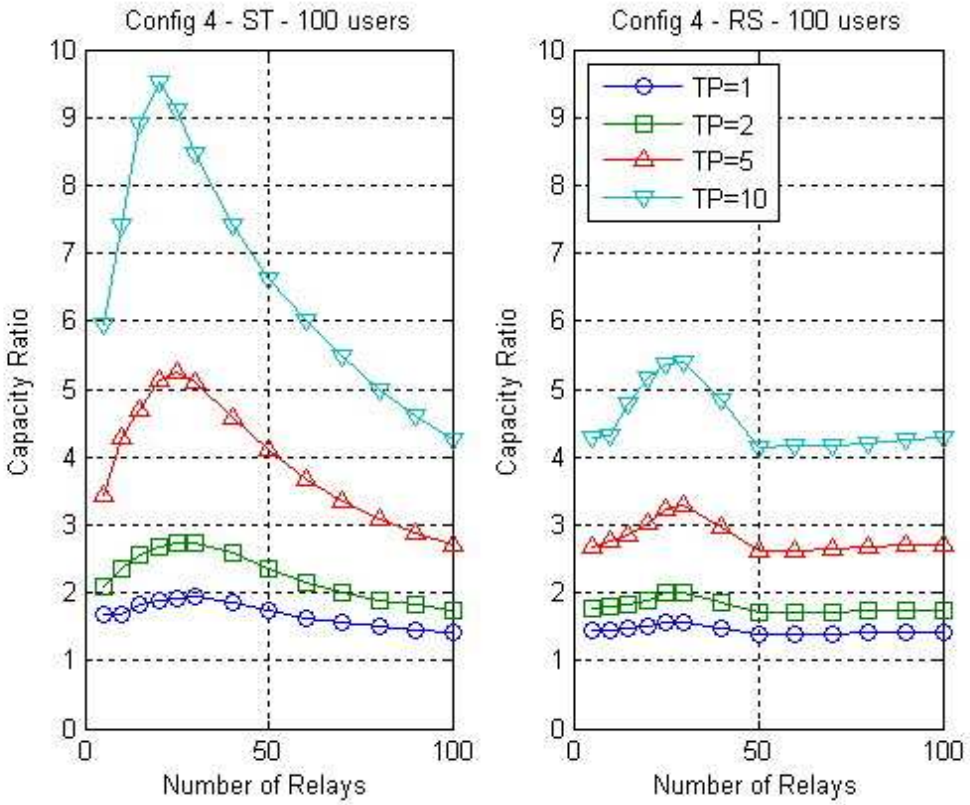


Figure 5.14: Performance comparison of two methods for $\overset{*}{\Upsilon}_3$ when the network is highly populated.



q

Figure 5.15: Performance comparison of two methods for $\overset{*}{\Upsilon}_4$ when the network is scarcely populated.

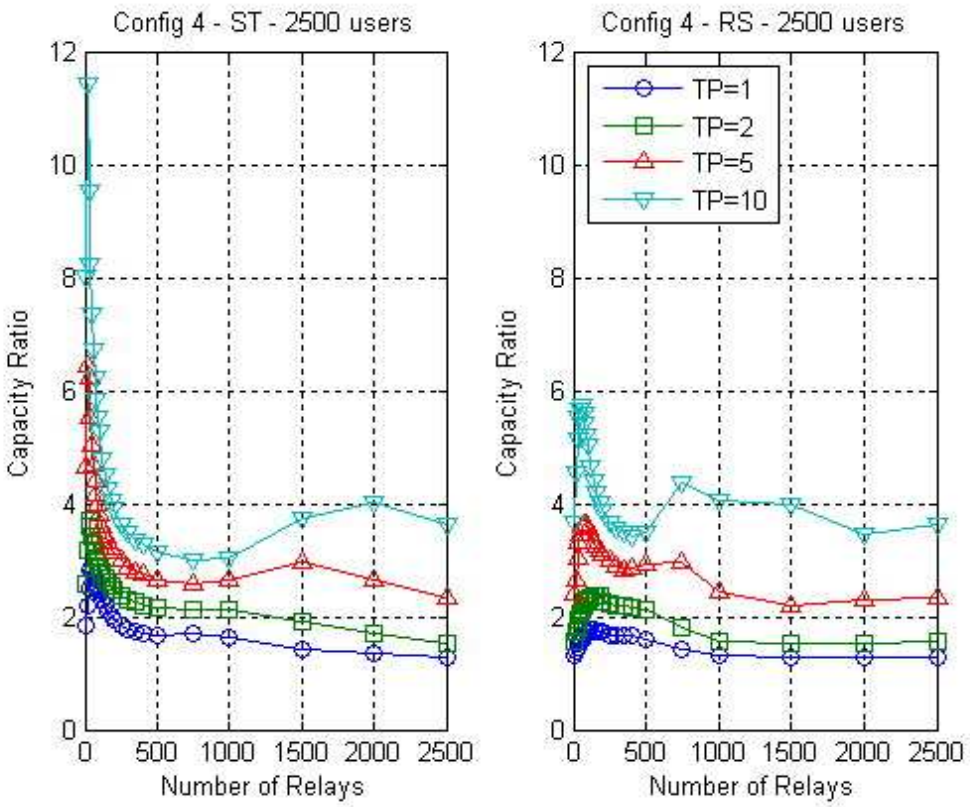


Figure 5.16: Performance comparison of two methods for $\overset{*}{\Upsilon}_4$ when the network is highly populated.

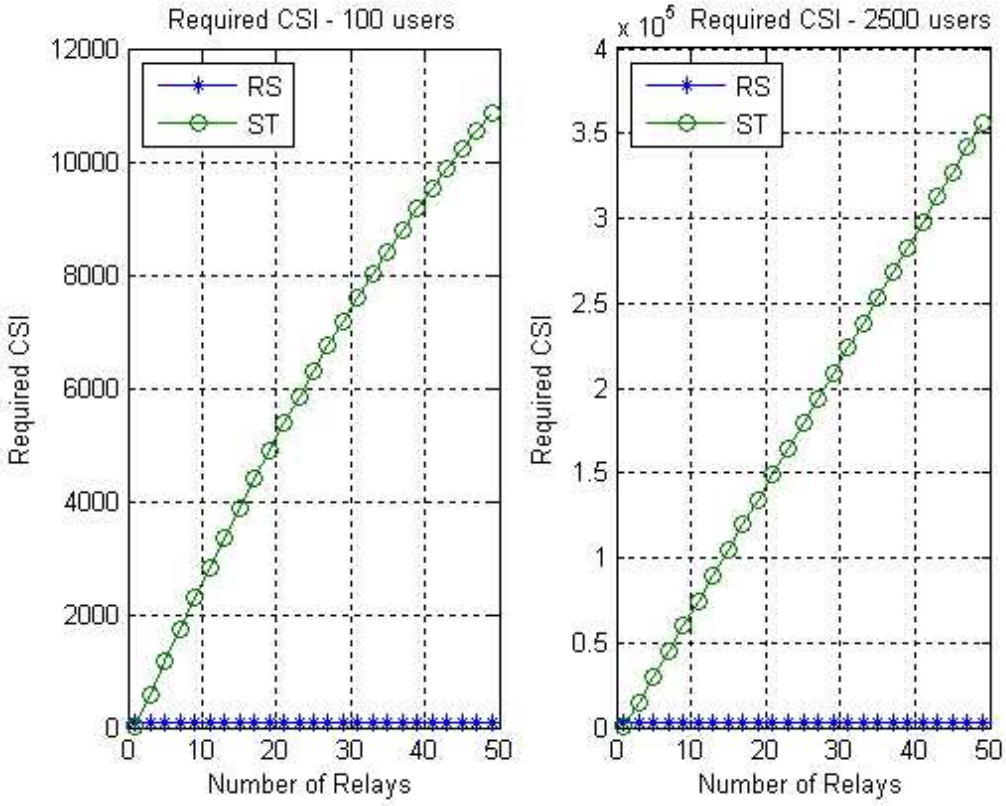


Figure 5.17: Required CSI vs. number of relay nodes for proposed heuristics.

Chapter 6

CONCLUSIONS

In this dissertation, we propose the idea of optimizing the system resources in a cross-layer fashion for wireless multimedia broadcasting and we show the significant gains in the overall system performance achieved by this idea. This thesis presents novel, multi-objective optimized wireless video broadcasting frameworks using a cross-layer approach. All the work done in this thesis considers practical implementations of the proposed system designs and simulations are done using the latest available software and accurate channel models mentioned in the wireless system standards. From this perspective, results found in this thesis may represent a good approximation of the real values that can be achieved by the current wireless systems. In addition, the proposed frameworks can be applied over the existing 3G and the next generation wireless standards. On the whole, the proposed wireless system designs and analysis answer some of the open questions in the literature on wireless broadcasting.

In the first part of the dissertation, we concentrate on the feasibility of using scalable video coding for wireless video broadcasting over the existing 3G wireless standards. Our proposed framework divides forward link code space among multiple layers of the scalable coded video stream optimally in order to maximize the average user throughput and minimize the service outage probability. Our system design includes multi-objective optimization technique for finding a best compromise operating point between two conflicting objectives of the system. The simulations include all the impediments of a wireless channel and channel models are based on a recommendation of ITU for two types of mobility schemes, Pedestrian A and Vehicular B. The result of this study shows that high data rates, such as 614.4 kbps, can be achieved with zero outage for stationary users while high mobility users experience modest outage. However, the use of scalable video coding is not desirable for this scenario since the addition of a second layer degrades the performance for mobile users and brings an inconsiderable improvement for stationary users. This is due to the limited number of

the available physical layer operating modes in 1xEV-DO, which operates at a narrowband of 1.25 MHz. So, if the wireless system operates at wideband regime then the proposed SVC encoded video broadcasting system may increase the system performance.

Following the results of the first part, we focus on a cross-layer optimization for wireless video broadcasting using H.264/AVC. We jointly optimize several system and video coding parameters for the broadcasting service. In addition, this work includes the PER optimization of the wireless system where PER has not been questioned for wireless video broadcasting in the literature. The study differentiates itself from the rest since PER is jointly optimized with FEC instead of fixing PER to 1%,at which all the existing 3G wireless systems operate, and optimizing FEC to this value. A joint source-channel coding approach is taken with MOO, where two conflicting objectives are optimized with respect to the utopia point. Simulation results show that cross-layer optimized wireless video broadcasting outperforms a non-cross-layer design. Also, the video content effects the optimal operating mode of the system. Although cdma2000 1xEV-DO is used in the simulations, the proposed scheme can be applied to any wireless system supporting multiple data rates for transmission. Simulation layout for wireless transmission is similar to the one used in the first study but the video stream is encoded using H.264/AVC instead of SVC.

The proposed SVC encoded and the cross-layer optimized video broadcasting frameworks can be applied to any wireless system supporting multiple data rates for transmission. As cdma2000 1xEV-DO is used in the simulations, the results of our proposed design are dependent to the limitations at the physical layer of 1xEV-DO. Especially, the result on the feasibility of using scalable video coding for wireless video broadcasting is counterintuitive. However, multiple-objective optimization of code space among SVC encoded video layers and cross-layer multiple-objective optimization of physical and application layer parameters can be applied to different wireless system standards.

Lastly, a theoretical analysis of achievable broadcast capacity on wireless multihop networks is performed. Due to the scarce resources on wireless multihop broadcast capacity analysis with practical considerations, this work tries to fill this gap in the literature. The idea is to investigate the behavior of the broadcast capacity with respect to the number of relays in the network. However, the optimal broadcast capacity achieving relay selection problem is complicated. In addition, total energy consumption, interference from multiple

transmitters and utilization of signals by multiple fingers at the receiver via diversity are taken into account in the analysis. Therefore, we compute broadcast capacity limits when there are infinitely many nodes in the network in order to simplify the relay selection problem. Then, we suggest two heuristics, one centralized and one distributed, as suboptimal but practical solutions to this selection problem. In these heuristics, the parameters are selected according to the results of the theoretical model done for infinitely many user case. The performance of the heuristics are compared in different system settings and results are computed with respect to direct transmission (no relays) capacity. The results show that the proposed randomized and distributed relay selection method brings significant capacity gains on the average, even in the worst case scenario.

Appendix A

LEMMAS FOR PROVING ACHIEVABLE MULTIHOP BROADCAST CAPACITY THEOREMS

All of the theorems below are written for one-ring scenario with one-finger at the receiver.

A.1 All-Interfere Scenario (AIS)

Lemma 1 *Assume that receiver can not utilize more than one transmission, thus all the transmitters except the intended one are interferers. Let* $p_0 \in \mathbf{\Pi}$ *is a Voronoi intersection point and* $p_3 \in \mathbf{\Pi}$ *is a point at the intersection of Voronoi tessellation and the boundary of* $\mathbf{\Gamma}$ *as shown in Fig. 5.2. Under this assumption, the SINR function is minimized either at* p_0 *or* p_3*.*

Proof 6 *The proof is intutive. Without any utilization, the SINR function for a point* $p \in \Pi$ *can be written as,* $\Psi_p = \frac{Pd_{0,p}^{-\alpha}}{N_0B + \sum_{m=1}^{n} Pd_{m,p}^{-\alpha}}$*, if* $p \in V(0)$ *and* $\Psi_p = \frac{Pd_{1,p}^{-\alpha}}{N_0B + \sum_{\substack{m=0 \\ m\neq 1}}^{n} Pd_{m,p}^{-\alpha}}$*, if* $p \in V(1)$*. As signal power decays exponentially, signal power decreases when distance increases, which is achieved at either* p_0 *or* p_3*. In addition, interference from closest interfering transmitters increases as point* p *moves towards to the boundary of the region* $\mathbf{\Pi}$*. Therefore, any interior point would have a better SINR than the minimum SINR over the region. Thus, the SINR function is minimized either at* p_0 *or* p_3*.* ■

A.2 Strong Interferer Scenario (SIS)

Lemma 2 *Assume there exists two points,* p *and* t*, on a Voronoi tessellation of nodes* 0 *and* i *with* $\tau_0 \neq \tau_i$ *and out of the utilization region. If* $d_{i,p} > d_{i,t}$ *then* $\Psi_t > \Psi_p$*.*

Proof 7 *For simplicity, we can set* $i = 1$ *due to the circular symmetry and ignore the effect of all transmitters except* 0, 1, 2, *and* n*. Besides, assume* p *and* t *are on* $\overline{p_2p_5}$ *as shown in Fig. 5.2 for visualization. For one-finger case, SINR values can be written as,* $\Psi_p =$

$\frac{Pd_{1,p}^{-\alpha}}{N_0B+Pd_{0,p}^{-\alpha}+Pd_{2,p}^{-\alpha}+Pd_{n,p}^{-\alpha}}$ *and* $\Psi_t = \frac{Pd_{1,t}^{-\alpha}}{N_0B+Pd_{0,t}^{-\alpha}+Pd_{2,t}^{-\alpha}+Pd_{n,t}^{-\alpha}}$ *where* $d_{0,p} = d_{1,p} < \min_{k\in\{2,\ldots,n\}}(d_{k,p})$, $d_{0,t} = d_{1,t} < \min_{k\in\{2,\ldots,n\}}(d_{k,p})$. *Given* $d_{1,p} > d_{1,t}$, p *is closer to* 2 *than* t, *so* $d_{2,p} < d_{2,t}$ *and* $d_{n,p} > d_{n,t}$. *As* 2 *is horizontally symmetrical with* n, $d_{n,p} > d_{n,t} \geq d_{2,t} > d_{2,p}$, *equality is achieved when* t *is on the x-axis in Fig. 5.2. Since signal power decays exponentially with path-loss exponent* $\alpha \geq 2$, $d_{2,p}^{-\alpha} + d_{n,p}^{-\alpha} > d_{2,t}^{-\alpha} + d_{n,t}^{-\alpha}$ *can be written due to the convexity of the exponential function. Let us write SINR values in a compact form,* $\Psi_p = \frac{S_p}{S_p+I_p}$ *and* $\Psi_t = \frac{S_t}{S_t+I_t}$ *where* $S_p = Pd_{1,p}^{-\alpha}$, $I_p = N_0B + Pd_{2,p}^{-\alpha} + Pd_{n,p}^{-\alpha}$, $S_t = Pd_{1,t}^{-\alpha}$, $I_t = N_0B + Pd_{2,t}^{-\alpha} + Pd_{n,t}^{-\alpha}$. *We know that* $S_p < S_t$ *and* $I_p > I_t$. *By contradiction, assume* $\Psi_t \leq \Psi_p$, *then* $\frac{S_t}{S_t+I_t} \leq \frac{S_p}{S_p+I_p}$ $\implies S_tS_p + S_tI_p \leq S_tS_p + S_pI_t \implies S_tI_p \leq S_pI_t \implies \frac{S_t}{S_p} \leq \frac{I_t}{I_p}$, *which can not hold since* $\frac{S_t}{S_p} > 1$ *and* $\frac{I_t}{I_p} < 1$. *Thus,* $\Psi_t > \Psi_p$. ■

A.3 Weak Interferer Scenario (WIS)

Lemma 3 *Assume there exists two points,* $p \in V(i)$ *and* $t \in V(i)$, *on the boundary of the circular region and out of the utilization region. If* $d_{i,p} > d_{i,t}$ *then* $\Psi_t > \Psi_p$.

Proof 8 *For simplicity, we can set* $i = 1$ *due to the circular symmetry and ignore the effect of all transmitters except* 1, 2, *and* n. *Besides, assume* p *and* t *are on* $\overline{p_4p_6}$ *as shown in Fig. 5.2 for visualization. For one-finger case, following the same analysis in SIS, we can write the SINR values in compact form,* $\Psi_p = \frac{S_p}{I_p}$ *and* $\Psi_t = \frac{S_t}{I_t}$ *where* $S_p = Pd_{1,p}^{-\alpha}$, $I_p = N_0B + Pd_{2,p}^{-\alpha} + Pd_{n,p}^{-\alpha}$, $S_t = Pd_{1,t}^{-\alpha}$, $I_t = N_0B + Pd_{2,t}^{-\alpha} + Pd_{n,t}^{-\alpha}$. *We know that* $S_p < S_t$ *and* $I_p > I_t$. *By contradiction, assume* $\Psi_t \leq \Psi_p$, *then* $\frac{S_t}{I_t} \leq \frac{S_p}{I_p} \implies S_tI_p \leq S_pI_t \implies \frac{S_t}{S_p} \leq \frac{I_t}{I_p}$, *which can not hold since* $\frac{S_t}{S_p} > 1$ *and* $\frac{I_t}{I_p} < 1$. *Thus,* $\Psi_t > \Psi_p$. ■

Appendix B

MULTIPLE-OBJECTIVE OPTIMIZATION (MOO)

Multi-objective optimization aims to find the solution of an optimization problem with the set of multiple objectives. A solution is called globally Pareto-optimal if any one of the objectives cannot be improved without degrading the other objectives for this solution [72].

Assume that the optimization problem under investigation consists of distinct, and possibly conflicting objective functions. Without any loss of generality, assume further that the problem in hand requires all of the objective functions to be minimized. Then, a Pareto-optimal solution, s^* exists if there is no other feasible solution, s, that satisfies

$$f_p(s) \leq f_p(s^*),\ \forall p \in \{1, 2, ..., P\} \tag{B.1}$$

with at least one strict inequality. In other words, there is no other feasible solution that is at least as good as this Pareto-optimal solution in all of the objective functions and also is strictly better in one or more objective functions. In our formulation, $P = 2$ and the objective functions are given by (3.5) and (3.8).

For single objective optimization problems, it is possible to have multiple optimal solutions resulting in a unique optimal functional value. It is also possible to have multiple Pareto-optimal solutions in multi-objective optimization problems. However, unlike the single objective problems, the multiple Pareto-optimal solutions do not necessarily result in a unique functional value. In many instances, as different objective functions represent different system aspects on a specific scale, variance, and units of measurement, it is difficult to discriminate between these Pareto-optimal points and determine which one is better than the others. However, if relative importance weights for each of the objective functions is specified, a so-called best compromise solution may be determined.

In order to find the best compromise solution among the objective functions, f_p's, one has to first re-scale their range of values to lie in the intervals $[0, w_p]$, where w_p is the importance weight of the p'th objective function:

$$f_{p,scaled}(m) = w_p \frac{f_p(m) - f_{min}(m)}{f_{max}(m) - f_{min}(m)} \tag{B.2}$$

where $f_{min}(n)$ and $f_{min}(p)$ are the minimum and maximum values of the objective function, $f_p(m)$, respectively. Once scaling is done, all feasible operating points are mapped onto the P-dimensional space where each dimension represents one of the objectives. For $P = 2$ and $w_1 = w_2 = 1$, this is illustrated in Fig. B.1.

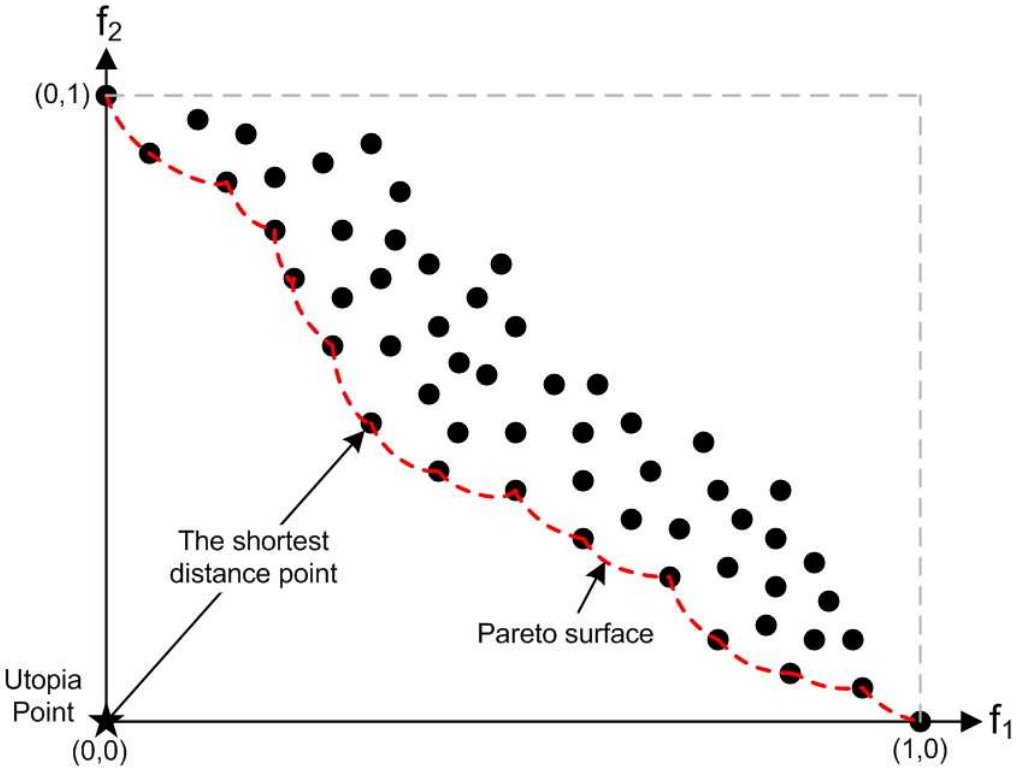


Figure B.1: Pareto surface and the solution with the shortest distance to the utopia point.

In multi-objective optimization, an unfeasible operating point that optimizes all of the objective functions simultaneously is called the utopia point. When $P = 2$, and both of the scaled objectives need to be minimized, this corresponds to the (0,0) point in the two-dimensional objective space as illustrated in Fig. B.1. The best compromise solution is then found as the feasible point that is closest to the utopia point in the Euclidean-distance sense.

It may be necessary to impose constraints to the multi-objective optimization problem. In the proposed framework, for a given number of layers, the data rates for the base and enhancement layers as well as the code space division are determined for the best compromise operating point. The two objectives of maximization of the base layer video coverage and maximization of the average decodable video data rate may be achieved by selecting a very low data rate for the base layer requiring a small portion of the code space, and allocating a remaining resources towards the transmission of high data rate enhancement layers. However, this may result in an unacceptably poor basic video quality for users capable of decoding only the base layer. In practice, one may choose to put a lower bound on the base layer data rate such that an acceptable basic video quality is available for all. When such a constraint is imposed on the multi-objective optimization, the best compromise solution becomes the point with the smallest Euclidean distance from the utopia point among operating points satisfying the constraint.

VITA

CAGDAS ATICI was born in Istanbul, Turkey on October 26, 1982. He received his B.Sc. degree in Electrical and Electronics Engineering with high honors, also from Koç University in 2005 where he had the second highest GPA in his graduating class. He received Dean's Award in 2005 due to his significant contribution to Koç University at large while maintaining a high scholarly achievement. He is a recipient of the Tübitak Ph.D. Fellowship during his Ph.D. studies. From February 2005 to January 2009, he worked as a teaching and research assistant in Koç University, Turkey. He had studied on *CDMA, scalable video coding, cross-layer optimization, multimedia broadcasting.* He has authored articles in these areas at a refereed journal, *IEEE Transactions on Broadcasting,* and international conferences, such as *ICC2006, SIU2007, PIMRC2007, MoViD2008 and BMSB2009.*

PUBLICATIONS

Journals

- C. Atici and M. O. Sunay, "High Data-Rate Video Broadcasting over 3G Wireless Systems," *IEEE Trans. on Broadcasting*, vol. 53, no. 1, pp. 212-223, Mar. 2007.

- C. Atici and M. O. Sunay, "Cross-Layer Design for Wireless Video Broadcasting," submitted.

- C. Atici and M. O. Sunay, "Achievable Broadcast Capacity of Wireless Multihop Interference Networks," submitted.

Conferences

- C. Atici and M. O. Sunay, "Multi-Layered Video Broadcasting over 1xEV-DO Using Multiple Objective Optimization," in *Proc. of the IEEE ICC 2006*, Istanbul, Turkey, June 11-15, 2006.

- C. Atici and M. O. Sunay, "Cross-Layer Design for Wireless H.264/AVC Video Broadcasting System," in *Proc. of the IEEE PIMRC 2007*, Athens, Greece, September 3-7, 2007.

- C. Atici and M. O. Sunay, "Kablosuz Video Yayini icin H.264/AVC Kullanarak Katmanlar Arasi Tasarim," in *Proc. of the IEEE SIU 2007*, Eskisehir, Turkey, June 11-13, 2007.

- V. Dedeoglu, C. Atici, S. Salman and M. O. Sunay, "System Optimization for Peer-to-Peer Multi Hop Video Broadcasting in Wireless Ad Hoc Networks," in *Proc. of the IEEE Workshop on Mobile Video Delivery (MoViD)*, Newport Beach, CA, June 23, 2008.

- C. Atici, M. O. Sunay, "Improving the Performance of Wireless H.264 Video Broadcasting Through a Cross-Layer Design " accepted for publication in *Proc. of the IEEE Broadband Multimedia Systems and Broadcasting (BMSB),* 2009.

BIBLIOGRAPHY

[1] *Advanced Video Coding for Generic Audiovisual Services*, ITU-T, Rec. H.264, Mar. 2005.

[2] A. Agarwal and P. R. Kumar, "Capacity Bounds for Ad hoc and Hybrid Wireless Networks," *ACM SIGCOMM Computer Comm. Review*, vol. 34, no. 3, pp. 71-81, July 2004.

[3] A. Algans, K. I. Pedersen, and P. E. Mogensen, "Experimental Analysis of the Joint Statistical Properties of Azimuth Spread, Delay Spread, and Shadow Fading," *IEEE Journal on Sel. Areas in Comm.*, vol. 20, no. 3, pp. 523-531, Apr. 2002.

[4] H. W. Arnold, D. C. Cox, and R. R. Murray, "Macroscopic Diversity Performance Measured in the 800-MHz Portable Radio Communications Environment," *IEEE Trans. on Anten. and Prop.*, vol. 36, no. 2, pp. 277-281, Feb. 1988.

[5] C. Atici and M. O. Sunay, "High Data-Rate Video Broadcasting over 3G Wireless Systems," *IEEE Trans. on Broadcasting*, vol. 53, no. 1, pp. 212-223, Mar. 2007.

[6] T. Aulin, "A Modified Model for the Fading Signal at a Mobile Radio Channel, " *IEEE Trans. on Veh. Tech.*, vol. 28, no. 3, pp. 182-203, Aug. 1979.

[7] I. V. Bajic, "Efficient Cross-Layer Error Control for Wireless Video Multicast," *IEEE Trans. on Broadcasting*, vol. 53, no. 1, pp. 276-285, Mar. 2007.

[8] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and A. Viterbi, "CDMA/HDR: A Bandwidth-Efficient High-Speed Wireless Data Service for Nomadic Users " *IEEE Comm. Mag.*, vol. 38, no. 7, pp. 70-77, July 2000.

[9] N. Bhushan, C. Lott, P. Black, R. Attar, Y.-C. Jou, M. Fan, D. Ghosh, and J. Au, "CDMA2000 1xEV-DO Revision A: A Physical Layer and MAC Layer Overview, " *IEEE Comm. Mag.*, vol. 44, no. 2, pp. 75-87, Feb. 2006.

[10] Q. Bi, R. R. Brown, D. Cui, A. D. Gandhi, C. Y. Huang, and S. Vitebsky, "Performance of 1xEV-DO Third-Generation Wireless High-Speed Data Systems, " *Bell Labs Technical Journal*, vol. 7, no. 3, pp. 97-107, 2003.

[11] British patent No. 12,039, Date of Application 2 June 1896; Complete Specification Left, 2 March 1897; Accepted, 2 July 1897.

[12] M. Cagalj, J. P. Hubaux, and C. Enz, "Minimum-Energy Broadcast in All-Wireless Networks: NP-Completeness and Distribution Issues," in *Proc. of the MobiCom*, 2002.

[13] Z. Cakareski, N. Ahmed, A. Dhar, and B. Aazhang, "Multilevel Coding of Broadcast Video over Wireless Channels,"in *Proc. of the IEEE ICASSP*, 2002.

[14] CDMA Development Group, *3Q 2008 CDMA Subscribers*, http://www.cdg.org/worldwide/report/083Q_cdma_subscriber_report.pdf, Sep. 2008.

[15] M. R. Chari, F. Ling, A. Mantravadi, R. Krishnamoorthi, R. Vijayan, G. K. Walker, and R. Chandhok, "FLO Physical Layer: An Overview,"*IEEE Trans. on Broadcasting*, vol. 53, no. 1, pp. 145-160, Mar. 2007.

[16] S. Chung and A. Goldsmith, " Degrees of Freedom in Adaptive Modulation: A Unified View, " *IEEE Trans. on Comm.*, vol. 49, no. 9, pp. 1561-1571, Sep. 2001.

[17] R. H. Clarke, " A Statistical Theory of Mobile Radio Reception, " *Bell System Technical Journal*, vol. 47, no. 6, pp. 957-1000, 1968.

[18] *Coding of Audio-Visual Objects - Part 10: Advanced Video Coding*, ISO/IEC, 14496-10, ver. 4, 2008.

[19] *Coding of Audio-Visual Objects - Part 2: Visual*, ISO/IEC, 14496-2, ver. 3, 2004.

[20] D. C. Cox, R. Murray, and A. Norris, "800 MHz attenuation measured in and around suburban houses," *AT & T Bell Labs Technical Journal*, vol. 673, July-Aug. 1984.

[21] P. Dent, G. E. Bottomley, and T. Croft, "Jakes Fading Model Revisited," *Electronics Letters*, vol. 29, no. 13, pp. 1162-1163, June 1993.

[22] Q. Du, V. Faber, and M. Gunzburger, "Centroidal Voronoi Tessellations: Applications and Algorithms," *SIAM Review*, vol. 41, no. 4, pp.637-676, 1999.

[23] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical Solutions for the 3G Long-Term Evolution," *IEEE Comm. Mag.*, vol. 44, no. 3, pp. 38-45, Mar. 2006.

[24] A. Ephremides, "Energy Concerns in Wireless Networks, "*IEEE Wireless Communications,* vol. 9, no. 4, pp. 48-59, Aug. 2002.

[25] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi, "An Empirically Based Path Loss Model for Wireless Channels in Suburban Environments, " *IEEE Journal on Sel. Areas in Comm.*, vol. 17, no. 7, pp. 1205-1211, July 1999.

[26] *Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H)*, ETSI, EN 302 304, v1.1.1, Nov. 2004.

[27] *Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines*, ETSI, TR 102 377, v1.2.1, Nov. 2005.

[28] M. Etoh and T. Yoshimura, "Wireless Video Applications in 3G and Beyond, "*IEEE Wireless Comm.*, vol. 12, no. 4, pp. 66-72, Aug. 2005.

[29] European Cooperative in the Field of Science and Technical Research EURO-COST 231, "Urban transmission models for mobile radio in the 900 and 1800 MHz bands, " rev. 2, The Hague, Sep. 1991.

[30] G. Faria, J. A. Henriksson, E. Stare, and P. Talmola, "DVB-H: Digital Broadcast Services to Handheld Devices,"in *Proc. of the IEEE*, vol. 94, no. 1, pp. 194–209, Jan. 2006.

[31] M. J. Gans, "A Power-Spectral Theory of Propagation in the Mobile-Radio Environment," *IEEE Trans. on Veh. Tech.*, vol. 21, no. 1, pp. 27-38, Feb. 1972.

[32] M. Gastpar and M. Vetterli, "On The Capacity of Wireless Networks: The Relay Case," in *Proc. of the IEEE INFOCOM*, 2002.

[33] *Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video*, ISO/IEC, 13818-2, ver. 2, 2000.

[34] A. Ghosh, D. R. Wolter, J. G. Andrews and R. Chen, "Broadband Wireless Access with WiMax/802.16: Current Performance Benchmarks and Future Potential," *IEEE Comm. Mag.*, vol. 43, no. 2, pp. 129-136, Feb. 2005.

[35] D. Giancristofaro, "Correlation Model for Shadow Fading in Mobile Radio Channels," *Electronics Letters*, vol. 32, no. 11, pp. 958-959, May 1996.

[36] A. J. Goldsmith and P. P. Varaiya, "Capacity of Fading Channels with Channel Side Information "*IEEE Trans. on Info. Theory*, vol. 43, no. 6, pp. 1986-1992, Nov. 1997.

[37] A. J. Goldsmith and S. B. Wicker, "Design Challenges for Energy-Constrained Ad Hoc Wireless Networks,"*IEEE Wireless Comm.*, vol. 9, no. 4, pp. 8-27, Aug. 2002.

[38] A. Goldsmith, *Wireless Communications,* Cambridge University Press, 2005.

[39] F. Graziosi and F. Santucci, "A General Correlation Model for Shadow Fading in Mobile Radio Systems," *IEEE Comm. Letters*, vol. 6, no. 3, pp. 102-104, Mar. 2002.

[40] M. Grossglauser and D. N. C. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks," *IEEE/ACM Trans. on Networking*, vol. 10, no. 4, pp. 477-486, Aug. 2002.

[41] M. Gudmundson, "Correlation Model for Shadow Fading in Mobile Radio Systems, " *Electronics Letters*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[42] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. on Info. Theory*, vol. 46, no. 2, pp. 388-404, Mar. 2000.

[43] J. Hagenauer and T. Stockhammer, "Channel Coding and Transmission Aspects for Wireless Multimedia, " *Proc. of the IEEE*, vol. 87, no. 10, pp. 1764-1777, Oct. 1999.

[44] "Harbour," *Test Sequence*. Available: ftp.tnt.uni-hannover.de/pub/svc/testsequences/

[45] F. Hartung, U. Horn, J. Huschke, M. Kampmann, T. Lohmar, and M. Lundevall, "Delivery of Broadcast Services in 3G Networks, " *IEEE Trans. on Broadcasting,* vol. 53, no. 1, pp. 188-199, Mar. 2007.

[46] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio Services, " *IEEE Trans. on Veh. Tech.,* vol. 29, no. 3, pp. 317-325, Aug. 1980.

[47] Z. He, J. Cai and C. W. Chen, "Joint Source Channel Rate-Distortion Analysis for Adaptive Mode Selection and Rate Control in Wireless Video Coding," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 12, no. 6, pp. 511-523, June 2002.

[48] R. Hekmat and P. V. Mieghem, "Interference in Wireless Multi-Hop Ad-Hoc Networks and Its Effect on Network Capacity," *Wireless Networks,* vol. 10 no. 4, pp. 389-399, July 2004.

[49] P. Herhold, E. Zimmermann, and G. Fettweis, "A Simple Cooperative Extension to Wireless Relaying," in *Proc. of the IEEE International Zurich Seminar on Comm.,* 2004.

[50] U. Horn, G. Girod, and B. Belzer, "Scalable Video Coding for Multimedia Applications and Robust Transmission over Wireless Channels,"in *Proc. of the International Workshop on Packet Video*, 1996.

[51] F. Ikegami, T. Takeuchi, and S. Yoshida, "Theoretical Prediction of Mean Field Strength for Urban Mobile Radio, " *IEEE Trans. on Anten. and Prop.,* vol. 39, no. 3, pp. 299-302, Mar. 1991.

[52] IS-856-A, *cdma2000 High Rate Packet Data Air Interface Specification,* 3GPP2, C.S0024-A, v3.0, Sep. 2006.

[53] IS-856-B, *cdma2000 High Rate Packet Data Air Interface Specification,* 3GPP2, C.S0024-B, v2.0, Apr. 2007.

[54] ITU-R, "Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000,"Recommendation ITU-R M.1225, 1997.

[55] W. C. Jakes and D. C. Cox, "Microwave Mobile Communications " *Wiley-IEEE Press*, 1994.

[56] H. Jenkac, T. Stockhammer, and W. Xu, "Cross-Layer Assisted Reliability Design for Wireless Multimedia Broadcast,"*Signal Proc.*, vol. 86, no. 8, pp. 1933-1949, Aug. 2006.

[57] H. Jiang and W. Zhuang, "Cross-Layer Resource Allocation for Integrated Voice/Data Traffic in Wireless Cellular Networks,"*IEEE Trans. on Wireless Comm.*, vol. 5, no. 2, pp. 457-468, Feb. 2006.

[58] N. Jiang, "Wireless Broadcast Services,"M.Sc. Thesis, Rutgers University, Oct. 2002.

[59] A. Jovicic, P. Viswanath, and S. R. Kulkarni, "Upper Bounds to Transport Capacity of Wireless Networks, " *IEEE Trans. on Info. Theory,* vol. 50, no. 11, pp. 2555-2565, Nov. 2004.

[60] D. Julian, M. Chiang, D. O'Neill and S. Boyd, "QoS and Fairness Constrained Convex Optimization of Resource Allocation for Wireless Cellular and Ad Hoc Networks, "in *Proc. of the IEEE INFOCOM*, June 2002.

[61] JVT, AHG Report: JSVM & WD Test, SVC Software JVT of ISO/IEC MPEG & ITU-T VCEG Document, JVT-O007, Apr. 16–22, 2005.

[62] A. Keshavarz-Haddad and R. Riedi, "On the Broadcast Capacity of Multihop Wireless Networks: Interplay of Power, Density and Interference," in *Proc. of the SECON*, 2007.

[63] M. Khansari, A. Jalali, E. Dubois, and P. Mermelstein, "Low Bit-Rate Video Transmission over Fading Channels for Wireless Microcellular Systems," *IEEE Trans. on Circ. and Sys. for Video Tech.*, vol. 6, no. 1, pp. 1–11, Feb. 1996.

[64] M. Kornfeld and G. May, "DVB-H and IP Datacast—Broadcast to Handheld Devices " *IEEE Trans. on Broadcasting*, vol. 53, no. 1, pp. 161-170, Mar. 2007.

[65] M. Kornfeld and K. Daoud, "The DVB-H Mobile Broadcast Standard [Standards in a Nutshell] " *IEEE Sig. Proc. Mag.*, vol. 25, no. 4, pp. 118-122, 127, July 2008.

[66] A. Ksentini, M. Naimi, and A. Gueroui, "Toward an Improvement of H.264 Video Transmission over IEEE 802.11e through a Cross-Layer Architecture," *IEEE Comm. Mag.*, vol. 44, no. 1, pp. 107-114, Jan. 2006.

[67] H. Kwon, Y. Kim, J. K. Han, D. Kim, H. W. Lee, and Y. K. Kim, "Performance Evaluation of High-Speed Packet Enhancement on cdma2000 1xEV-DV," *IEEE Comm. Mag.*, vol. 43, pp.67–73, Apr. 2005.

[68] S. Kwon, K. R. Rao, O.Kwon, and T. Kim, "Joint Bandwidth Allocation for User-Required Picture Quality Ratio Among Multiple Video Sources," *IEEE Trans. on Broadcasting*, vol. 51, no. 3, pp. 287–295, Sep. 2005.

[69] T. Kwon, H. Lee, S. Choi, J. Kim, D. H. Cho, S. Cho, S. Yun, W. H. Park, and K. Kim, "Design and Implementation of a Simulator Based on a Cross-Layer Protocol between MAC and PHY Layers in a WiBro Compatible IEEE 802.16e OFDMA System," *IEEE Comm. Mag.*, vol. 43, no. 12, pp. 136-146, Dec. 2005.

[70] W. C. Y. Lee, *Mobile Communications Engineering: Theory and Applications,* McGraw-Hill, 1997.

[71] O. Leveque and E. Preissmann, "Scaling Laws for One-Dimensional Ad Hoc Wireless Networks," *IEEE Trans. on Info. Theory*, vol. 51, no. 11, pp. 3987-3991, Nov. 2005.

[72] Y. Lim, P. Floquet, and X. Joulia, "Multiobjective Optimization Considering Economics and Environmental Impact,"in *Proc. of the ECCE2*, 1999.

[73] B. Liu, Z. Liu, and D. Towsley, "On the Capacity of Hybrid Wireless Networks," in *Proc. of the IEEE INFOCOM*, 2003.

[74] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-Layer Scheduling with Prescribed QoS Guarantees in Adaptive Wireless Networks,"*IEEE Journal on Sel. Areas in Comm.*, vol. 23, no. 5, pp. 1056-1066, May 2005.

[75] R. Love, A. Ghosh, X. Weimin, and R. Ratasuk, "Performance of 3GPP High Speed Downlink Packet Access (HSDPA),"in *Proc. of the IEEE VTC-Fall*, 2004.

[76] V. H. MacDonald, "The cellular concept," *Bell System Technical Journal,* vol. 58, no. 1, pp.15-41, 1979.

[77] U. Madhow, M. L. Honig and K. Steiglitz, "Optimization of Wireless Resources for Personal Communications Mobility Tracking,"*IEEE/ACM Trans. on Networking*, vol. 3, no. 6, pp. 698-707, Dec. 1995.

[78] H. S. Malvar, A. Hallapuro, and M. Karczewicz, and L. Kerofsky "Low-Complexity Transform and Quantization in H.264/AVC," *IEEE Trans. on Circ. and Syst. for Video Techn.*, vol. 13, no. 7, pp. 598-603, July 2003.

[79] I. Maric and R. D. Yates, "Cooperative Multihop Broadcast for Wireless Networks," *IEEE Journal on Sel. Areas in Comm.,* vol. 22, no. 6, pp. 1080-1088, Aug. 2004.

[80] D. Marpe, H. Schwarz, and T. Wiegand, "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard," *IEEE Trans. on Circ. and Syst. for Video Techn.*, vol. 13, no. 7, pp. 620-636, July 2003.

[81] D. Marpe, T. Wiegand, and G. J. Sullivan, "The H.264/MPEG4 Advanced Video Coding Standard and its Applications,"*IEEE Comm. Mag.*, vol. 44, no. 8, pp. 134-143, Aug. 2006.

[82] A. Mawira, "Models for the spatial correlation functions of the log-normal component of the variability of VHF/UHF field strength in urban environment,"in *Proc. of the IEEE PIMRC,* 1992.

[83] B. Sirkeci-Mergen and M. Gastpar, "On the Broadcast Capacity of Wireless Networks," in *Proc. of the Info. Theory and App. Workshop*, 2007.

[84] A. Molisch, *Wireless Communications,* Wiley-IEEE Press, 2005.

[85] *Multimedia Broadcast/Multicast Service (MBMS) User Services; Stage 1 (Release 8),* 3GPP TS 22.146, v9.0.0, June 2008.

[86] *Multimedia Broadcast/Multicast Service; Stage 1 (Release 9),* 3GPP TS 22.246, v8.5.0, Mar. 2008.

[87] *Multimedia Broadcast/Multicast Service (MBMS); Architecture and Functional Description (Release 8),* 3GPP TS 23.246, v8.2.0, June 2008.

[88] *Multimedia Broadcast/Multicast Service (MBMS); Protocols and Codecs (Release 8),* 3GPP TS 26.346, v8.1.0, Dec. 2008.

[89] M. J. Neely and E. Modiano, "Capacity and Delay Tradeoffs for Ad Hoc Mobile Networks," *IEEE Trans. on Info. Theory*, vol. 51, no. 6, pp. 1917-1937, June 2005.

[90] R. Negi and A. Rajeswaran, "Capacity of power constrained ad-hoc networks," in *Proc. of the IEEE INFOCOM*, 2004.

[91] Nokia, *1xEV-DO Evaluation Methodology,* 3GPP2 Technical Specification C30-DOAH-20030505-004, May 5, 2003.

[92] J. R. Ohm, "Advances in Scalable Video Coding,"in *Proc. of the IEEE*, vol. 93, no. 1, pp. 42–56, Jan. 2005.

[93] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi "Video Coding with H.264/AVC: Tools, Performance, and Complexity," *IEEE Circ. and Syst. Mag.*, vol. 4, no. 1, pp. 7-28, First Quarter, 2004.

[94] T. Ozcelebi, M.O. Sunay, A.M. Tekalp, and M.R. Civanlar, "Cross-Layer Optimized Rate Adaptation and Scheduling for Multiple-User Wireless Video Streaming," *IEEE Journal on Sel. Areas in Comm.*, vol. 25, no. 4, pp. 760-769, May 2007.

[95] S. Parkvall, E. Englund, M. Lundevall, and J. Torsner, " Evolving 3G Mobile Systems: Broadband and Broadcast Services in WCDMA, " *IEEE Comm. Mag.,* vol. 44, no. 2, pp. 30-36, Feb. 2006.

[96] J. D. Parsons, *The Mobile Radio Propagation Channel,* John Wiles & Sons, 2000.

[97] A. Qayyum, L. Viennot, and A. Laouiti, "Multipoint Relaying for Flooding Broadcast Messages in Mobile Wireless Networks," in *Proc. of the HICSS*, 2002.

[98] J. Qiu and G. Zhu, "An Adaptive Cross-Layer Video Transmission Scheme over Wireless Channels,"in *Proc. of the IEEE ISPACS*, 2005.

[99] Qualcomm, "MediaFlo: Flo Technology Overview,"December 11, 2008 [Online]. Available: http://www.mediaflo.com/news/pdf/tech_overview.pdf

[100] S. O. Rice, "Statistical properties of a sine wave plus noise," *Bell System Technical Journal*, vol. 27, pp. 109–157, Jan. 1948.

[101] A. Scaglione and Y.-W. Hong, "Opportunistic Large Arrays: Cooperative Transmission in Wireless Multihop Ad Hoc Networks to Reach Far Distances," *IEEE Trans. on Signal Proc.,* vol. 51, no. 8, pp. 2082-2092, Aug. 2003.

[102] A. Scaglione and M. van der Schaar, "Cross-Layer Resource Allocation for Delay Constrained Wireless Video Transmission,"in *Proc. of the IEEE ICASSP*, 2005.

[103] M. van der Schaar and J. Meehan, "Robust Transmission of MPEG-4 Scalable Video over 4G Wireless Networks,"in *Proc. of the IEEE ICIP*, 2002.

[104] M. van der Schaar and S. Shankar N, "Cross-Layer Wireless Multimedia Transmission: Challenges, Principles, and New Paradigms," *IEEE Wireless Comm.*, vol. 12, no. 4, pp. 50-58, Aug. 2005.

[105] T. Schierl, H. Schwarz, D. Marpe, and T. Wiegand, "Wireless Broadcasting using the Scalable Extension of H.264/AVC,"in *Proc. of the ICME*, 2005.

[106] H. Schwarz, D. Marpe, T. Schierl, and T. Wiegand, "Combined Scalability Support for the Scalable Extension of H.264/AVC,"in *Proc. of ICME*, 2005.

[107] H. Schwarz, D. Marpe, and T. Wiegand, "Hierarchical B Pictures,"JVT Standards Contribution, JVT_014, 23-29 July 2005.

[108] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard, "*IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 17, no. 9, pp. 1103-1120, Sep. 2007.

[109] H. Seferoglu, Y. Altunbasak, O. Gurbuz, and O. Ercetin, "Rate Distortion Optimized Joint ARQ-FEC Scheme for Real-Time Wireless Multimedia,"in *Proc. of the IEEE ICC*, 2005.

[110] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, " Cross-Layer Design for Wireless Networks," *IEEE Comm. Mag.*, vol. 41, no. 10, pp. 74-80, Oct. 2003.

[111] G. Sharma, R. Mazumdar, N. B. Shroff, "Delay and Capacity Trade-Offs in Mobile Ad Hoc Networks: A Global Perspective," *IEEE/ACM Trans. on Networking*, vol. 15, no. 5, pp. 981-992, Oct. 2007.

[112] K. Shen and E. J. Delp, " Wavelet Based Rate Scalable Video Compression, "*IEEE Trans. on Circ. and Syst. for Video Tech.*, vol. 9, no. 1, pp. 109–122, Feb. 1999.

[113] B. Sirkeci-Mergen, A. Scaglione, and G. Mergen, "Asymptotic Analysis of Multistage Cooperative Broadcast in Wireless Networks," *IEEE Trans. on Info. Theory,* vol. 52, no. 6, pp. 2531-2550, June 2006.

[114] L. Song and D. Hatzinakos, "Broadcasting Energy Efficiency Limits in Wireless Networks," *IEEE Trans. on Wireless Comm.*, vol. 7, no. 7, pp. 2502-2511, July 2008.

[115] V. Srivastava and M. Motani, "Cross-Layer Design: A Survey and the Road Ahead,"*IEEE Comm. Mag.*, vol. 43, no. 12, pp. 112-119, Dec. 2005.

[116] T. Stockhammer, T. Wiegand, T. Oelbaum, and F. Obermeier, " Video Coding and Transport Layer Techniques for H.264/AVC-based Transmission over Packet-Lossy Networks," in *Proc. of the IEEE ICIP*, 2003.

[117] T. Stockhammer, "Is Fine-Granular Scalable Video Coding Beneficial for Wireless Video Applications?, "in *Proc. of the IEEE ICME*, 2003.

[118] T. Stockhammer and M. M. Hannuksela, "H.264/AVC Video for Wireless Transmission, "*IEEE Wireless Comm.*, vol. 12, no. 4, pp. 6-13, Aug. 2005.

[119] G.L. Stüber, Principles of Mobile Communication, Kluwer Academic Publishers, 2001.

[120] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of Video Transmission over Lossy Channels, "*IEEE Journal on Sel. Areas in Comm.*, vol. 18, no. 6, pp. 1012-1032, June 2000.

[121] TIA/EIA/IS-856, *cdma2000 High Rate Packet Data Air Interface Specification,* 3GPP2, C.S0024-0, v4.0, Oct. 2002.

[122] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication,* Cambridge University Press, 2005.

[123] G. L. Turin, "Introduction to Spread Spectrum Antimultipath Techniques and Their Application to Urban Digital Radio," *Proc. of the IEEE,* vol. 68, no. 3, pp. 328-353, Mar. 1980.

[124] U.S. Patent 0,586,193 "Transmitting electrical signals," filed December 1896, patented July, 1897.

[125] *Video Coding for Low Bit Rate Communication,* ITU-T, H.263, ver. 2, Feb. 1998.

[126] A. J. Viterbi, "Very Low Rate Convolutional Codes for Maximum Theoretical Performance of Spread-Spectrum Multiple-Access Channels," *IEEE Journal on Sel. Areas in Comm.*, vol. 8, no. 4, pp. 641–649, May 1990.

[127] J. Walfisch and H. L. Bertoni, "A Theoretical Model of UHF Propagation in Urban Environments, " *IEEE Trans. on Anten. and Prop.,* vol. 36, no. 12, pp. 1788-1796, Dec. 1988.

[128] T. Wang, A. Cano, G. B. Giannakis, and J. N. Laneman, "High-Performance Cooperative Demodulation With Decode-and-Forward Relays," *IEEE Trans. on Comm.,* vol. 55, no. 7, pp. 1427-1438, July 2007.

[129] S. Wei, D. L. Goeckel, and M. C. Valenti, "Asynchronous Cooperative Diversity," *IEEE Trans. on Wireless Comm.*, vol. 5, no. 6, pp. 1547-1557, June 2006.

[130] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," *IEEE Trans. on Circ. and Syst. for Video Tech.,* vol. 13, no. 7, pp. 560-576, July 2003.

[131] M. Wien, H. Schwarz, and T. Oelbaum, "Performance Analysis of SVC," *IEEE Trans. on Circ. and Syst. for Video Tech.,* vol. 17, no. 9, pp. 1194-1203, Sep. 2007.

[132] J. E. Wieselthier, G. D. Nguyen, C. M. Barnhart and A. Ephremides, "A Problem of Constrained Optimization for Bandwidth Allocation in High-Speed and Wireless Communication Networks, "in *Proc. of the IEEE Decision and Control*, Dec. 1996.

[133] W. W. S. Wong and E. S. Sousa, "Performance Optimization of Single Frequency Broadcast Systems in FDD-CDMA Cellular Bands for Wireless Multimedia Services, "in *Proc. of the IEEE VTS-Fall*, Sep. 2000.

[134] J. W. Woods, *Multidimensional Signal, Image, and Video Processing and Coding,* Academic Press, 2006.

[135] D. Wu, Y. T. Hou, and Y. Q. Zhang, "Scalable Video Coding and Transport over Broad-Band Wireless Networks,"in *Proc. of the IEEE*, vol. 89, no. 1, pp. 6–20, Jan. 2001.

[136] L.-L. Xie and P. R. Kumar, "A Network Information Theory for Wireless Communication: Scaling Laws and Optimal Operation," *IEEE Trans. on Info. Theory*, vol. 50, no. 5, pp. 748-767, May 2004.

[137] M. D. Yacoub, J. E. V. Bautista, and L. Guerra de Rezende Guedes, "On Higher Order Statistics of the Nakagami-m Distribution, " *IEEE Trans. on Veh. Tech.*, vol. 48, no. 3, pp. 790-794, May 1999.

[138] R. Zheng, "Information Dissemination in Power-constrained Wireless Networks," in *Proc. of the IEEE INFOCOM*, 2006.