Cross-layer Design for Wireless Mesh Networks with Advanced Physical and Network Layer Techniques

Mina Yazdanpanah

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montréal, Québec, Canada

August 2012

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis proposal prepared

By: **Mina Yazdanpanah**

Entitled: **Cross-layer Design for Wireless Mesh Networks with Advanced Physical and Network Layer Techniques**

and submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____________________ Dr. Peter Grogono, Chair

____________________ Dr. François Gagnon, External examiner

____________________ Dr. Clement Lam, Examiner

____________________ Dr. Anjali Agarwal, Examiner

____________________ Dr. Dongyu Qiu, Examiner

____________________ Dr. Yousef Shayan, Co-supervisor

____________________ Dr. Chadi Assi, Co-supervisor

Approved by ______________________________

Chair of the ECE Department

____

Dean of Engineering

ABSTRACT

Cross-layer Design for Wireless Mesh Networks with Advanced Physical and Network Layer Techniques

Mina Yazdanpanah, Ph.D.

Concordia Unviersity, 2012

Cross-layer optimization is an essential tool for designing wireless network protocols. We present a cross-layer optimization framework for wireless networks where at each node, various smart antenna techniques such as beam-forming, spatial division multiple access and spatial division multiplexing are employed. These techniques provide interference suppression, capability for simultaneous communication with several nodes and transmission with higher data rates, respectively. By integrating different combinations of these multi-antenna techniques in physical layer with various constraints from MAC and network layers, three Mixed Integer Linear Programming models are presented to minimize the scheduling period. Since these optimization problems are combinatorially complex, the optimal solution is approached by a Column Generation (CG) decomposition method. Our numerical results show that the resulted directive, multiple access and multiplexing gains combined with scheduling, effectively increase both the spatial reuse and the capacity of the links and therefore enhance the achievable system throughput.

The introduced cross-layer approach is also extended to consider heterogeneous networks where we present a multi-criteria optimization framework to model the design problem with an objective of jointly minimizing the cost of deployment and the scheduling period. Our results reveal the significant benefits of this joint design method.

We also investigate the achievable performance gain that network coding (with opportunistic listening) when combined with Successive Interference Cancellation (SIC) brings to a multi-hop wireless network. We develop a cross-layer formulation in which SIC enables concurrent receptions from multiple transmitters and network coding reduces the transmission time-slot for minimizing the scheduling time. To solve this combinatorially complex non-linear problem, we decompose it to two linear sub-problems; namely opportunistic network coding aware routing, and scheduling sub-problems. Our results affirm our expectation for a remarkable performance improvement when both techniques are jointly used.

Further, we develop an optimization model for combining SIC with power control (PC). Our model optimally adjusts the transmission power of nodes to avoid interference on unintended receivers and properly embraces undesired interference through SIC. Therefore, it provides a balance between usage of PC and SIC at the transmitting and receiving sides, respectively. Our results show considerable throughput improvement in dense and heavily loaded networks.

To my dearest parents, Fatemeh and Mohamad

and to my beloved husband, Mehran

ACKNOWLEDGEMENTS

During my challenging PhD studies I was incredibly lucky to meet, work and become friend with such great people who helped me to progress further in my life. I would like to take this opportunity to name a few of them and express my sincere gratitude to them.

First, I would like to thank my supervisors; Dr. Yousef Shayan and Dr. Chadi Assi for the time and effort they dedicated to this research. Dr. Shayan has been a wise and honest professor who always helped me overcome the difficulties along the road. I am always grateful to him for all invaluable advice, constant supports, encouragements, and opportunities he provided me with to learn and discover my potential abilities. Dr. Assi was always a hardworking professor whose enormous passion for research, his continuous follow-ups, prompt feedbacks, supports and encouragements always brought great enjoyment to my research. I am always grateful to him for all the invaluable ideas, advice and suggestions.

I would like to extend my appreciation to my doctoral committee members, Dr. François Gagnon, Dr. Clement Lam, Dr. Cameron Skinner, Dr. Anjali Agarwal and Dr. Dongyu Qiu. Also, I would like to thank Dr. Samir Sebbah for the insightful and enlightening discussions I had with him. I also extend my thanks to the faculty and staff of Electrical and Computer Engineering Department at Concordia University, specially Ms. Pamela J.Fox, Ms. Diane Moffat and Ms. Connie Cianciarelli.

I would especially like to thank all of my great friends in Montreal, with whom I could share my sadness and happiness and be confident of receiving their supports. In this limited space I can only name a few of them, Fatemeh, Elham, Zohreh, Juel, Behnoush, Azam, Zeinab, Motahareh, Heather, Nassim, Fatemeh, Mohsen, John, Hirbod, Amin, Toma, Shahin, Mojtaba, Farjad and Mostafa.

Finally, I am indebted to my best parents Fatemeh and Mohamad for their invaluable and never-ending love, support and encouragements. They have always been examples to follow as honest and hardworking people with strong motivations. Also many thanks go to Poone and Minoo for being such wonderful sisters to me. Last but not least, I wish to thank my best friend and beloved husband, Mehran, for his endless love, patience, support and understanding.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations and Symbols

AP	Access Point
BER	Bit Error Rate
BF	Beam-forming
BLP	Binary Linear Programming
BM	Base Model
BW	Bandwidth
CDMA	Code Division Multiple Access
CG	Column Generation
CPU	Computation Time
CSI	Channel State Information
CSMA	Carrier Sense Multiple Access
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
DoF	Degrees of Freedom
DPC	Dirty Paper Coding
ILP	Integer Linear Programming
LP	Linear Programming
MAC	Medium Access Control
MILP	Mixed Integer Linear Programming
MIMO	Multiple Input Multiple Output
MMSE-SIC	Minimum Mean Square Estimation combined with SIC
MPR	Multiple Packet Reception
MPT	Multiple Packet Transmission
NC	Network Coding
NLP	Non-Linear Programming

NC-MPR	Network Coding combined with Multi-Packet Reception
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PER	Packet Error Rate
PC	Power Control
RC	Reduced Cost
RF	Radio Frequency
RM	Restricted Master
Rx	Receiving side of the link
SC-FDMA	Single Carrier Frequency Division Multiple Access
SIC	Successive Interference Cancellation
SINR	Signal to Interference and Noise Ratio
SDMA	Spatial Division Multiple Access
SDM	Spatial Division Multiplexing
SPC	Super Position Coding
TP	Transmission Pattern
TDMA	Time Division Multiple Access
Tx	Transmitting side of the link
UWB	Ultra Wide Band
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WMN	Wireless Mesh Network
C	Set of interfering links without any common nodes
C_i^+	$= \{j \in L : (i,j) \in \mathbf{C}\}$
C_i^-	$= \{j \in L : (j,i) \in \mathbf{C}\}$
$G_{n,n'}$	Path loss from node n to node n'
H	Channel coefficients matrix

Ir_n	Interference range of node n
L	Set of all links in the network
L_n	Set of all links connected to node n
L_n^+	$= \{i \in L : t(i) = n\}$
L_n^-	$= \{i \in L : r(i) = n\}$
$Lngth^p$	Length of configuration p
M	Set of all traffic sessions
M_i^p	A big integer for linearizing SINR constraints i p
M_i^s	A big integer for linearizing SINR constraints in s
$M_{n,i}^p$	An integer for linearizing power control constraints
N	Set of all nodes in the network
N_a	Set of all active nodes
N_R	Number of antenna elements at Rx
N_T	Number of antenna elements at Tx
P	Constant transmission power
P_{max}	Maximum transmission power
P_n	Transmitting power at n
P_n^p	Transmitting power of n at configuration p
$\mathcal{P}$	Set of all the configurations
$\mathcal{P}_0$	A sub-set of all the configurations
$\bar{P}$	Total number of configurations
Q	Maximum number of coding components
R_m	Commodity of session m
$\mathcal{S}$	Set of all time slots in the current configuration
Tr_n	Transmission range of node n
U	Weight vector at Tx
V	Weight vector at Rx

$W_{\xi_n^q}$	Maximum codable flow by coding component ξ_n^q
W_i	Uncoded traffic on link i
$W_{(n_i,n_j)}$	Remaining traffic to be sent from n_i to n_j
c_i	Time-invariant capacity of link i
d_m	Destination node of session m
$d_{n,n'}$	Distance between nodes n and n'
d_0	A close-in distance to n with P_n as the received power
$e_{n,i}^p$	A binary variable for linearizing power control constraints
f_i^m	Amount of flow from session m on link i
$h_{i,j}$	Channel coefficient between antennas i and j
$h_{n,i}^p$	$= e_{n,i}^p P_n^p$
i	One wireless link
k_n	Maximum number of decodable packets at n
m	One session
n	One node
p	One configuration
$r(i)$	Receiver of link i
$r(t)$	Received signal
s	One time slot within a configuration
s_m	Source node of session m
$s(t)$	Transmitted signal
$t(i)$	Transmitter of link i
t_i^p	$= 1$ if i carries uncoded packets not to be encoded
u_n^p	$= 1$ if node n is active in p
$u_{i,\xi_{r(i)}^q}^p$	$= 1$ if i carries uncoded packets to be coded using $\xi_{r(i)}^q$
v_i^p	$= 1$ if link i is active in p
v_i^s	$= 1$ if link i is active in s

v^s_{i,ζ^h_n}	$= 1$ if active link i is involved in transmission patternζ^h_n
$x_j(t)$	Received propagated signal
x^s	$= 1$ if there is at leat one active link in s
$\bar{y}_i$	Dual value of the BW constraint for link i
z^p_i	Number of active streams on i in configuration p
$z^p_{i,\xi^q_{t(i)}}$	$= 1$ if i carries coded packets using $\xi^q_{t(i)}$
Γ	SINR threshold
Ξ_n	Set of all available coding components at node n
$\Upsilon^{\zeta^h_n}$	$= 1$ if ζ^h_n is to be scheduled in current configuration
Ψ_m	Set of predetermined k-shortest paths for session m
α_n	Effective transmit DoFs
$\alpha v(i,j)$	$= \alpha_{t(i)} v^p_j$
β	Path loss exponent
β_n	Effective receive DoFs
$\beta v(i,j)$	$= \beta_{r(i)} v^p_j$
$\gamma^p_{i,j}$	$= 1$ if $t(i)$ nullifies its signal at $r(j)$
δ	Interference range coefficient
δ^i_ψ	$= 1$ if link i is traversed by path ψ
$\delta^{i,j}_\psi$	$= 1$ if two consecutive links i and j, are traversed by ψ
ζ^h_n	The h^{th} transmission pattern at node n
η	Background noise power in the operation frequency band
$\theta^p_{i,j}$	DoFs assigned by $t(i)$ to nullify its signal at $r(j)$
$\kappa^{\zeta^h_n}$	Number of links on which n transmits in ζ^h_n
λ^p	Frequency of configuration p being active
$\mu^p_{i,j}$	$= 1$ if $r(j)$ suppresses interference of $t(i)$
ν^i	Dual value of BW constraints at i when $t^p_i = 1$
ξ^j	TP corresponding to the j^{th} coding component

ξ_n^q	The q^{th} coding component at node n
ρ	Average number of DoFs at all nodes
$\sigma^i_{\xi_n^q}$	Dual value of BW constraints at i when $z^p_{i,\xi^q_{t(i)}} = 1$
ς^i	TP corresponding to unicast transmission on link i
$\tau^i_{\xi_n^q}$	Dual value of BW constraints at i when $u^p_{i,\xi^q_{r(i)}} = 1$
$\phi^p_{i,j}$	DoFs assigned by $r(j)$ to suppress $t(i)$'s signal at $r(j)$
ψ	A predetermined path from the source to the destination
$\omega^{n_1 n_2 n_3}$	Amount of traffic flow from n_1 to n_3 through n_2
ω^m_ψ	Amount of session m on path ψ

Chapter 1

Introduction

Wireless Mesh Networks (WMNs) have recently emerged as a solution for the last mile access network to provide high speed Internet conectivity. These networks are envisioned to replace their wired counterpart, due to their easy and low cost deployment and maintenance.

A WMN is typically a multihop network consisting of a number of stationary wireless routers interconnected by wireless links to provide a backbone over which the clients can access the Internet. The clients of WMNs are mobile nodes and end-users in residential areas and offices. The coverage of wireless communication networks in this technology is extended by supporting multi-hop connections as shown in Figure 1.1. Wireless mesh routers beyond each other's transmission range can still communicate through message forwarding provided by the intermediate routers. To establish multi-hop connections, each individual router in the network coordinates with others to automatically find and maintain the routes to deliver the traffic. When a node joins the network for the first time, it searches for neighbor nodes and connects to the network through them. This self-configuration and self-organization feature is achieved without or with minimal human intervention. In addition, if one of the nodes disconnects from the network, other nodes can find

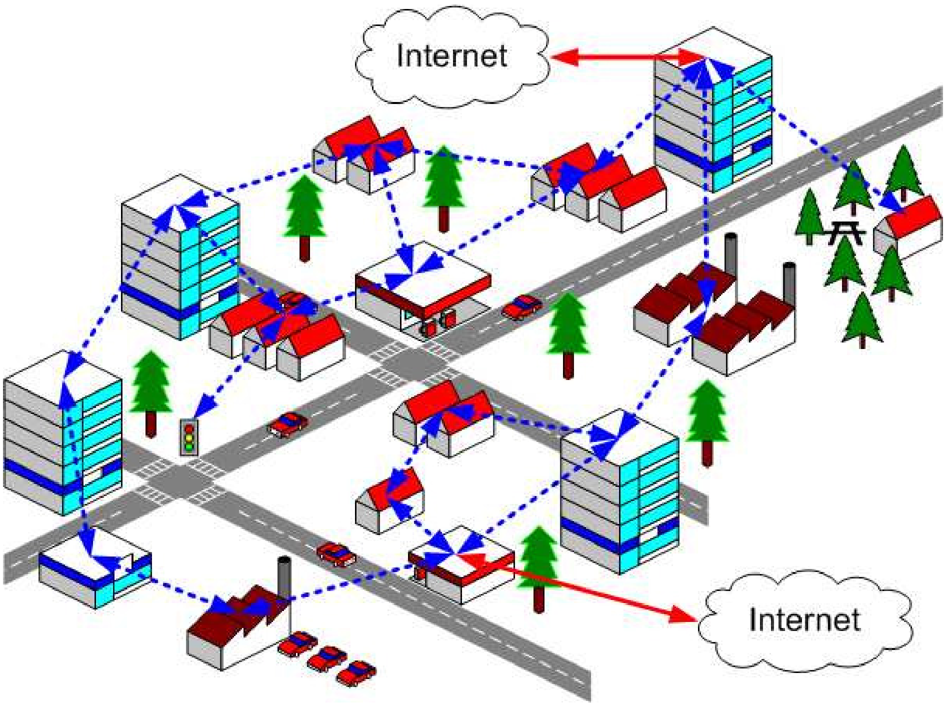


Figure 1.1: Conceptual illustration of multi-hop networks [41]

their routes through the rest of the nodes; this feature is known as the self-healing property of WMNs.

In WMNs only few of the routers are responsible for having wired connections with the Internet; these routers are commonly known as mesh gateways. A mesh gateway is responsible for delivering traffic between the Internet and end-users using mesh routers and multi-hop routes. WMNs are expected to provide various attractive applications, such as broadband Internet access, information storage, multimedia services, etc., to be competitive with their wired counterpart. Therefore, a major requirement for WMNs is to provide high-throughput services for end-users. As a result, studying the maximum throughput of WMNs and designing appropriate mesh protocols have been the subject of many research activities over the past few years. The throughput of such networks depends on the capacity of individual links between the senders and receivers which can be studied from Information Theory and physical layer point of view. To enhance the performance and throughput of

WMNs, advanced wireless radio technologies are employed, such as reconfigurable radios, frequency agile/cognitive radios, multiple antennas, multi-radio and multi-channel systems [46].

Now, in contrast to their wired counterparts, which have fixed capacities, wireless links suffer from inter-path and intra-path interference, which creates detrimental effect on their reliability and ultimately the overall network performance. Therefore, to benefit from any radio (physical layer) technology in enhancing the network throughput, the capabilities of this technique should be exploited at higher layer protocols as well. These issues should be considered and properly managed when evaluating the overall throughput of such systems. There are several practical approaches for minimizing the effect of these interferences. Mechanisms that deal with interference problems are usually defined and handled at the Medium Access Control (MAC) layer. A MAC protocol introduces a set of rules for scheduling or activating links in the network while maintaining interference under control; either through random access (e.g. Aloha, Carrier Sense Multiple Access (CSMA), CSMA with Collision Avoidance (CSMA/CA)and CSMA with Collision Detection (CSMA/CD)) or resource allocation based schemes (e.g. Orthogonal or Single Carrier Frequency Division Multiple Access (OFDMA/SC-FDMA) and Time Division Multiple Access (TDMA)). Therefore, the overall throughput of the network depends on how the links are scheduled to be active without causing interference on other neighboring active links and the way demands are routed over these scheduled links. Consequently, cross-layer design [63, 76, 137] is very important for maximizing the throughput of the network.

1.1 Motivation

Researchers have recently shown an increased interest in multi-antenna technology due to its capability in improving the throughput of wireless networks through several novel techniques. For example, multi-antenna Beam-forming (BF) technique suppresses the interference[1]coming from other active neighboring links and thus improves the spectrum spatial reuse [41]. Spatial Division Multiplexing (SDM), on the other hand, provides higher transmission data rate on each active link without requiring additional spectrum or power [6,37]. Finally, Spatial Division Multiple Access (SDMA) enables a single-radio node to communicate concurrently with several neighboring nodes in the network [11]. Some of these multi-antenna technologies have been included in wireless networking standards such as IEEE 802.11n [117], 802.11AC [128] and WiMAX [39,59]. Numerous companies such as Cisco, Entrasys, Trapeze Networks, Proxim Wireless, Xirrus and Motorola [1], have also started developing such technologies in new products and several multi-antenna-based networks are or will be deployed in the near future.

Nevertheless, to benefit from multi-antenna technology in enhancing the network throughput, the capabilities of this technique should be exploited at higher layer protocols as well. In other words, since both inter-flow and intra-flow interference are serious issues for wireless multihop networks and have detrimental effect on their performance, the overall throughput of the network will depend on how wireless links are scheduled to be active without causing interference on other neighboring active links and the way traffic demands are routed. It has been shown [75, 126] that although employing smart antennas significantly increases the performance, however, if both MAC and routing aspects are not properly considered in network design, then little or no performance gain may be achieved. Therefore, coordinating mechanisms at the physical layer, MAC layer

[1]More explanation on interference suppression is provided in Section 7.1.1.

and network layer is vital to obtain the potential capacity of multi-antenna wireless networks. Accordingly, cross-layer optimization is necessary for designing multi-antenna-based networks, and has been attracting considerable research attention [8, 10, 15, 17, 24, 25, 51, 54, 60, 67, 77, 78, 92, 93, 98, 99, 109, 120, 122, 129, 141].

In addition to empowering nodes with multi antennas, Multiple Packet Reception (MPR) [18,79,114,122,140,142] has been recently introduced to further mitigate the detrimental effect of interference by exploiting, rather than avoiding, it. MPR refers to the nodes capability of decoding more than one packet from an aggregated signal of simultaneous transmissions on the air. While this technique increases the per-node throughput, it also yields significant improvement in network throughput by allowing more concurrent transmissions [114]. MPR is indeed a physical layer method, as opposed to a MAC layer method used for resolving collisions and congestions [96]. Extensive research has been done to investigate the benefits of MPR which is implemented through sophisticated signal processing techniques, such as SDMA through multiple antennas [9, 55], Ultra Wideband (UWB) [45], Code Division Multiple Access (CDMA) [20, 106], Directional Antennas (DA) [31], Parallel and Successive Interference Cancellation (SIC) [61, 86, 87, 90, 102].

Other techniques (e.g. network coding) have recently emerged for improving the performance of wireless networks by reducing the transmission time slot overhead. Network coding improves the network capacity in both wired and wireless networks with multicast traffic [2]. This technique has shown to be more attractive in wireless networks due to the broadcast nature of the wireless transmission medium. For example, packets received by an intermediate node may be combined together using a linear XOR operation and broadcasted. Coded packets are decoded at the next hop of each broadcast transmitter where all packets, excluding the desired one, are formerly available. This simple scheme for network coding is referred to as COPE [66] and will be studied further in this thesis. There has been extensive cross-layer design and optimization to characterize the network throughput for the case of multiple unicast sessions with different opportunistic network coding for more than one packet through mathematically modeling the problem [4,7,35,74,94,105,107,136,137]. Despite these recent efforts to characterize the advantages of NC and MPR [38,124] on the performance of wireless multi-hop networks, there is still a lack of a thorough analysis of the benefits of these two advanced techniques when used concurrently. This will be a subject for further investigation in this thesis.

There are two main research directions we undertake in this thesis. The first one is to investigate the achievable performance gain that network coding when combined with Successive Interference Cancellation brings to a multi-hop wireless network, whereas the second one is to evaluate the effect of various multi-antenna techniques in wireless multi-hop networks. These objectives are elaborated next.

1.2 Problem Statement and Objectives

In this thesis, we investigate the effect of various physical layer and network layer techniques on the throughput of multihop wireless networks. Our main objective is to study the achievable network performance gains which successive interference cancellation, power control, network coding, beam-forming, spatial division multiple access and spatial division multiplexing bring to a multihop wireless network. To achieve our objective, we follow a cross-layer design approach. We mathematically formulate the complex problem of joint routing and link scheduling and solve it using standard decomposition methods. Our objective is to minimize the system activation time to satisfy certain traffic demands, where the system activation time refers to the schedule length during which the links in the network should be active for supporting a set of traffic commodities. In other words, given a certain traffic load, our system activation time refers to the period of time (seconds) the network would take to deliver all the sessions. It should be noted that the lower bound on the system activation time is equivalent to the upper bound of the network throughput.

While network architecture traditionally adopts layering to network coordination, where each layer controls a subset of the decision variables, and classically observes a subset of constant parameters and the variables from other layers [23], considering separated modules for each task and allowing the modules to communicate only with their adjacent modules is not efficient especially in wireless network designs. For instance, adjusting the resource allocation in the physical layer changes the average link rates, influences the optimal routing, and alters the achievable network throughput [62]. Therefore, by jointly designing mechanisms based on which physical-layer resource allocations, link scheduling and routing the traffic, is a necessity for recent network protocols [29, 69, 112]. While designing a layered scheme is intuitively scalable, however, optimization in a cross-layer framework is sophisticated and requires advanced optimization techniques [76, 137]. Hence, our modeling throughout this thesis will follow the cross-layer approach that has been widely used in designing and studying the performance of wireless networks.

1.3 Contributions

The significant achievements in this thesis are summarized as follows.

- We develop an original optimization model for combining successive interference cancellation (SIC) with power control at the physical layer to manage the interference and scheduling conflict-free links. While these two methods have shown remarkable effects on network throughput, to the best of our knowledge there is no work proposed in the literature which investigates their combination benefits. Our model optimally adjusts the transmission power of nodes to avoid causing interference on unintended receivers through power control. Moreover, it properly embraces undesired wireless interference through applying SIC at receiving nodes. Therefore, our model minimizes the total scheduling time by providing a balance between usage of power control at the transmitting side and SIC at the receiving side. Our results show the considerable benefits of this combination in dense and heavily loaded networks.

- We advocate an optimization framework for the joint routing, scheduling and network coding where all types of network coding components (with/without opportunistic listening) are exploited and we allow a number of packets to be coded. We introduce a novel scheduling scheme for activating network coding components instead of links. To the best of our knowledge, our proposed scheme is the first TDMA-based scheduling that manages network coding with opportunistic listening. We develop a decomposition method based on column generation for solving it. We make this problem linear and tractable by solving the network coding-based routing sub-problem (Restricted Master), and selecting network coding components and link scheduling in another sub-problem (Pricing).

- We investigate the achievable performance gain that network coding when combined with Successive Interference Cancellation (SIC) brings to a multi-hop wireless network. While SIC enables concurrent receptions from multiple transmitters, network coding reduces the transmission time-slot overhead and each of these techniques has shown great benefits for network performance. We present a cross-layer formulation for the joint routing and scheduling problem in a wireless network with network coding (with opportunistic listening) and SIC capabilities. We use the realistic signal to interference plus noise ratio (SINR) interference model. To solve this combinatorially complex non-linear problem, we decompose it (using Column Generation) to two linear sub-problems; namely opportunistic network coding aware routing, and scheduling subproblems. Our scheduling sub-problem consists of activating network coding components, rather than links, which do not interfere with each other and will be used to route the traffic. Our results affirm our expectation for a remarkable throughput improvement. NC combined with SIC shows more than 50% throughput gain in a medium-size dense network.

- We present a cross-layer optimization framework for wireless mesh networks where at each node, various smart antenna techniques such as beam-forming, spatial division multiple access and spatial division multiplexing are employed. These techniques provide interference suppression, capability for simultaneous communication with several nodes and transmission with higher data rates, respectively, through multiple antennas. By integrating different combinations of the multi-antenna techniques in physical layer with various constraints from MAC and network layers, three Mixed Integer Linear Programming (MILP) models are presented to minimize the system activation time. Since these optimization problems are complex to solve, the optimal solution is approached by a Column Generation decomposition method. The numerical results for different network scenarios with various node densities, number of antennas, transmission ranges and number of sessions are provided. It is shown that the resulted directive, multiple access and multiplexing gains combined with scheduling, effectively increase both the spectrum spatial reuse and the capacity of the links and therefore, enhance the achievable system throughput.

- We consider a multi-criteria optimization framework for the design problem in heterogeneous networks. In this thesis, a heterogeneous network refers to a multi-antenna network with various number of antennas at each node. The proposed model of optimization is a multi-criteria MILP that jointly minimizes the system activation time and the cost of deployment. We minimize the cost of deployment by obtaining the minimum number of antennas required at each node. To obtain the optimal solution of this large-sized multi-criteria problem, we have combined the ϵ-constraint method [34] with column generation decomposition approach and provided an MILP model which yields the optimal solution relatively fast. To the best of our knowledge, this is the first attempt to jointly optimize these two cost functions.

1.4 Structure of the Thesis

We present an overview of the literature and the optimization techniques used in this research in Chapter 2. In Chapter 3, we explain the basic principles of the mathematical modeling for joint routing and scheduling in multihop wireless networks. In Chapter 4, we introduce a cross-layer design for wireless networks when successive interference cancellation is exploited and node transmit with constant or adjustable power. In Chapter 5, we propose a joint routing and scheduling when network coding is applied. Then, we extend this model to a case where both network coding and successive interference cancellation are used in the same chapter. We present the numerical results of models for successive interference cancellation, network coding and their combinations in Chapter 6. In Chapter 7, we introduce three different models for varied multi-antenna techniques, namely, beam-forming, spatial division multiple access and spatial division multiplexing. Then, we propose a multi-criteria optimization model for minimizing both the scheduling time and the usage of antenna elements in a multihop wireless heterogeneous network. In Chapter 8, we present the numerical results of multi-antenna models. Finally, we conclude the thesis in Chapter 9 and we provide some future directions for this research.

Chapter 2

Literature Review and Optimization Techniques

We start this chapter by giving an overview of the literature pertaining to this thesis research. Then, we present a brief explanation about the optimization methods which are used in this thesis.

2.1 Literature Review

This section is divided into four parts where the first one surveys the literature related to the joint routing and scheduling with smart antenna technologies. The second one addresses different approaches to achieve multi-packet reception in wireless networks. In the third part, we overview previous research on network coding and the last part presents the literature pertaining to network coding combined with multi-packet reception.

2.1.1 Joint Routing and Scheduling with Smart Antenna Technologies

In recent years, there has been an increasing amount of research on cross-layer design in networks equipped with smart antennas [8, 10, 15, 17, 24, 25, 51, 54, 60, 67, 77, 78, 92, 93,98,99,109,120,122,129,141]. A sizable portion has proposed newly-designed MAC and routing protocols suitable for multi-antenna networks [24,25,54,93,109,129]. For instance, in [93] the authors considered both SDM and BF techniques in ad hoc networks and introduced their MAC protocol called HYB that provides the maximal utilization of multiple antennas under any environment. An algorithm is proposed in [25] to maximize the network throughput by taking advantage of physical layer channel information to opportunistically schedule cooperative spatial multiplexed transmissions between nodes. Joint MAC and routing protocols are investigated in [24, 54]. While the former focused on BF-based WMNs, the latter studied SDM-based ad hoc networks. In [129], a heuristic routing and scheduling algorithm is proposed for wireless backhaul networks with smart antennas where the interference aware tree is constructed. In [109] scheduling in adaptive cooperative relay transmission in multi-antenna ad hoc networks is studied for both centralized and distributed systems and a MAC protocol is proposed for the distributed scheduling case.

However, cross-layer design in multi-antenna networks have been addressed through mathematical modeling and optimization as well [8, 10, 15, 17, 51, 60, 67, 77, 78,92,120,122]. In [60,92,120], a cross-layer optimization between MAC and physical layer in SDM-based multi-antenna networks is addressed. The authors of [92] found a sub-optimal solution by heuristics and in [120], the problem of joint bandwidth allocation, element assignment and scheduling is studied. In [60], an analytical tractable model is proposed for the physical and link layers whereby the performance of a multi-hop multi-antenna network enabled with SDM and BF techniques is optimized. In [8], a heuristic algorithm is introduced for joint stream control and scheduling optimization in an SDM-basd multi-hop network and the authors of [15] introduced a non-linear programming model whereby BF, power control and scheduling in a multiple-input single-output downlink are jointly optimized for maximizing the minimum weighted rate among all users. Joint routing and scheduling optimization in SDMA-based multi-antenna networks has been studied in [122] where the sub-optimal solution for the network throughput is obtained through heuristics. However, the authors only considered multi-packet reception scenario of SDMA in their work.

Cross-layer optimization over network, MAC and physical layers is studied in [10, 17, 67, 77, 78]. An LP model is provided in [77] for distributed link scheduling, power control and routing in multi-hop multi-antenna wireless networks. The problem of routing, power allocation and bandwidth assignment in an FDMA multi-antenna ad hoc network is solved in [78]. A sub-optimal network throughput is obtained by proposing a heuristic centralized algorithm for joint routing, scheduling and stream controlling subject to fairness constraints in [10].

Among all of these works, only a few of them are closely related to our work. In [17] and [67], a centralized cross-layer optimization in BF-based multi-hop wireless backhaul networks is proposed. The authors have provided a primal-dual decomposition for joint routing and scheduling and adopted a heuristic algorithm to assist in determining interference-free feasible schedule. However, the problem formulation in our work is different where in addition to BF, SDMA and SDM techniques are employed. Moreover, the protocol model is adopted for modeling interference in our network.

Our mathematical optimization model to evaluate the throughput of the network is similar to [51], however the authors of [51] optimized the throughput by maximizing the sum of all flows in routing and did not consider scheduling in their work.

Moreover, they only evaluated the effect of BF and SDM techniques on their design. Additionally, they used some LP relaxation that provided sub-optimal solutions. While we have considered a centralized scheduling in the network, there are some other papers in the literature focused on distributed link scheduling [77, 98, 99, 141].

Although there is a large amount of research about optimization in heterogeneous networks, only [120] has focused on minimizing antenna elements at different nodes of the network and the rest have considered other parameters in their network designing.

2.1.2 Multi-packet Reception in Multi-hop Wireless Networks

In their seminal work, the authors of [48] have studied the capacity of random wireless networks and presented their study using both the physical and the protocol interference model. It has however been shown that the simple protocol model underestimates the network interference and thus yields overestimated capacity [58]. The authors of [124] have analyzed the capacity and energy efficiency of ad hoc networks with MPR under the physical model. They extracted a closed gap between upper and lower bounds of the throughput capacity. Furthur, interference under protocol model has been exploited in an MPR scheme for joint routing and scheduling in ad hoc networks [123].

MPR in MIMO-based networks links have been discussed in [55] and [9], where [55] proposes the first MAC protocol designed for MIMO/orthogonal frequency division multiplexing (OFDM) based wireless local area networks (WLANs). [9] introduced an asynchronous MPR method for the physical layer and also designed a compatible MAC for WLANs which reduces the congestion. The authors of [45] studied the throughput performance of a peer-to-peer asynchronous Code Division Multiple Access (CDMA) network supported by a full duplex Ultra Wideband (UWB) and showed the necessity of MPR as a feature to improve the throughput. CDMA has also been used in [20] and [106] to highlight the advantage of MPR scheme in multimedia networks and slotted ALOHA systems, respectively. The authors of [31] have proposed an LP to model the MPR through Directional Antennas(DA).

A ZigBee prototype of SIC has been built in [49] by using software radios and used experimental results to validate that SIC is an effective way to improve the system throughput. The authors of [84] have developed a greedy heuristic scheduling algorithm based on conflict set graph and studied the scheduling problem in an ad hoc network with SIC. The NP-hard link scheduling problem for wireless networks with SIC was also studied in [87], but the authors did not consider the practical physical model with the effect of combined interference. Later in [85], they used a new weighted graph to characterize the sequential detection nature of SIC. However, they found a near-optimal maximal feasible set of links which is constructed based on their greedy algorithm. The transmission cost of such a set is interpreted as the bound for the approximation performance. Asymptotic transmission capacity of ad hoc networks with SIC was studied in [14, 125] where the first work considered a simplified model where all signals from transmitters within a specific radius up to one hop, can be successfully decoded. However in [14] a more realistic SIC model is used. [61] proposes a cross-layer optimization framework that incorporates variables at physical, link, and network layers to maximize sum of the weighted flow rates where SIC and interference avoidance are combined. A CG-based LP formulation for joint routing and scheduling has been proposed in [90] to study the optimal max-min throughput of wireless multi-hop networks with SIC. The authors have considered predetermined number of decodable packets as 2 or 3. However, in [61] the authors have found that the limitation on the number of packets is based on the physical parameters in their system.

Among all the research work on SIC, only one of them ([102]) has claimed that the potential gain would be marginal. However, this result is not generally valid because their network example consisted of only two links; thus, the benefits of SIC would not be fully exploited.

2.1.3 Joint Routing and Scheduling with Network Coding

The benefits of network coding on multicast networks were characterized in [2]. However, these benefits are not fully clarified for multiple unicast sessions yet. Among a large body of work which is done on NC for unicast sessions, only some of them have done cross-layer optimization through LP or MILP model which are related to our work [7,35,64,74,105,107]. However, the interference has been considered under the protocol model in [7,74,105]. Physical interference model has been employed in the rest.

[35, 107] have studied joint routing and scheduling and NC for wireless multihop networks where physical interference model is used. However, network coding components with opportunistic listening were not considered in any of them. Recently, [64] introduced k-tuple coding, a generalization of pairwise coding with next-hop decodability where fully characterized the region of arrival rates where the network queues can be stabilized under this coding strategy.

Column Generation decomposition approach has also been employed for NC-based [35,90,107] or SIC-based [108] optimizations to tackle the problem and to find the optimal solution.

2.1.4 Network Coding Incorporated with Multi-packet Reception

Recently, the combination of advanced physical layer techniques has shown considerable improvement on wireless mesh network capacity [108] where SIC, Dirty

Paper Coding (DPC) and Superposition Coding (SPC) are combined. Interaction between interference and NC has also been studied in some literature. In [82], a new algorithm in two-path successive relay systems has been introduced to cancel the inter-relay interference through network coding.In [21], a novel cross layer method for interference cancellation and network coding that significantly increases the capacity of multi-hop wireless networks is proposed. In their work, the multiuser interference is cancelled through dirty-paper decoding. The authors of [121] have proposed an interference-known network coding-aware routing metric that considers the trade-off between increasing coding opportunity and decreasing wireless interference. They have used this metric in an LP to achieve a better throughput and less delay. In [32], a systematic mechanism for studying achievable rates in multiple unicast networks is obtained through linear NC and interference alignment.

To the best of our knowledge, there are only three works proposed in the literature that studied both MPR and NC [27,38,100]. The individual performance of NC and MPR has been highlited in [38]. However, the benefits of their combination are studied in [27, 100]. In [100], the authors analyzed the trade-off between NC without opportunistic listening and MPR in a special case of fully connected networks. In [27], a cross-layer design for the joint use of NC and MPR is proposed to relieve the congestion problem in 802.11 MAC. Although both of these works have shown promising benefits of NC incorporated with MPR, they may not accurately highlight the benefits since they have used the protocol scheme to model the interference.

Another type of network coding, namely Analog NC or Physical NC, was introduced and discussed in [65,110,139]. This method may show compatible performance to combined NC and MPR only if no opportunistic listening is used.

To the best of our knowledge, there is no approach tat jointly evaluates SIC incorporated with NC in a multi-hop network, and this is one of the main focus of this research.

2.2 Optimization Methods

In this section, we briefly overview the optimization methods that we use throughout this thesis. We start by explaining the following constrained minimization

$$\begin{aligned} &\min \quad F(x) \\ \text{subject to} & \\ &x \in S \end{aligned}$$

where $F(x)$ is called the *cost function*, the vector $x = (x_1, x_2, ...x_n)$ is the variable of this optimization problem and S denotes the feasibility region. This minimization problem can be generally formulated to the standard form of:

$$\begin{aligned} &\min \quad F(x) \\ \text{subject to} & \\ &g_i(x) \leq 0 \qquad i = 1, 2, ..., m \\ &x \in \mathbf{R}^n \end{aligned} \tag{2.1}$$

where the set of constraints $g_i(x) \leq 0$ are used to determine the boundaries of region S. We note that if the optimization problem is a maximization one, then $F(x)$ is called as the *utility function*. The solution to problem (2.1) is an $x^\star$ where:

$$F(x^\star) \;<\; F(x) \qquad \forall x \in S, x \neq x^\star$$

2.2.1 Linear and Non-linear Programming problem

In the optimization problem (2.1), if all the objective and the constraints functions are linear, i.e. $F(x) = \sum_{j=1}^{n} c_j x_j$ and $g_i(x) = \sum_{j=1}^{n} a_{ij} x_j - b_i,\ \forall i,\ \forall j,\ a_{ij}, b_i \in \mathbf{R}$, the problem is known as the *Linear Programming* (LP), and (2.1) modifies to:

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} c_j x_j \\ \text{subject to} \quad & \\ & \sum_{j=1}^{n} a_{ij} x_j \geq b_i \qquad i = 1, 2, ..., m \\ & x_j \geq 0 \qquad j = 1, 2, ...n \end{aligned} \tag{2.2}$$

However, in the case that any of these functions are non-linear, the problem is named *Non-Linear Programming* (NLP).

While there are several methods (e.g. Simplex algorithm [52]) for effectively solving LP problems, NLP problems are more sophisticated to deal with. One of the most efficient methods for solving NLPs is linearizing them into LPs. In Chapter 5 of this thesis we have linearize the non-linear objective of our optimization by decomposing the problem into sub-problems with linear objective functions.

2.2.2 Integer Linear Programming problem

In the optimization problem (2.1), if all the variables, which are the entries of x, are integers (i.e., $x \in \mathbf{Z}$), the problem is known as an *Integer Linear Programming* (ILP). *Binary Linear Programming* (BLP) is a especial case of ILP where the variables can be either 0 or 1. In the case that only some of the variables are integer or binary, the problem is called *Mixed Integer Linear Programming* (MILP). We note that solving ILP, BLP and MILP problems are more complex and time consuming in comparison with LPs [127]. There are different methods for solving these problems where one of them used in this research is known as Column Generation [26]. This method is described next.

2.2.3 Column Generation Decomposition Method

Column Generation (CG) is a mathematical technique which finds the exact optimal solution for LP problems [26, 80] and is suitable to deal with problems with large number of variables and constraints. This technique has been introduced to solve MILP problems through Branch-and-Price method which is a combination of CG and Branch-and-Bound [16] techniques. In CG approach, an MILP problem is decomposed into two smaller sub-problems; namely, the Master and the Pricing problems. Let us denote the Master problem by:

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}} c_j x_j \\ \text{subject to} \quad & \\ & \sum_{j \in \mathcal{J}} a_j x_j \geq b \\ & x_j \geq 0 \qquad j \in \mathcal{J} \end{aligned} \tag{2.3}$$

where a_j represents one column, as shown in Figure 2.1, which is a potential solution of the Pricing problem. We assume $\mathcal{A}$ is an non-empty set where a_j $(j \in \mathcal{J})$ belong to it. The Master problem is solved over only a subset of available columns $\mathcal{J}_0 \subset \mathcal{J}$ and it is called the Restricted Master (RM) problem which finds the solution relatively faster and easier due to the fact that $|\mathcal{J}_0| \leq |\mathcal{J}|$. Therefore, the RM problem is given by:

$$\begin{aligned} \min \quad & \sum_{j \in \mathcal{J}_0} c_j x_j \\ \text{subject to} \quad & \\ & \sum_{j \in \mathcal{J}_0} a_j x_j \geq b \\ & x_j \geq 0 \qquad j \in \mathcal{J}_0 \end{aligned} \tag{2.4}$$

The RM and the Pricing problems exchange some parameters and are solved iteratively until eventually finding the optimal solution of the RM. The solution of each

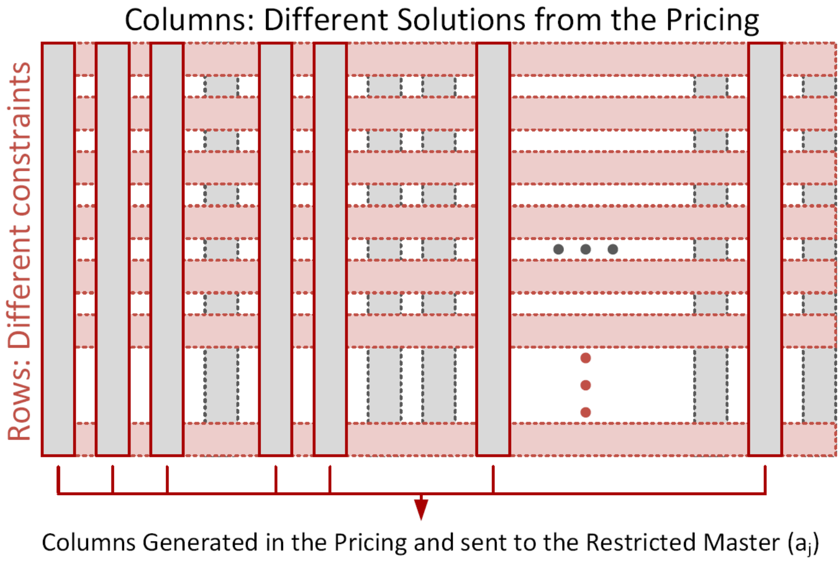


Figure 2.1: Illustration of an MILP getting solved by CG approach

sub-problem is treated as a parameter in the other one. In each iteration of solving the Pricing, a new column is generated and sent to the RM. However, for solving the RM in its first iteration, an initial set of columns which is a feasible solution to the Pricing is available to the RM. Let us denote the non-negative dual variables of the RM by the vector y. In each iteration, we denote the corresponding dual variables to the optimal solution of the RM, i.e. $x^\star$, by $y^\star$ which is passed to the Pricing to construct the objective function as:

$$\min \quad RC = \{c(a) - {y^\star}^T a | a \in \mathcal{A}\} \tag{2.5}$$

where the cost function is referred to as the Reduced Cost (RC) and y^T is the transposed of vector y. In the Pricing problem where a_j is the variable, the optimal solution is an ${c_j}^\star = c_j - {y^\star}^T a_j^\star$ where $j \in \mathcal{J}$ and the generated column $a_j^\star$ minimized the RC. This newly generated column is sent to the RM to improve its solution.

If the ${c_j}^\star \geq 0$, the obtained solution in the RM problem, i.e. $x^\star$ optimally solves the MP as well [26]. On the other hand, if ${c_j}^\star < 0$, the Pricing generates another column $a_j^\star$ and sends it to the RM sub-problem. This procedure keeps iterating until finding a non-negative ${c_j}^\star$.

2.2.4 Multiple Criteria Optimization

Assume we need to find an optimal solution for a problem with k objectives

$$
\begin{aligned}
&\min \quad F_1(x) \\
&\min \quad F_2(x) \\
&\vdots \\
&\min \quad F_k(x) \\
\text{subject to} \\
&x \in S
\end{aligned}
$$

where all $F_1(x) \cdots F_k(x)$ are the objective functions. There are two different approaches introduced to deal with multi-criteria optimizations, the weighted sum method and the ϵ-constraint method [34].

In the weighted sum method, the optimization model is modified to

$$
\begin{aligned}
&\sum_{j=1}^{k} \delta_j F_j(x) \\
\text{subject to} \\
&x \in S
\end{aligned}
$$

where δ_j is the weight for the j^{th} objective function while the summation of all δs has to be equal to 1. Different efficient solutions are obtained by varying δ_js; these weights are chosen based on the importance of each objective function.

However, in the ϵ-constraint method, only one of the criteria is kept as the objective function and others are transformed to new constraints shown as:

$$
\begin{aligned}
& \min \quad F_1(x) \\
\text{subject to} & \\
& F_2(x) \leq \epsilon_2 \\
\vdots & \\
& F_k(x) \leq \epsilon_k \\
& x \in S
\end{aligned}
$$

where $\epsilon_i \in R^k$. However, it is very important to choose a proper range for functions $F_2(x) \dots F_n(x)$ over the efficient set. The most common way to calculate these ranges is to use the payoff table; the table indicating the results of the individual optimization of each $k-1$ objective functions. In this way, the range of each of the objective functions is obtained and then is divided to narrower intervals after each run, until finding the optimal solution. The ϵ-constraint method has several superiorities over weighted sum solution, which are explained as follows [88]:

- Solving linear problems by weighting method for different selection of weights may be redundant in several runs. Because it is always solved over the original feasible region and may provide the same efficient extreme solution (the corner solution). However the ϵ-constraint method solves the problem over different feasible regions and produces non-extreme efficient solutions.

- In the sum weighted approach, scaling the objective functions has a strong effect on the results. Nevertheless, no scaling is necessary in the ϵ-constraint method.

- In the ϵ-constraint method, by properly adjusting the range intervals for the objective function, the number of the generated efficient solutions can be controlled. However, in the weighted sum method there is no control on the number of efficient solutions.

Due to the advantages of the ϵ-constraint method, we use this approach for solving our multi-criteria optimization problem, which is a column generation model.

2.3 Conclusion

In this chapter, after presenting the related literature to our research, we gave an overview of the optimization methods that we used in this thesis. We have used CG approach for solving the MILP problem in Chapters 4, 5 and 7. This approach is used in Chapters 5 for linearizing the non-linear objective function, as well. In Chapter 7, we have extended our cross-layer optimization model to a multiple criteria model where we used ϵ-constraint method to be compatible with CG approach for solving the problem.

Chapter 3

Cross-layer Optimization in Multihop Wireless Networks

In this chapter, we explain the essentials that has to be considered in a cross-layer optimization model in wireless networks. We introduce the popular schemes for modeling the interference in a wireless environment, followed by describing the scheduling constraints. We explain the routing problem accordingly. Next, we mention the objective functions which have been considered in the literature for the purpose of cross-layer optimization in wireless networks and clarify our objective function.

3.1 Interference Models

In this section, we describe two interference models widely popular in the literature. The first one which is based on the signal to interference plus noise ratio (SINR model), is a realistic scheme for modeling the interference. However, due to its complexity, the second model which is a simple one (Protocol model) has been introduced and used more vastly in the analysis of wireless networks, albeit it is less accurate than the SINR model.

3.1.1 The SINR Model

The Signal to Interference plus Noise Ratio (SINR) model is also known as additive interference model or physical model in the literature [47, 58, 110]. In this model, a receiving node treats the sum of all interference, which are undesired received signals from other on-going transmissions, as noise. Thus, the SINR at the receiver of link i is given by:

$$SINR_i = \frac{G_{t(i),r(i)}P_{t(i)}}{\eta + \sum\limits_{\forall n \in N_a - \{t(i)\}} G_{n,r(i)}P_n} \tag{3.1}$$

where $t(i)$ and $r(i)$ are the transmitter and receiver of link i, respectively; η is the background noise power in the frequency band of operation, N_a is the set of all active nodes in the network, P_n is the transmitting power at n, and $G_{n,n'}$ is the signal attenuation from node n to node n'. $G_{n,n'}$ is known as path loss and is modeled as:

$$G_{n,n'} = (\frac{d_{n,n'}}{d_0})^{-\beta} \tag{3.2}$$

where β is the path loss exponent and d_0 is a close-in distance to the transmitter, where the received power is measured as $P_{t(i)}$.

A link is feasible if the packet from its transmitter to its receiver is decodable. According to the SINR model, link i is feasible if the signal to noise ratio at $r(i)$ is above the SINR threshold of the receiver radio. The SINR threshold is the minimum required signal to noise ratio at the receiver which guarantees the tolerable Bit Error Rate (BER) of the link. Denote that higher SINR reduces the BER of the link. The SINR threshold, Γ, however depends on the acceptable Packet Error Rate (PER) which depends on the packet length and BER. In this case, the following inequality has to be satisfied for link i to be feasible:

$$\frac{G_{t(i),r(i)}P_{t(i)}}{\eta + \sum\limits_{\forall n \in N_a - \{t(i)\}} G_{n,r(i)}P_n} \geq \Gamma \tag{3.3}$$

3.1.2 The Protocol Model

Suppose we need to have a successful transmission on link i. In this model, all the nodes are assumed to communicate with the same transmission power. Let us define the transmission range of a node n by Tr_n. Thus, the interference range for each node n' is defined as:

$$Ir_{n'} = (1+\delta)\ Tr_n \qquad \forall n, n' \in N \tag{3.4}$$

where δ $(0 \leq \delta \leq 1)$ is the interference range coefficient. In this model, an active transmitter at a distance less than the interference range, is assumed to cause strong interference to the signal at the receiver. Thus, link i is feasible if the following conditions hold:

1. The receiver node, $r(i)$, is inside the transmission range of the transmitter node, $t(i)$.

   $$|r(i) - t(i)| \leq Tr_{t(i)}$$

   where $|x - y|$ denotes the Euclidean distance between nodes x and y.

2. Except $t(i)$, there is no active transmitter located inside the interference range of $r(i)$.

In this model, two links i and i' may be active conflict-free in any of the following cases:

- The transmitter of each is outside the interference range of the receiver of the other (shown in Figure 3.1), i.e.,

  $$\begin{cases} |r(i) - t(i')| \geq (1+\delta)\ Tr_{t(i)} \\ |r(i') - t(i)| \geq (1+\delta)\ Tr_{t(i')} \end{cases} \qquad \forall i, i' \in L$$

- Interference suppression techniques are employed to cancel the interference coming from unintended transmitter on any of recoveries $r(i)$ and $r(i')$.

Figure 3.1: Transmission range and interference range in the Protocol model.

Therefore, while i and i' in Figure 3.1 can be simultaneously active, node n cannot concurrently transmit since it interferes with them.

3.2 System Model and Parameters

We model a multihop wireless network as a directed graph $G = (N, L)$ where N is the set of all nodes equipped with multiple antennas and L is the set of all feasible transmission links in the network. We assume a set of M end-to-end (unicast) traffic sessions in this network. The m^{th} session is denoted by $S_m = \{(s_m, d_m, R_m) : s_m \in N, d_m \in N, R_m > 0, m = 1, ...M\}$. Therefore, in the m^{th} session, the source node s_m sends the commodity R_m to the destination node d_m.

3.3 Link Scheduling Constraints

Suppose a Time Division Multiple Access (TDMA) scheme has been employed for activating links in the network. In a TDMA-based MAC layer, time is divided into equal duration slots. At each time slot, a set of links can be active together without violating the requirement for successful communication, i.e. collision-free packet reception at nodes. The set of links that can be active concurrently in the same time slot is defined as a *configuration*, denoted by p. However, there are typically two types of limitations on link activations at each configuration p, known as *radio constraints* and *interference constraints* which are explained next.

3.3.1 Radio Constraints

These constraints essentially depend on the physical layer capabilities of the nodes. Noting that each node n in a basic multi-hop network has only one radio set, it can be active on a single link in each configuration p. Thus, the radio constraints are given as:

$$\sum_{i \in L_n} v_i^p \leq 1 \qquad \forall n \in N \tag{3.5}$$

where $L_n \subset L$ denotes the set of all links connected to node n, and the binary variable v_i^p is equal to 1 if link i is active in configuration p.

3.3.2 Interference Constraints

These constraints are generally established based on the following factors:

1. The physical-layer capabilities of nodes to deal with concurrent transmissions which are undesired to them and is treated as interference.

2. The scheme which is used to model the interference, i.e. either SINR or Protocol models.

In a simple wireless network where nodes have one simple radio, they are incapable of dealing with another ongoing transmission in their vicinity. Therefore, we describe the following interference constants for the SINR and the Protocol model accordingly.

SINR-based Interference Constraints: According to the explanation given in Section 3.1.1, link i is feasible if the overall signal to noise plus interference ratio at $r(i)$ is above the threshold Γ. Thus, the interference constraints are given as:

$$\frac{PG_{t(i),r(i)}}{\eta + \sum_{\forall n \in N - \{t(i)\}} PG_{n,r(l)} u_n^p} \geq \Gamma \qquad \forall i \in L \tag{3.6}$$

where u_n^p is defined as a binary variable with value 1, if node n is active in p and 0 otherwise.

Protocol-based Interference Constraints: For this scheme, let us define $\mathbf{C}$ as a set including all pairs of interfering links without any common nodes in between. This set is obtained from (3.4), which determines the links with their transmitters located in the interference range of the receiver of one link, i.e. interfering links to that. To guarantee that no pairs of interfering links are simultaneously active, we write:

$$v_i^p + v_j^p \leq 1 \qquad \forall (i,j) \in \mathbf{C} \tag{3.7}$$

3.4 Multihop Routing Constraints

In this section, we introduce two types of routing which have been used in the literature.

3.4.1 Link-based Routing

The multihop routing constraints are built based on two basic principles as follows.

1. **Flow-balance Constraints:** The flows of each session should be balanced at each node. Therefore, for each session m, whether a node is the source,

destination or just one intermediate hop, we preserve the flows by (3.8).

$$\sum_{i \in L_n^+} f_i^m - \sum_{j \in L_n^-} f_j^m = \begin{cases} 0 & \text{if } \; n \neq s_m, d_m \\ R_m & \text{if } \; n = s_m \\ -R_m & \text{if } \; n = d_m \end{cases} \qquad \forall n \in N, \forall m \in M \tag{3.8}$$

where f_i^m denotes the amount of flow from session m on link i; L_n^+ and L_n^- denote all outgoing links from, and incoming links to, node n, respectively; other parameters were defined earlier in Section 3.2.

2. **Bandwidth Constraints:** The flow of traffic on each link should not exceed the link capacity. Thus, one needs to guarantee that a link is active during enough time slots to be able to carry the amount of traffic flowing through it. Similar to [68], this can be mathematically presented as:

$$\sum_{m \in M} f_i^m \leq c_i \sum_{p \in \mathcal{P}} \lambda^p v_i^p \qquad \forall i \in L \tag{3.9}$$

where c_i is the time-invariant capacity of link i, λ^p is the number of time slots in which configuration p is active, $\mathcal{P}$ is the set of all configurations, and v_i^p was defined earlier as a binary variable with value 1 if link i is active during configuration p.

3.4.2 Path-based Routing

It has been shown in a previous work [35], when the number of variables is very large, the mathematical model becomes unscalable if we give all the links in the network the same chance to be involved for routing each of the sessions. Despite the wide usage of routing methods where sessions are liberally transformed through any link, i.e. link-based routing, they have shown that using limited predetermined paths for each of the sessions (so called *path-based* routing) is more scalable. Moreover, if the number of predefined paths for each session is large enough, path-based routing performs as good as linked-based. In this work, we have used the modified Dijkstra's [33] and Yen's [135] algorithms to find the shortest path and the k-shortest paths between each pair of source and destination for each session, respectively.

Flow-balance Constraints: We will develop our routing constraints according to the path-based method in Chapter 5, given the scale of the problem. Clearly, the larger k is, the better the performance becomes. We will later show the impact of value k on the performance of our model in the numerical results of Chapter 6. Using the path-based routing method, we modify the flow balance conservations given in equation (3.8) to:

$$\sum_{i \in L_n^+} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m - \sum_{j \in L_n^-} \sum_{\psi \in \Psi_m} \delta_\psi^j \omega_\psi^m = \begin{cases} 0 & \text{if } n \neq s_m, d_m \\ R_m & \text{if } n = s_m \\ -R_m & \text{if } n = d_m \end{cases} \tag{3.10}$$

where Ψ_m denotes the set of predetermined k-shortest paths for session m, ω_ψ^m is a variable that shows the amount of session m on path ψ and δ_ψ^i is defined as a binary parameter with value 1 for any link i which is traversed by path ψ. Equations (3.10) preserve the flow balance at all the nodes in the network.

Bandwidth Constraints: Based on k-shortest path routing, several packets of one session at the source are divided into k different groups where each of them, ω_ψ^m, is routed through an individual path ψ toward the destination. In this case, the amount of traffic flowing on link i is equal to $\sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m$. Therefore, the bandwidth constraints given in (3.9), has to be modified to (3.11).

$$\sum_{p \in \mathcal{P}} v_i^p c_i \lambda^p \geq \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m, \qquad \forall i \in L \tag{3.11}$$

3.5 Objective Function

There have been a number of objective functions in optimizing the performance of wireless networks proposed in the literature. These objectives can be summarized as:

1. A utility function used for maximizing:

   - Sum of the flows on the links in the network [122]:

     $\max \sum_{m \in M, i \in L} f_i^m$

   - Sum of the weighted flows on the links [61]:

     $\max \sum_{m \in M, i \in L} w(f_i^m) f_i^m$

   - Sum of the traffic sessions in the network [51]:

     $\max \sum_{m \in M} R_m$

2. A utility function employed to maximize the minimum traffic session in the network [83, 107, 108] which also provides fairness:

   $\max \min_{\forall m \in M} R_m$

3. A cost function defined to minimize the total scheduling time while the network is responsible to deliver a certain amount of traffic sessions [35]:

   $\min \sum_{p \in \mathcal{P}} \lambda^p$

Throughout this thesis, our objective will be to improve the network performance by minimizing the total system activation time or the scheduling period, which is the third method. We modify this objective function, however, in Chapter 5. In contrast to the two first methods where the amount of traffic for each session m, i.e. R_m, is defined as a variable, in our method R_m is predefined.

3.6 Conclusion

In this chapter, we explained the basic principles of joint routing and scheduling optimization in multihop wireless networks. We described the required constraints and the objective functions for mathematically modeling the problem and will extend these models in the next chapters for the cases of employing advanced techniques at physical and network layers.

Chapter 4

Scheduling with Successive Interference Cancellation

In this chapter, first we describe the Successive Interference Cancellation (SIC) technique. Then, we introduce a decomposed model for joint routing and scheduling in wireless multihop networks when SIC is employed. The numerical results for this model is provided in Chapter 6.

4.1 Successive Interference Cancellation

Interference cancellation is fundamentally the idea of removing some parts of the interference from the aggregate received signal to improve the SINR. In other words, with interference cancellation, some terms in the summation of denominator in equation (3.1) are cancelled out and the value of SINR is increased. The interference can conveniently be detected and removed due to its data-like structure. Different methods for interference cancelation have been employed in the literature; namely, parallel [119], successive [49], combination of successive and parallel [118] and iterative [3]. In this thesis, we are focusing on Successive Interference Cancelation (SIC) which is shown to be the best scheme for the case of unequal power reception [97].

Figure 4.1: SIC process [61]. Assume $SINR_{1n} > SINR_{2n} > ... > SINR_{k_n n}$. Corresponding nodes are shown in Figure 4.2.

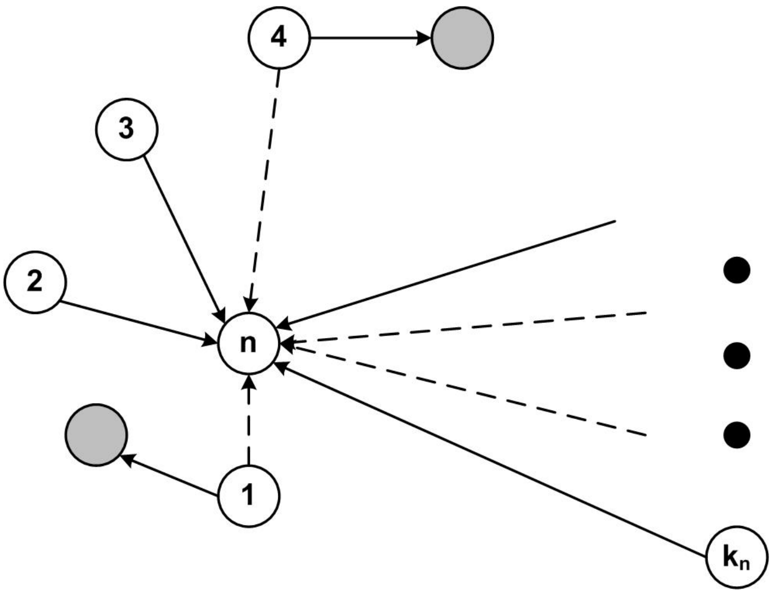


Figure 4.2: Node indexing in SIC.

As shown in Figure 4.1, the packet with the highest power is decoded in the first step of SIC process. Since we have considered equal transmission power for all nodes, the signal with the highest power is received from the closest active node. As illustrated in Figure 4.1, suppose concurrent transmitting nodes are indexed from 1 to k_n based on their distance from a node n, where $d_{1,n} < d_{2,n} < ... < d_{k_n,n}$. Denote k_n as the maximum number of decodable packets at this node. Thus, the first packet is decodable if equation (3.3) holds as:

$$\frac{G_{1,n} P_1}{\eta + \sum_{\forall n' \in N_a - \{1\}} G_{n',n} P_{n'}} \geq \Gamma$$

where P_1 is the transmitting power of node 1. To extract the second strongest packet, however, we need to determine and subtract the interference caused by the first packet from the aggregate signal. Therefore, we reconstruct a part of the received signal which is related to the first packet. This reconstruction is done by using the decoded packet and an estimation of the path loss of the corresponding link. In this work, we assume that packet decoding and reconstruction of the received signals are error free. After removing the interference caused by the first packet, the receiver checks if the threshold requirement for the second packet is met:

$$\frac{G_{2,n} P_2}{\eta + \sum_{\forall n' \in N_a - \{1,2\}} G_{n',n} P_{n'}} \geq \Gamma$$

The receiver continues decoding up to k_n packets; however, it stops whenever the SINR threshold is no longer satisfied. From Figure 4.2 we can observe two main advantages for SIC. The first one is the capability of receiving multiple desired packets from concurrent intended transmitters: $2, 3, ...k_n$. The second one is the capability of rejecting the interference from unintended transmitters: $1, 4, ...$. For instance, receiver n is strongly interfered by transmitter 1. However, by employing SIC this interference can be cancelled and n is still able to receive its desired packets, even with lower signal power. While the first benefit allows for transfusing more data to a node, the second one increases the spatial reuse for other transmissions. These two advantages of SIC indeed improve the throughput significantly. Although decoding more packets and cancelling more interference provides higher capacity to the system, we should note that it causes more latency due to the successive nature of the process [5]. In [61], an upper bound for the number of successful decoding at node n is obtained based on physical parameters of the system:

$$k_n \leq 1 + \log_{\Gamma+1}(\frac{PG_{n',n}}{\eta\, \Gamma}) \qquad n' \text{ is the closest node to } n \tag{4.1}$$

Denote that k_n includes the number of intended and unintended packets.

4.2 System Model

In this work, a multi-hop wireless network is modeled using a directed graph $G = (N, L)$, where N is the set of all nodes and L is the set of all transmission links in the network. Considering the same transmitting power, modulation and coding scheme at all nodes, the capacity of all the links, named c, is the same which is normalized to unity. We assume a set of M end-to-end (unicast) traffic sessions in this network. The m^{th} session is denoted by $S_m = \{(s_m, d_m, R_m) : s_m \in N, d_m \in N, R_m > 0, m = 1, ...M\}$. Therefore, in the m^{th} session, the source node s_m sends the commodity R_m to the destination node d_m. We assume the SINR model [47] is used for modeling the interference in this network which we explained it in Section 3.1.1.

4.3 Joint Routing and Scheduling with SIC

In this section, We provide the optimization constraints for the cross-layer design of wireless multihop networks according to the basics we described in Chapter 3.

4.3.1 Scheduling

As we discussed in Section 3.3, two sets of constraints have to be considered for the scheduling which are introduced in the following parts.

Interference Constraints

We assume a Time Division Multiple Access (TDMA) system. In a TDMA-based MAC layer, time is divided into equal duration slots. At each time slot, a set of links can be active together without violating the requirement for successful communication, i.e. collision-free packet reception at nodes. The set of links that can be active concurrently in the same time slot is defined as a configuration, denoted by p. A binary variable that indicates whether a node is transmitting in p or not is defined as:

$$u_n^p = \begin{cases} 1 & \text{if node } n \text{ is transmitting in } p \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the SINR requirement, or interference constraint for a link l can be written as:

$$\begin{aligned} &\frac{G_{t(i),r(i)}P}{\eta + I_1 - I_2} \geq \Gamma \\ &I_1 = \sum_{\forall n \in N - \{t(i)\}} G_{n,r(i)} P u_n^p \\ &I_2 = \sum_{\forall n' \in N - \{t(i)\} : G_{t(i),r(i)} \leq G_{n',r(i)}} G_{n',r(i)} P u_{n'}^p \end{aligned} \tag{4.2}$$

where we use P to denote the transmission power at any transmitting node, I_1 is the total interference received at $r(l)$ and I_2 is the cancelled inference. As I_2 shows, the interference from transmitters closer than $t(l)$ are cancelled when the packet of $t(l)$ is getting decoded. Equation (4.2) is then simplified and written as:

$$\frac{PG_{t(i),r(i)}}{\eta + \sum_{\forall n \in N - \{t(i)\} : G_{t(i),r(i)} \geq G_{n,r(i)}} PG_{n,r(i)} u_n^p} \geq \Gamma \tag{4.3}$$

We define binary the variable v_i^p to indicate whether link i is active in p or not as:

$$v_i^p = \begin{cases} 1 & \text{if link } i \text{ is active in } p \\ 0 & \text{otherwise.} \end{cases}$$

In this case, if equation (4.13) holds, then v_i^p can be equal to 1; and if it is not satisfied, then $v_i^p = 0$. To reflect this statement in a mathematical programming format, we modify equation (4.13) as:

$$PG_{t(i),r(i)} + M_i^p(1 - v_i^p) \geq \Gamma(\eta + \sum_{n \neq t(i), G_{t(i),r(i)} \geq G_{n,r(i)}} PG_{n,r(i)} u_n^p) \qquad (4.4)$$

where M_i^p is a constant parameter satisfying:

$$M_i^p \geq \Gamma(\eta + \sum_{n \neq t(i), G_{t(i),r(i)} \geq G_{n,r(i)}} PG_{n,r(i)}) \qquad (4.5)$$

We can see that whenever $v_i^p = 1$, (4.4) reduces to (4.13), and whenever $v_i^p = 0$, (4.4) is still satisfied because of (4.5).

Radio Constraints

To achieve a network-wide link scheduling, free of conflicts, we have to consider radio constraints as well. Let L_n be a subset of L that includes all the links connected to node n. $L_n^+ = \{i \in L : t(i) = n\}$ is the set of all links whose transmitter is node n and $L_n^- = \{i \in L : r(i) = n\}$ is the set of all links whose receiver is node n. Therefore, $L_n = \{L_n^+ \cup L_n^-\}$. The radio constraints which include half-duplex properties and limitation of receiving k_n packet at node n must be satisfied for the proper link scheduling in the network and are given in equations (4.27) and (4.7), respectively.

$$v_i^p + v_j^p \leq 1 \qquad \forall n \in N : i \in L_n^-,\ j \in L_n^+ \qquad (4.6)$$

$$\sum_{i \in L_n^-} v_i^p \leq k_n \qquad \forall n \in N \qquad (4.7)$$

We note that a node can only transmit to a single receiver in a general scheme, (i.e., no multi-packet transmission capability) which is taken care by equation (4.28).

$$\sum_{i \in L_n^+} v_i^p \leq u_n^p \qquad \forall n \in N \tag{4.8}$$

4.3.2 Routing

The two sets of constraints required in the routing in this problem are addressed next.

Flow-balance Constraints

Although routing the flows in a MWN with SIC properties is possible through the links, we use the path-based approach which was described in Section 3.4.

$$\sum_{i \in L_n^+} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m - \sum_{j \in L_n^-} \sum_{\psi \in \Psi_m} \delta_\psi^j \omega_\psi^m = \begin{cases} 0 & \text{if } \ n \neq s_m, d_m \\ R_m & \text{if } \ n = s_m \\ -R_m & \text{if } \ n = d_m \end{cases} \tag{4.9}$$

where Ψ_m denotes the set of predetermined k-shortest paths for session m, ω_ψ^m is a routing variable that shows the amount of traffic of session m on path ψ and δ_ψ^i is defined as a binary parameter with value 1 for any link i which is traversed by path ψ. Equations (4.9) preserve the flow balance at all the nodes in the network.

Bandwidth Constraints

Based on k-shortest path routing, several packets of one session at the source are divided into k different groups where each of them, ω_ψ^m, is routed through an individual path ψ toward the destination. In this case, the amount of traffic flowing on link i is equal to $\sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m$. Therefore, the bandwidth constraints given in 3.9, has to be modified to (3.11).

$$\sum_{p \in \mathcal{P}} v_i^p c_i \lambda^p \geq \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m, \qquad \forall i \in L \tag{4.10}$$

4.3.3 Problem Formulation

As explained in Chapter 3, our objective is to maximize the throughput of the network by minimizing the overall scheduling time, which is represented as:

$$\min \sum_{p=1}^{\bar{P}} \lambda_p \tag{4.11}$$

where $\bar{P}$ is the total number of configurations. This minimization is done through a joint routing and scheduling model, considering all the radio, interference, flow conservation and bandwidth constraints given in the previous sections. In this problem, all the configurations in $\mathcal{P}$ should be generated and the routing problem should be solved over all of them to yield the optimal solution. However, by using CG decomposition approach, the optimal solution is obtained without enumerating all configurations. Thus, the problem is decomposed into smaller subproblems which are the Restricted Master (RM) problem and the Pricing problem. More details in solving the problem by CG technique are explained in the following paragraphs.

Master and Pricing sub-problems

The routing of demands is determined in the RM problem. Figure 4.3 illustrates the flowchart of the CG algorithm. The objective of the RM is the same as the objective of the original problem and the constraints are flow conservations (4.9) and bandwidth constraints (4.10). In the first iteration, a set of initial feasible configurations, $\mathcal{P}_0 \subseteq \mathcal{P}$ is available to solve the first instance of the RM. It should be noted that v_i^p is not a variable in the RM. The solution of the Restricted Master is the best routing over all possible configurations ($\mathcal{P}_0$), which is a local optimum. When the RM obtains the local optimal solution, it generates the dual values of the bandwidth constraints, as $\{\bar{y}_i,\ i \in L\}$ and sends them to the Pricing to find the best link scheduling. Then the Pricing sub-problem applies these dual values to

Figure 4.3: The flowchart of Column Generation method

construct the Pricing objective as:

$$\min(1 - \sum_{i \in L} \bar{y}_i v_i^p) \tag{4.12}$$

where $(1 - \sum_{i \in L} \bar{y}_i v_i^p)$ is called the Reduced Cost (RC). The constraints of the Pricing are the radio and interference constraints (4.13)-(4.28).

From now on, in each iteration, a feasible configuration (known as a column) is generated in the Pricing sub-problem. The value of the Pricing objective is always checked to determine the optimality of the solution. If (4.12) is a non-negative value, the obtained solution in the RM is the optimal solution to the main problem. In this case, generating more columns in the Pricing would not provide further improvement because the optimal solution has already been found. A proof of optimality of the solution is given in [26]. On the other hand, if (4.12) has a negative value, the Pricing generates another configuration named p and sends it to the RM problem. Then, the RM adds p into the set of previous configurations, $\mathcal{P}_0 \biguplus p \rightarrow \mathcal{P}_0$, and resolves the routing problem. Whenever a solution is produced by the Restricted Master model, the corresponding dual values are passed to the Pricing and the procedure continues until finding the optimal solution.

CG-based MPR Model

Our decomposed model based on CG approach is given in Table 4.1

Table 4.1: Column Generation model of SIC

[RM problem] :

$$\min \sum_{p \in \mathcal{P}_0} \lambda_p$$

subject to

$$\sum_{i \in L_n^+} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m - \sum_{j \in L_n^-} \sum_{\psi \in \Psi_m} \delta_\psi^j \omega_\psi^m = \begin{cases} 0 & \text{if } n \neq s_m, d_m \\ R_m & \text{if } n = s_m \\ -R_m & \text{if } n = d_m \end{cases}$$

$$\sum_{p \in \mathcal{P}} v_i^p c_i \lambda^p \geq \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m$$

[Pricing problem] :

$$\min(1 - \sum_{i \in L} \bar{y}_i v_i^p)$$

subject to

$$PG_{t(i),r(i)} + M_i(1 - v_i^p) \geq \Gamma(\eta + \sum_{n \neq t(i), G_{t(i),r(i)} \geq G_{n,r(i)}} G_{n,r(i)} u_n^p P)$$

$$M_i \geq \Gamma(\eta + \sum_{n \neq t(i), G_{t(i),r(i)} \geq G_{n,r(i)}} G_{n,r(i)} P)$$

$$v_i^p + v_j^p \leq 1 \qquad \forall n \in N : i \in L_n^-, \ j \in L_n^+$$

$$\sum_{i \in L_n^-} v_i^p \leq k_n \qquad \forall n \in N$$

$$\sum_{i \in L_n^+} v_i^p \leq u_n^p \qquad \forall n \in N$$

4.4 Joint Routing, Scheduling and Power Control

We assume here that nodes can adjust their power of transmission; therefore, in the case of employing SIC, the interference constraint for a feasible link i which was given as:

$$SINR_i = \frac{PG_{t(i),r(i)}}{\eta + \sum\limits_{\forall n \in N - \{t(i)\} :\ G_{t(i),r(i)} \geq\ G_{n,r(i)}} PG_{n,r(i)} u_n^p} \geq \Gamma \tag{4.13}$$

for the case of constant transmission power will be changed to:

$$SINR_i = \frac{P_{t(i)}^p G_{t(i),r(i)}}{\eta + \sum\limits_{\forall n \in N - \{t(i)\} : P_{t(i)}^p\ G_{t(i),r(i)} \geq P_n^p\ G_{n,r(i)}} P_n^p G_{n,r(i)} u_n^p} \geq \Gamma \tag{4.14}$$

where $P_{t(i)}^p$ and P_n^p are the transmission power of nodes $t(i)$ and n, respectively. This constraint is written in an LP format as follows:

$$P_{t(i)}^p\ G_{t(i),r(i)} + M_i^p\ (1 - v_i^p) \geq \Gamma\ (\eta + \sum_{n \neq t(i), P_{t(i)}^p\ G_{t(i),r(i)} \geq P_n^p\ G_{n,r(i)}} P_n^p\ G_{n,r(i)}) \tag{4.15}$$

The maximum transmission power at nodes is denoted by P_{max} and the power allocated to an inactive node n must be zero. Therefore, we can write:

$$P_n^p \leq u_n^p P_{max} \tag{4.16}$$

We also consider a P_{min} for each node through which it can connect to the closest node in the absence of any interference. Therefore:

$$P_n^p \geq u_n^p P_{min} \tag{4.17}$$

However, the constraints in (4.15) are not represented in an LP format due to the condition $P_{t(i)}^p\ G_{t(i),r(i)} \geq P_n^p\ G_{n,r(i)}$ which is in the summation. Since this condition includes some variables (i.e. $P_{t(i)}^p$ and P_n^p), it is not implementable in LP format. However, we overcome this obstacle as will be explained next.

4.4.1 Linear Interference Constraints

Let us define a new binary variable:

$$e^p_{n,i} = \begin{cases} 1 \quad \text{if} \quad P^p_{t(i)}\ G_{t(i),r(i)} \geq P^p_n\ G_{n,r(i)} \\ 0 \quad \text{if} \quad P^p_{t(i)}\ G_{t(i),r(i)} < P^p_n\ G_{n,r(i)} \end{cases} \qquad \forall i \in L, \forall n \in N - \{t(i), r(i)\}$$

which can be obtained through the following constraints:

$$P^p_{t(i)} G_{t(i),r(i)} - P^p_n G_{n,r(i)} + (1 - e^p_{n,i}) M^p_{n,i} \geq 0 \tag{4.18}$$

$$P^p_{t(i)} G_{t(i),r(i)} - P^p_n G_{n,r(i)} - e^p_{n,i} M^p_{n,i} < 0 \tag{4.19}$$

$$M^p_{n,i} \geq |P^p_{t(i)} G_{t(i),r(i)} - P^p_n G_{n,r(i)}| \tag{4.20}$$

where $M^p_{n,i}$ is an integer parameter. Therefore, the interference constraints (4.15) modifies to:

$$P^p_{t(i)}\ G_{t(i),r(i)} + M^p_i\ (1 - v^p_i) \geq \Gamma\ (\eta + \sum_{n \neq t(i),r(i)} e^p_{n,i}\ P^p_n\ G_{n,r(i)}) \tag{4.21}$$

which is still non-linear because of the term $e^p_{n,i} P^p_n$. Let us resolve this non-linearity by defining a new variable $h^p_{n,i}$ where

$$h^p_{n,i} = e^p_{n,i} P^p_n \tag{4.22}$$

However, this equality is implicitly achievable through the following LP constraints:

$$h^p_{n,i} \leq P^p_n + (1 - e^p_{n,i}) P_{max} \tag{4.23}$$

$$h^p_{n,i} \geq P^p_n - (1 - e^p_{n,i}) P_{max} \tag{4.24}$$

$$h^p_{n,i} \leq e^p_{n,i} P_{max} \tag{4.25}$$

where if $e^p_{n,i} = 0$, then $h^p_{n,i}$ becomes 0; and whenever $e^p_{n,i} = 1$, then $h^p_{n,i} = P^p_n$. Using $h^p_{n,i}$ in (4.21), the linear interference constraints are given by:

$$P^p_{t(i)}\ G_{t(i),r(i)} + M^p_i\ (1 - v^p_i) \geq \Gamma\ (\eta + \sum_{n \neq t(i),r(i)} h^p_{n,i}\ G_{n,r(i)}) \tag{4.26}$$

4.4.2 Radio Constraints

The radio constraints for the case of SIC with power control are similar to the case of constant power which are given in (4.27)-(4.28).

$$v_i^p + v_j^p \leq 1 \qquad \forall n \in N : i \in L_n^-,\ j \in L_n^+ \tag{4.27}$$

$$\sum_{i \in L_n^+} v_i^p \leq u_n^p \qquad \forall n \in N \tag{4.28}$$

Table 4.2: Column Generation model of SIC with Power Control

[RM problem] :

$$\min \sum_{p \in \mathcal{P}_0} \lambda_p$$

subject to

$$\sum_{i \in L_n^+} f_i^m - \sum_{j \in L_n^-} f_j^m = \begin{cases} 0 & \text{if } n \neq s_m, d_m \\ R_m & \text{if } n = s_m \\ -R_m & \text{if } n = d_m \end{cases}$$

$$\sum_{m \in M} f_i^m \leq c_i \sum_{p \in \mathcal{P}} \lambda^p v_i^p$$

[Pricing problem] :

$$\min(1 - \sum_{i \in L} \bar{y}_i v_i^p)$$

subject to

$$P_{t(i)}^p\ G_{t(i),r(i)} + M_i^p\ (1 - v_i^p) \geq \Gamma\ (\eta + \sum_{n \neq t(i),r(i)} h_{n,i}^p\ G_{n,r(i)})$$

$$P_n^p \leq u_n^p P_{max}$$

$$P_n^p \geq u_n^p P_{min}$$

$$P_{t(i)}^p G_{t(i),r(i)} - P_n^p G_{n,r(i)} + (1 - e_{n,i}^p) M_{n,i}^p \geq 0$$

$$P_{t(i)}^p G_{t(i),r(i)} - P_n^p G_{n,r(i)} - e_{n,i}^p M_{n,i}^p < 0$$

$$h_{n,i}^p \leq P_n^p + (1 - e_{n,i}^p) P_{max}$$

$$h_{n,i}^p \geq P_n^p - (1 - e_{n,i}^p) P_{max}$$

$$h_{n,i}^p \leq e_{n,i}^p P_{max}$$

$$v_i^p + v_j^p \leq 1$$

$$\sum_{i \in L_n^+} v_i^p \leq u_n^p$$

$$M_i^p, M_{n,i}^p = 1000$$

4.4.3 CG-based model for SIC with Power Control

Table 4.2 shows the decomposed problem of joint routing and scheduling when SIC is used with variable transmission power at nodes. In this model we have used the link-based routing constraints in the RM problem which was explained in Section 3.4.1.

4.4.4 CG-based model for Power Control

In Table 4.3, we have provided the the decomposed problem of joint routing and scheduling when only transmission power control (and no SIC) is employed at nodes. We will later use this model to discuss the benefits of the interefrence avoidance acheived by power control technique and interfernce exploition provided by SIC technique.

4.5 Conclusion

In this chapter, we provided a cross-layer optimization scheme for MWNs where nodes are capable of cancelling the inference through successive interference cancellation method. We extended our model to a case where transmission power is adjustable to reduce the wireless interference. We made our models tractable by decomposing them with column generation approach. We defer the presentation of the numerical results of this model to Chapter 6.

Table 4.3: Column Generation model of Power Control (PC)

[RM problem] :

$$\min \sum_{p \in \mathcal{P}_0} \lambda_p$$

subject to

$$\sum_{i \in L_n^+} f_i^m - \sum_{j \in L_n^-} f_j^m = \begin{cases} 0 & \text{if } \; n \neq s_m, d_m \\ R_m & \text{if } \; n = s_m \\ -R_m & \text{if } \; n = d_m \end{cases}$$

$$\sum_{m \in M} f_i^m \leq c_i \sum_{p \in \mathcal{P}} \lambda^p v_i^p$$

[Pricing problem] :

$$\min(1 - \sum_{i \in L} \bar{y}_i v_i^p)$$

subject to

$$PG_{t(i),r(i)} + M_i^p (1 - v_i^p) \geq \Gamma(\eta + \sum_{n \neq t(i)} G_{n,r(i)} P_n^p)$$

$$P_n^p \leq u_n^p P_{max}$$

$$P_n^p \geq u_n^p P_{min}$$

$$v_i^p + v_j^p \leq 1$$

$$\sum_{i \in L_n^+} v_i^p \leq u_n^p$$

$$M_i^p = 1000$$

Chapter 5

Optimal Routing and Scheduling with Network Coding and SIC

We start this chapter by providing general information about the the simple wireless network coding scheme COPE introduced in [66] and we describe various structures which can be used as network coding components. We analyze the benefits and the gain that each coding component provides. Then, we introduce our network coding aware routing scheme by using some illustrative examples and then formulating the routing constraints. After than we discuss the problem of TDMA-based scheduling scheme when network coding with opportunistic listening is employed and we introduce and formulate our novel scheduling method. Finally, we formulate the cross-layer optimization and provide the decomposed model. The presentation od the numerical results is however deferred to Chapter 6.

5.1 Network Coding Model and Components

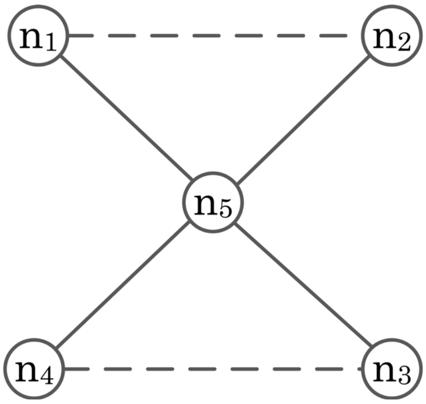
The fundamental idea of network coding is that a relay node combines several packets, intended for various receivers, into one packet and broadcasts it. Providing that each recipient has a priori knowledge about other packets, it can decode the desired packet from the aggregate packet. Therefore, the relay node is capable of forwarding more data within one transmission which eventually improves the overall throughput.

In this work, we focus on COPE scheme proposed in [66], where packets are linearly coded through a simple XOR operation. We assume that each coded packet is decodable at the next hop of a broadcast transmission. Therefore, several network coding topologies, which we call *coding components*, can be formed in the network (see Figure 5.1). In a coding component, an *edge node* is a transmitter and/or receiver of different traffic packets. A recipient edge node must know any uncoded packet except its desired one through either: 1) overhearing the link through which the packet was transmitted, or 2) being its previous transmitter. Coding components constructed based on the former scheme of obtaining knowledge, are referred to as *network coding with opportunistic listening*, while the latter one is known as network coding without opportunistic listening in the literature [66]. In any particular coding component, the node responsible for combining native received packets from other nodes in the same coding component (edge nodes), is referred to as the *relay node.* We denote that in each coding component, only packets from different sessions can be encoded together. Figure 5.1 illustrates various coding components, constructed based on the above rules, which we explain as follows:

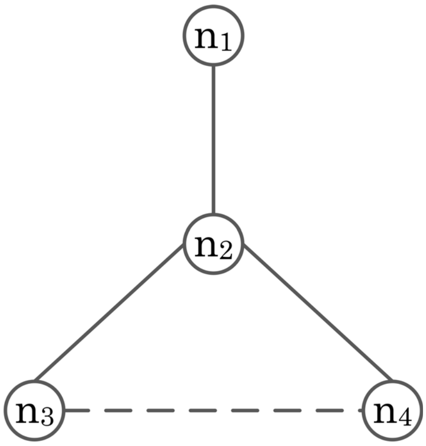
Chain component

A maximum of two packets from two sessions traversing in reverse directions are coded without opportunistic listening. For instance, relay node n_2 upon receiving packets from n_1 and n_3 (in two consecutive time slots) performs XOR operation and broadcasts the coded packet for both nodes n_1 and n_3 in the third time slot. These two edge nodes subsequently can decode the coded packet by XORing it with their own native one to extract their desired packet. This component reduces the required

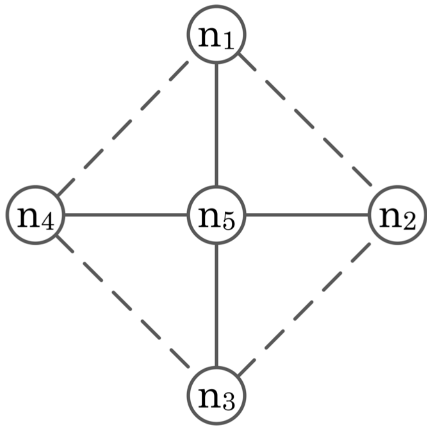
(a) Chain component.



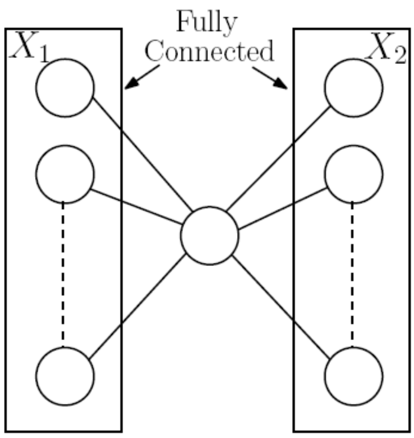
(b) X component.



(c) Bell component.



(d) Cross component.



(e) extended X component [27].

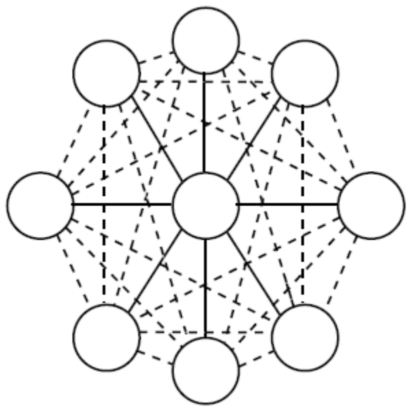

(f) extended Cross component [27].

Figure 5.1: Different coding components. Solid lines show intended transmission links and dashed lines show overhearing links.

packet delivery from 4 to 3 time slots which is a 25% improvement [66].

X component

A maximum of two packets which are sent in two consecutive time slots and intersecting at the relay node, are encoded. The destination of each packet obtains the other unintended native packet by listening to its transmission (opportunistic listening). Later, the overheard packet is used to decode the intended packet. The performance of X component is similar to the Chain and provides 25% improvement by reducing required packet delivery time from 4 to 3.

Bell component

A maximum of two packets are encoded where only one of the destinations obtains its unintended packet through opportunistic listening. This component improves the delivery time by 25%, because we need 3 slots instead of 4, similar to the Chain and X components.

Cross component

A maximum of 4 packets are coded through opportunistic listening, similar to the approach for other components. We observe that the X component reduces the delivery time from 8 to 5 which is a 37.5% gain.

The structure of a Partial Cross component is similar to the Cross component, however one or two of the flow sessions are zero; i.e. only three or two packets are involved. In the former case, the three native packets are coded and broadcast in one transmission. Since instead of 6 slots for transmission (3 for the edge and 3 for the relay nodes), 4 time slots (3 for the edge and 1 for the relay nodes) are used, the delivery time improvement is 33.33%. However, the latter case with two flows, is essentially a Chain or X component depending on the absent flows.

Extended X component

Here, all nodes in one group of X_1 or X_2 are fully connected beside being connected to the relay. Thus, they can listen to each other's transmission. Assume there are k_1 packets to be sent from nodes in X_1 to nodes in X_2, and k_2 packets to be sent from group X_2 to group X_1. Without loss of generality, we assume $k_1 \leq k_2$. Therefore, there is a maximum of k_1 coded packets where each one contains two native packets. The remaining $k_2 - k_1$ packets have to be transmitted as uncoded or native. In this case, the total required time to complete the packets deliveries is equal to $2k_2 + k_1$ slots. Because we need $k_2 + k_1$ slots for unicast transmissions from edge nodes to the relay, k_1 slots for broadcasting coded packets and $k_2 - k_1$ slots for unicast transmissions from the relay to the intended edge nodes. On the other hand, in the case of no NC the required time is equal to $2k_1 + 2k_2$. Therefore, an improvement of $\frac{k_1}{2k_1+2k_2} \times 100\%$ is achieved. We discuss the gain for different values of k_1 and k_2 as follows:

- $k_1 = k_2 = k$**:** The improvement is $\frac{(k)\times 100}{4k}\%$ which is 25%. Denote that Chain, X and Bell components are especial cases of extended X component with $k_1 = k_2 = k = 1$.

- $k_1 \ll k_2$**:** The asymptotic gain is obtained as $\frac{100}{2k_2}\%$ due to the negligible k_1 comparing to k_2. This gain is very small for a large k_2, because the benefits of NC appears when the number of intersecting flows from each direction at the relay node are balanced.

The above discussion clarifies that the extended X component improves the delivery time by 25%, at most.

Extended Cross component

Suppose there are k edge nodes in this component, where k has to be an even number, greater than 4. Each of the edge nodes is connected to all the others except the one located at the furthest opposite direction. Assuming each of the edge nodes has packets to transmit, the required delivery time without NC is $2k$ and reduces to $k + 1$ when NC is used. We encode all the packets in one at the relay node. Therefore, the gain for this component is $\frac{k-1}{2k} \times 100\%$. If $k = 4$, we have the ordinary Cross component with gain of 37.5%. We understand that this component can be interpreted as an extended X component if: 1) $k_1 = k_2 = k/2$, and 2) different groups of X_1 and X_2 are defined for each edge node; all other edge nodes are in one group, X_1 for instance, and only the node at the furthest opposite direction is in the other group, X_2.

It is denoted that coding components in the format of Chain are more applicable in regular networks, however X and Cross components and their extensions are more useful in more dense networks with bottleneck and congestion problems [27]

Table 5.1: Delivery time, number of involved packets and coding component gains for spatial reuse, NC and NC+SIC cases

Coding Component	Number of Native Packets in		Time Slots			Gain	
	Total	One Coded Packet	BM	NC	NC+SIC	$\frac{NC}{BM}\%$	$\frac{NC+SIC}{BM}\%$
Chain	2	2	4	3	2	25%	50%
X	2	2	4	3	2	25%	50%
Bell	2	2	4	3	2	25%	50%
Partial Cross	3	3	6	4	3	33.3%	50%
Cross	4	4	8	5	3	37.5%	62.5%
Extended X	$k_1 + k_2$	2	$2k_1 + 2k_2$	$2k_2 + k_1$	$2k_2$	$\frac{(k_1)}{2k_1+2k_2} \times 100$	$\frac{(k_1)}{k_1+k_2} \times 100$
Extended Cross	k	k	$2k$	$k + 1$	$\frac{k}{2} + 1$	$\frac{(k-1)}{2k} \times 100$	$\frac{(1.5k-1)}{2k} \times 100$

5.2 Network Coding Incorporated with Successive Interference Cancellation

Consider the coding components in Figure 5.1 where all the nodes are endowed with SIC technique. Since SIC provides multi-packet reception, some of the edge nodes in the coding component can transmit their packets concurrently. Therefore, whether the receiver node is the relay or another edge node in listening mode, it can decode several intended packets and cancel the interference of unintended transmissions. For instance, in the Chain component (Figure 5.1(a)), the total required time reduces to 2, since nodes j and k transmit their native packets concurrently. Similarly, in the Cross and Bell components (Figures 5.1(b) and 5.1(c)), nodes n_1, n_2; and j, j' can transmit simultaneously to save one time slot.

In the Cross component (Figure 5.1(d)), however, each pair of unconnected edge nodes transmit in the same time slot while the other two nodes are listening. Therefore, we can save in total 2 time slots.

Alternatively, in the extended X component, instead of $2k_1$, only k_1 slots are needed for the relay to receive $2k_1$ native packets from both groups. For the case of extended Cross component, the required time for unicasting native packets from the edge nodes to the relay reduces from k to $k/2$. In Table 5.1, we have summarized the required delivery time and the gains in each case of: NC combined with SIC, NC, and the Base Model (BM), where no NC or SIC is applied.

Based on the gains expressed in this table, we observe that SIC incorporated with NC results in a double improvement for the Chain, Bell, X and the extended X components. SIC combined with NC shows 1.67 times improvement in the Cross component which reduces to 1.5 for large values of k in the extended Cross component. All of these results show the great significance of combining SIC with NC.

We denote that in none of these components (Figure 5.1), there are more than

2 concurrent transmissions; further, the capability of receiving more than 2 packets adds no more benefits. However, the capability of receiving more than 2 packets in larger networks with more than one unit of coding component, is intuitively promising. Because different coding components can be active concurrently without causing interference on each other (or the interference can be suppressed).

One of our main goals in this research is to investigate the impact of SIC on the throughput improvements in larger networks where several adjacent coding components are formed. We elaborate this subject is Section 5.5 where scheduling for NC combined with SIC is explained. In the following section, we illustrate various examples to show our method for selecting a coding component to route all the sessions. We illustrate how a relay node encodes packets from different sessions by using more than one coding component. Our other main objective in this work is to introduce a routing formulation which provides the best choice of coding components shown in Figure 5.1. We are interested in forming several of these components at a specific node or combining them at several neighboring nodes in larger networks to minimize the overall delivery time in the system.

Illustrative Example

Consider the following traffic cases in a Cross component network topology shown in Figure 5.1(d). Let $n_i \xrightarrow{R_k} n_j$, denote a traffic session containing R_k packets, which is from node n_i to node n_j. In the following examples, our objective is to employ coding component(s) which minimizes the overall scheduling time. We denote that the potential coding components are known at the relay node.

- **Example 1**
  $n_1 \xrightarrow{1} n_3, \quad n_3 \xrightarrow{1} n_1, \quad n_2 \xrightarrow{1} n_4, \quad n_4 \xrightarrow{1} n_2$**:**
  Since we seek for the optimal coding component(s) at n_5, we rationally start from the one with the highest individual gain, i.e. the Cross component. We

refer to this coding component as

$\xi_{n_5}^{\text{Cross}} =< n_5, \{n_1, n_3\}, \{n_3, n_1\}, \{n_2, n_4\}, \{n_4, n_2\} >$ where n_5 is the relay node, and each set of the edge nodes, $\{n_i, n_j\}$, represents a transmitter-receiver pair; or two consecutive links: (n_i, n_5) and (n_5, n_j) in this coding component. To find the maximum codable flow by coding component $\xi_{n_5}^{\text{Cross}}$, called as $W_{\xi_{n_5}^{\text{Cross}}}$, we have to calculate the common amount of flows on any two consecutive links in $\xi_{n_5}^{\text{Cross}}$, which is equal to the smallest one:

$$\begin{aligned} W_{\xi_{n_5}^{\text{Cross}}} &\leq \omega^{n_1 n_5 n_3} \\ W_{\xi_{n_5}^{\text{Cross}}} &\leq \omega^{n_2 n_5 n_4} \\ W_{\xi_{n_5}^{\text{Cross}}} &\leq \omega^{n_3 n_5 n_1} \\ W_{\xi_{n_5}^{\text{Cross}}} &\leq \omega^{n_4 n_5 n_2} \end{aligned} \tag{5.1}$$

where $\omega^{n_i n_5 n_j}$ is the amount of traffic flow to be sent from n_i to n_j through n_5. For simplicity, we combine the set of inequalities in (5.1) into one equation as:

$$W_{\xi_{n_5}^{\text{Cross}}} = \min_{\substack{\forall \{n_i, n_j\} \in \xi_{n_5}^{\text{Cross}} \\ \omega^{n_i n_5 n_j} \neq 0}} \omega^{n_i n_5 n_j}$$

Since $\omega^{n_1 n_5 n_3} = \omega^{n_2 n_5 n_4} = \omega^{n_3 n_5 n_1} = \omega^{n_4 n_5 n_2} = 1$, we find $W_{\xi_{n_5}^{\text{Cross}}} = 1$. Therefore, all packets are optimally encoded and routed in 3 time slots. Next, we have to check whether the routing is complete; or in other words, if there exists any undelivered packet, $W_{(n_i, n_j)}$, from n_i to n_j. Following these calculations:

$$\begin{aligned} W_{(n_1, n_3)} = W_{\xi_{n_5}^{\text{Cross}}} - \omega^{n_1 n_5 n_3} &= 0 \\ W_{(n_2, n_4)} = W_{\xi_{n_5}^{\text{Cross}}} - \omega^{n_2 n_5 n_4} &= 0 \\ W_{(n_3, n_1)} = W_{\xi_{n_5}^{\text{Cross}}} - \omega^{n_3 n_5 n_1} &= 0 \\ W_{(n_4, n_2)} = W_{\xi_{n_5}^{\text{Cross}}} - \omega^{n_4 n_5 n_2} &= 0 \end{aligned} \tag{5.2}$$

we conclude that no packet remains to be routed.

- **Example 2**

  $n_1 \xrightarrow{1} n_3, \quad n_3 \xrightarrow{2} n_1, \quad n_2 \xrightarrow{1} n_4, \quad n_4 \xrightarrow{1} n_2$**:**

  In this example, we find the Cross component $\xi_{n_5}^{\text{Cross}}$ as the optimal one (similar to the Example 1) which requires a total of 3 time slots for delivering 1 unit demands on each pair of links. However, after checking for the remaining packets (similar to equations (5.2)), we notice that $W_{(n_1,n_3)} = 1$, i.e. one packet remains to be sent from n_3 to n_1. Since we realize that no coding component is possible, this packet is delivered through two unicast transmissions, one from n_3 to n_5 and the second from n_5 to n_1. Thus, a total of 5 time slots is required to transfer all the sessions.

- **Example 3**

  $n_1 \xrightarrow{2} n_3, \quad n_3 \xrightarrow{2} n_1, \quad n_2 \xrightarrow{1} n_4, \quad n_4 \xrightarrow{1} n_2$**:**

  With the same logic, we start to encode the native packets by using the coding component with the highest gain, $\xi_{n_5}^{\text{Cross}}$, and we find this component feasible. After checking the remaining undelivered packets on each pair of links, we find that $W_{(n_1,n_3)} = 1$ and $W_{(n_3,n_1)} = 1$. In other words, the problem reduces to: $n_1 \xrightarrow{1} n_3, n_3 \xrightarrow{1} n_1$. Then, we search for the highest-gain feasible coding component at the relay, and find a Chain component $\xi_{n_5}^{\text{Chain}} =< n_5, \{n_1, n_3\}, \{n_3, n_1\} >$. By replacing the path flow variables $\omega^{n_1 n_5 n_3}$ and $\omega^{n_3 n_5 n_1}$ with the remaining packets on each pair of consecutive links, $\omega^{n_1 n_5 n_3} - W_{\xi_{n_5}^{\text{Cross}}}$ and $\omega^{n_3 n_5 n_1} - W_{\xi_{n_5}^{\text{Cross}}}$, respectively; we can write:

  $$W_{\xi_{n_5}^{\text{Chain}}} = \min_{\substack{\forall \{n_i, n_j\} \in \xi_{n_5}^{\text{Chain}} \\ \omega^{n_i n_5 n_j} \neq 0}} (\omega^{n_1 n_5 n_3} - W_{\xi_{n_5}^{\text{Cross}}})$$

  from which we obtain: $W_{\xi_{n_5}^{\text{Chain}}} = 1$. Therefore, the remaining unit flows on pair links $\{n_1, n_3\}$ and $\{n_3, n_1\}$ are encoded and sent through the Chain component, which uses 2 time slots. Thus, all the sessions are transfused in 5 time slots.

However, one may think if the order to apply coding components must be based on their individual gains, i.e. in descending order. Here and in Example 4, we probe this problem.

We repeat this example by applying one Chain component (Chain1), $\xi_{n_5}^{\text{Chain1}} =< n_5, \{n_1, n_3\}, \{n_3, n_1\} >$, first. From the following equation:

$$W_{\xi_{n_5}^{\text{Chain1}}} = \min_{\substack{\forall \{n_i, n_j\} \in \xi_{n_5}^{\text{Chain1}} \\ \omega^{n_i n_5 n_j} \neq 0}} \omega^{n_i n_5 n_j} \quad (5.3)$$

we obtain $W_{\xi_{n_5}^{\text{Chain1}}} = 2$, which requires a total of 4 time slots for delivery. Then we calculate the remaining flows and find: $W_{(n_1,n_3)} = W_{(n_3,n_1)} = 0$ and $W_{(n_2,n_4)} = W_{(n_4,n_2)} = 1$. Therefore, the only usable coding component is $\xi_{n_5}^{\text{Chain2}} =< n_5, \{n_2, n_4\}, \{n_4, n_2\} >$, that results in $W_{\xi_{n_5}^{\text{Chain2}}} = 1$, and needs 2 time slots to complete the routing. We observe that the required delivery time is increased from 5 to 6 slots.

We make the following proposition from our observation in this example.

Note 1: *For an optimal result, the consequence for coding packets at a relay node is to apply coding components in a decreasing gain manner.*

However, this proposition has to either be mathematically proven in the general cases, or disproved by finding a contradicting example.

- **Example 4**

$n_1 \xrightarrow{3} n_3, \quad n_3 \xrightarrow{2} n_1, \quad n_2 \xrightarrow{1} n_4, \quad n_4 \xrightarrow{0} n_2$:

Following our observation in Example 3, we select coding components based on descending order and start from the Cross component. However, given $n_4 \xrightarrow{0} n_2$, this component may not be selected. The next candidate is the Partial Cross component $\xi_{n_5}^{\text{P-Cross}} = (n_5, \{n_1, n_3\}, \{n_3, n_1\}, \{n_2, n_4\})$. From

$W_{\xi_{n_5}^{\text{P-Cross}}} = \min_{\forall \{n_i,n_j\} \in \xi_{n_5}^{\text{Cross}},\ \omega^{n_i n_5 n_j} \neq 0} \omega^{n_i n_5 n_j}$, we find $W_{\xi_{n_5}^{\text{P-Cross}}} = 1$ which uses 3 time slots for encoding and routing of 3 packets. After calculating the remaining packets on the links, the problem reduces to: $n_1 \xrightarrow{2} n_3, n_3 \xrightarrow{1} n_1$. Then, these packets are routed by using one Chain component, $W_{\xi_{n_5}^{\text{Chain}}(1)} = 1$, and two simple unicast transmissions. Thus, the total delivery time is obtained as 7.

On the other hand, by applying the Chain component $\xi_{n_5}^{\text{Chain2}} = (n_5, \{n_1, n_3\}, \{n_3, n_1\})$ at the begining, we find $W_{\xi_{n_5}^{\text{Chain2}}} = 2$ which uses 4 time slots. Then, the remaining flows: $n_1 \xrightarrow{1} n_3, n_2 \xrightarrow{1} n_4$ can be routed through using one X component, $\xi_{n_5}^{\text{X}} =< n_5, \{n_1, n_3\}, \{n_2, n_4\} >$ where $W_{\xi_{n_5}^{\text{X}}} = 1$. This component uses 2 time slots to complete the routing. Thus, all sessions are transfused to their destination in 6 time slots which is lower than the required time in the first method. This example contradicts the observation that we obtained in Example 3.

In these examples, we illustrated the method of routing and encoding packets by means of coding components. We showed how multiple coding components can be performed at any relay node. We also proved that the set of optimal choices of coding components is not necessarily achievable according to giving higher priority for higher-gain coding component. In the following section we explain how this set of optimal coding components are obtained in our mathematical model.

5.3 System Model and Parameters

We model a multihop wireless network as a directed graph $G = (N, L)$ where N is the set of all nodes equipped with multiple antennas and L is the set of all feasible transmission links in the network. We assume a set of M end-to-end (unicast) traffic sessions in this network. The m^{th} session is denoted by $S_m = \{(s_m, d_m, R_m) : s_m \in N, d_m \in N, R_m > 0, m = 1, ...M\}$. Therefore, in the m^{th} session, the source node s_m

sends the commodity R_m to the destination node d_m.

5.4 Opportunistic NC-Aware Routing Model

In this section, we introduce the required constraints to route all the sessions which are intersecting and can be potentially encoded together by determining the coded and uncoded amount of sessions on each link during. Our opportunistic coding-aware routing consists of flow-balance conservation, network coding, and bandwidth limitations and guarantees the optimal selection of routing paths along with coding components.

5.4.1 Flow-balance Constraints

In this work, we use the path-based approach which has shown to be both scalable and accurate [35] if enough routes are supplied to the model. In a path-based model, each source-destination pair is given a set of k alternate routes. Thus, the flow balance conservation are given as:

$$\sum_{i \in L_n^+} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m - \sum_{j \in L_n^-} \sum_{\psi \in \Psi_m} \delta_\psi^j \omega_\psi^m = \begin{cases} 0 & \text{if } \; n \neq s_m, d_m \\ R_m & \text{if } \; n = s_m \\ -R_m & \text{if } \; n = d_m \end{cases} \tag{5.4}$$

where Ψ_m denotes the set of predetermined k-shortest paths for session m, ω_ψ^m is a routing variable that shows the amount of traffic of session m on path ψ and δ_ψ^i is defined as a binary parameter with value 1 for any link i which is traversed by path ψ. Equations (5.4) preserve the flow balance at all the nodes in the network.

5.4.2 Network Coding Constraints

Packets from one session at the source may be divided into k different groups where each of them, ω_ψ^m, is routed through an individual path ψ toward the destination.

These packets are delivered either coded or uncoded. Let us define the binary parameter $\delta_{\psi}^{i,j}$ to be 1 if two consecutive links i and j, $((r(i) = t(j))$, are traversed by path ψ in this order; and otherwise 0. We use this parameter for determining the codable amount of ω_{ψ}^{m} through each available coding component in Figure 5.1.

Consider Ξ_n as the set of Q various coding components feasible at a relay node n, which are generated off-line. $W_{\xi_n^q}$ refers to the maximum codable flow by the q^{th} coding component $\xi_n^q \in \Xi_n$. As we showed in our illustrative examples, $W_{\xi_n^q}$, depends on the minimum flow sessions passing through each pair of sequential links $(i,j) \in \xi_n^q$. For now, suppose that coding components in Ξ_n can be ordered from 1 to Q, such that the first priority for encoding is reserved for ξ_n^1, the second priority is allocated to ξ_n^2 and so on, up to the least priority which is given to ξ_n^Q. Therefore, for each of ξ_n^q, $q = 1...Q$ we can write:

$$
\begin{aligned}
W_{\xi_n^1} &= \min_{\forall (i,j) \in \xi_n^1} \{ \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_{\psi}^{i,j} \omega_{\psi}^{m} \} \\
W_{\xi_n^2} &= \min_{\forall (i,j) \in \xi_n^q} \{ \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_{\psi}^{i,j} \omega_{\psi}^{m} - W_{\xi_n^1} \} \\
&\vdots \\
W_{\xi_n^Q} &= \min_{\forall (i,j) \in \xi_n^Q} \{ \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_{\psi}^{i,j} \omega_{\psi}^{m} - \sum_{\substack{\forall \xi_n^q \in \Xi_n: \\ (i,j) \in \xi_n^q, \\ q < Q}} W_{\xi_n^{q'}} \}
\end{aligned}
\tag{5.5}
$$

(5.5) show that for the coding component with lower priority, the available amount of traffic flowing on the pair (i,j) is updated and reduced based on the previous coded path flow. However, by summarizing these series of equations into (5.6), we do not need to distinguish preferences for each of the coding components; they will be eventually selected in an order according to the favor of the optimal solution.

$$
W_{\xi_n^q} = \min_{\forall (i,j) \in \xi_n^q} \{ \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_{\psi}^{i,j} \omega_{\psi}^{m} - \sum_{\substack{\forall \xi_n^{q'} \in \Xi_n: \\ (i,j) \in \xi_n^{q'}, \\ q \neq q'}} W_{\xi_n^{q'}} \} \tag{5.6}
$$

Note that $W_{\xi_n^q}$ not only represents the broadcast traffic at relay node n, but also shows the amount of each unicast traffic from any edge node in ξ_n^q toward the relay. We also denote that in our model, each coded packet is decodable at the next hop to reduce the practical issues for buffering the packets.

Thus, we have calculated the codable fraction of the traffic flowing on each path. Next, we determine the amount of uncoded packets that a relay node $t(i)$ unicasts to each individual neighbor $r(i)$. However, uncoded packets unicast by $t(i)$ will be either: 1) not encoded at the next hop $r(i)$, or 2) encoded at the next hop $r(i)$ by a $\xi_{r(i)}^q \in \Xi_{r(i)}$, where the quantity of them is equal to $\sum_{\forall \xi_{r(i)}^q \in \Xi_r(i): i \in \xi_{r(i)}^q} W_{\xi_{r(i)}^q}$. This uncoded traffic is denoted by W_i:

$$W_i = \sum_{m \in M} \sum_{\psi \in \Psi_m} \delta_\psi^i \omega_\psi^m - \sum_{\substack{\forall \xi_{t(i)}^q \in \Xi_{t(i)}: \\ i \in \xi_{t(i)}^q,}} W_{\xi_{t(i)}^q} - \sum_{\substack{\forall \xi_{r(i)}^q \in \Xi_{r(i)}: \\ i \in \xi_{r(i)}^q}} W_{\xi_{r(i)}^q} \tag{5.7}$$

5.4.3 Bandwidth Constraints

After computing the volume of each coded and uncoded traffic, we have to provide constraints to satisfy the second principle of the routing. According to the type of transmitted packets which we explained above, let us consider three transmission modes for a node n on the link i as: 1) broadcasting packets using any coding component $\xi_{t(i)}^q \in \Xi_{t(i)}$, 2) unicasting packets to be coded at the next hop using coding component $\xi_{r(i)}^q \in \Xi_{r(i)}$, and 3) the regular unicasting packets not to be coded. We define three different binary variables: $z_{i,\xi_{t(i)}^q}^p$, $u_{i,\xi_{r(i)}^q}^p$ and t_i^p, each corresponding to one of the transmission modes, respectively. Therefore, if in configuration p, link i is carrying coded packets related to the first transmission mode, then $z_{i,\xi_{t(l)}^q}^p = 1$. Similarly, if it is delivering uncoded packets for the purpose of becoming encoded in the next hop at time slot p, then $u_{i,\xi_{r(i)}^q}^p = 1$. In the case that i is carrying uncoded packets, whether $r(i)$ is the destination of that session, or $r(i)$ simply forwards them uncoded, then $t_i^p = 1$. By means of these variables and following (3.9), the set of

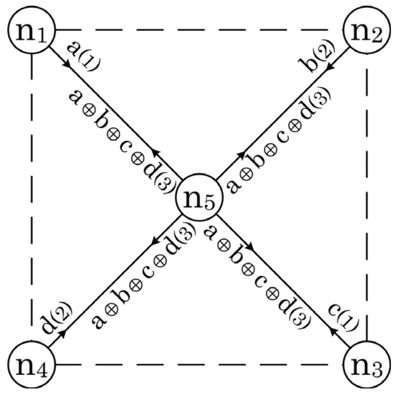


Figure 5.2: Scheduling for Cross component

bandwidth constraints for a link l can be written as:

$$\begin{cases} \sum_{p\in\mathcal{P}} z^p_{i,\xi^q_{t(i)}} c_i \lambda^p & \geq W_{\xi^q_{t(i)}} \quad \forall \xi^q_{t(i)} \in \Xi_{t(i)} \\ \sum_{p\in\mathcal{P}} \sum_{\substack{\forall \xi^q_{r(i)} \in \Xi_{r(i)}: \\ i \in \xi^q_{r(i)}}} u^p_{i,\xi^q_{r(i)}} c_i \lambda^p & \geq \sum_{\substack{\forall \xi^q_{r(i)} \in \Xi_{r(i)}: \\ i \in \xi^q_{r(i)}}} W_{\xi^q_{r(i)}} \\ \sum_{p\in\mathcal{P}} t^p_i c_i \lambda^p & \geq W_i \end{cases} \tag{5.8}$$

5.5 Scheduling for SIC and network coding with Opportunistic Listening

Earlier, a scheduling configuration was defined as a set of links that can be active concurrently, without violating the radio and interference constraints. This definition has been used in almost all the work in the literature [35, 107]. However, the problem of transmission scheduling in the presence of opportunistic listening NC is quite different.

We illustrate and elaborate this problem with an example in Figure 5.2 where four packets a, b, c, and d are supposed to be unicast from their sources, get encoded at the relay node n_5 and broadcast to all the destinations. Note that packets a and c can be unicast from n_1 and n_3 simultaneously due to MPR capability of n_5. During these transmissions, however, n_2 and n_4, which can receive multiple packets, are set into their receiving mode so that they can overhear both a and c, and later use them to detect their own desired packets from the coded one. Next, n_2 and n_4 unicast b and d to n_5 while n_1 and n_3 are overhearing these transmissions. Later, in the third time slot, all packets at n_5 are coded as $a \oplus b \oplus c \oplus d$ and transmitted. This example clarifies how the activity of different links in one coding component depends on each others. In other words, if our intention is to activate this coding component and we schedule a configuration which only includes the first transmissions of this coding component, there is no guarantee that other subsequent configurations will include the other set of unicasts or the required broadcast transmission at the relay. Indeed, a broadcast transmission at n_5 should be carried out only when the other two sets of unicast transmissions are performed and this can be guaranteed by scheduling the whole coding component in one configuration.

Therefore, in the case of network coding with opportunistic listening, we refer to a scheduling configuration as the set of coding components which are active concurrently. For simplicity, from now on we refer to a scheduling configuration as one configuration.

On the other hand, in scheduling coding components, we have to make sure that in each of the corresponding time slots, links are scheduled without violating the interference and radio constraints. Therefore, in this section we introduce a new scheduling scheme for NC with opportunistic listening where configurations are scheduled based on concurrent coding components. These coding components are scheduled in a way where their corresponding links in any time slot are not in conflict.

As we showed in Table 5.1, each of the coding components requires different number of time slots to complete its delivery of the packets. Hence, different configurations may consist of variable number of time slots depending on the scheduled coding components. In this work, the length of a configuration is equal to the length of the longest coding component, in terms of number of time slots (Figure 5.1), among all components scheduled in that configuration. For instance, if only unicast transmissions are scheduled, then the configuration consists of only one time slot. However, in the case where one or more of Chain, X or Bell components are scheduled, the configuration will consist of two time slots. Similarly, if the configuration has a Cross (also Partial Cross), Extended X or Extended Cross component, its length becomes 3, $2k_2$ or $\frac{k}{2}+1$, respectively. At a node n, let us define the h^{th} *Transmission Pattern* (TP), ζ_n^h, as either a coding component or a unicast transmission toward one of the neighbors. Therefore, we schedule transmission patterns instead of coding component to include unicast transmissions in our scheduling as well.

Transmission Pattern Scheduling

We refer to $\mathcal{S}$ as the set of all time slots in the current configuration. Thus, the length of the longest transmission pattern in this configuration is denoted by $|\mathcal{S}|$ which can be equal to 1 for unicast transmission and 2, 3, $2k_2$ or $\frac{k}{2}+1$ for any of coding components shown in Table 5.1. To determine the length of the current configuration, we need to find the number of time slots which has at least one active transmission due to a particular transmission pattern. We define the length of a configuration by $Lngth$, where:

$$Lngth = \sum_{s=1\ldots|\mathcal{S}|} x^s \tag{5.9}$$

$$x^s \geq \frac{\sum\limits_{i\in L} v^s_{i,\zeta_n^h}}{|L|} \qquad \forall s \in \mathcal{S} \tag{5.10}$$

where x^s is a binary variable which is equal to 1 whenever there is at leat one active link in time slot s; $|L|$ denotes the total number of the links in the network; and v^s_{i,ζ_n^h}

is defined as a binary variable showing whether link i is active in time slot s and it is involved in transmission pattern ζ_n^h. However, the first condition for a link to be capable of being active in this transmission pattern is being physically feasible. A link i is physically feasible if the radio and the interference constraints are met for it and thus, $v_i^s = 1$.

5.5.1 Radio Constraints

These constraints take care of the half-duplex and SIC properties of any node $n \in N$ with capability of receiving k_n packets simultaneously, and is given in (5.11) and (5.12), respectively.

$$v_i^s + v_j^s \leq 1 \qquad \forall n \in N : i \in L_n^-,\ j \in L_n^+ \tag{5.11}$$

$$\sum_{i \in L_n^-} v_i^s \leq k_n \qquad \forall n \in N \tag{5.12}$$

L_n is a subset of L that includes all the links connected to node n, $L_n^+ = \{i \in L : t(i) = n\}$ is the set of all links whose transmitter is node n and $L_n^- = \{i \in L : r(i) = n\}$ is the set of all links whose receiver is node n.

5.5.2 Interference Constraints

These constraints are provided based on SIC properties of the nodes is presented in a new format as:

$$PG_{t(i),r(i)} + M_i^s(1 - v_i^s) \geq \Gamma(\eta + I_1 - I_2) \tag{5.13}$$

$$I_1 = \sum_{j \neq i, j \in L} PG_{t(j),r(i)} \sum_{\forall \zeta_n^h \ni j} v_{j,\zeta_n^h}^s$$

$$I_2 = \sum_{l \neq i, j \in L : G_{t(i),r(i)} \geq G_{t(l),r(i)}} PG_{t(l),r(i)} \sum_{\forall \zeta_n^h \ni l} v_{l,\zeta_n^h}^s$$

which is presented by using the active links in the scheduled transmission patterns. In (5.13), I_1 denotes the overall interference at the receiver of link i, and I_2 represents that fraction of the interference which is decoded earlier and can be cancelled out. M_i^s in (5.13) is however a constant parameter satisfying (5.14) which is employed to reflect the interference constraints in a mathematical linear programming format.

$$M_i^s \geq \Gamma(\eta + \sum_{n \neq t(i), G_{t(i),r(i)} \geq G_{n,r(i)}} PG_{n,r(i)}) \tag{5.14}$$

5.5.3 Transmission Pattern Constraints

If the prospective scheduled TP is a coding component (rather than a unicast transmission), each of its links must be scheduled in a proper time slot. In the first step, we check whether a link i is physically feasible, by (5.15). In the case that $v_{i,\zeta_n^h}^s$ is related to a coding component where transmission on link i must be overheard by other nodes in the same component, the overhearing links should also be physically feasible, which is checked by equation (5.16). Moreover, constraints (5.17) guarantee that in a coding component, links i and j are concurrently active or inactive if they are supposed to be (e.g. exploited by Cross component in the case of NC+SIC in Table 5.1):

$$v_{i,\zeta_n^h}^s \leq v_i^s \qquad \forall s, \forall i \in \zeta_n^h \tag{5.15}$$

$$v_{i,\zeta_{r(i)}^h}^s \leq v_j^s \qquad \forall s, \forall j : t(j) = t(i), r(j) : \text{an opportunistic listener in } \zeta_{r(i)}^h \tag{5.16}$$

$$v_{i,\zeta_{r(i)}^h}^s = v_{j,\zeta_{r(i)}^h}^s \qquad \forall s, \forall i,j \in \zeta_{r(i)}^h : i,j \text{ must be active/inactive concurrently} \tag{5.17}$$

Let the binary variable $\Upsilon^{\zeta_n^h}$ be equal to 1, when the h^{th} transmission pattern is selected at node n to be scheduled in the current configuration. Equations (5.18) guarantee that in the whole configuration, each node is allowed to be active as a relay in only one coding component. However, (5.18) does not limit the number of unicast transmissions that a node can perform. We note that in any of the transmission patterns, an edge node which is not in transmitting, receiving or listening modes, can be active in an adjacent TP. In other words, we allow the transmission patterns to overlap.

$$\sum_{\forall \zeta_n^h : \zeta_n^h \text{is a coding component}} \Upsilon^{\zeta_n^h} \leq 1 \qquad \forall n \in N \tag{5.18}$$

On the other hand, a transmission pattern ζ_n^h can be selected to be scheduled in the current configuration if and only if every link in ζ_n^h can properly be scheduled in any time slots of this configuration. We show these constraints in (5.19). Constraints (5.20) guarantee that whenever a TP is scheduled at node n, it broadcasts the coded packet on all the links associated with this TP. We define a parameter $\kappa^{\zeta_n^h}$ to represent the number of links that the relay node n will transmit on in a transmission pattern ζ_n^h. Thus, if ζ_n^h is a unicast transmission, $\kappa^{\zeta_n^h} = 1$, and if ζ_n^h is one of the coding components in Table 5.1, $\kappa^{\zeta_n^h}$ is equal to the number of native packets hidden in one coded packet.

$$\Upsilon^{\zeta_n^h} = \sum_{\forall s \in \mathcal{S}} v_{i,\zeta_n^h}^s \qquad \forall i \in \zeta_n^h \qquad \forall n \in N \tag{5.19}$$

$$\sum_{\forall i \in \zeta_n^h : t(i)=n} \sum_{\forall s \in \mathcal{S}} v_{i,\zeta_n^h}^s = \Upsilon^{\zeta_n^h} . \kappa^{\zeta_n^h} \qquad \forall n \in N \tag{5.20}$$

5.6 Problem Formulation

We consider a set of traffic sessions (in terms of packet volume) and we are interested in finding the minimum scheduling period during which these sessions can be delivered. We suppose time is divided into slots and each configuration consists of one or more time slots. As we discussed earlier, one or more transmission patterns maybe active concurrently in one configuration while links active in each time slot should not interfere with each other. Let us denote a configuration by p and assume that $\mathcal{P}$ is the set of all feasible configurations. Therefore, we can represent the objective of our optimization problem by

$$\min \sum_{p \in \mathcal{P}} Lngth^p . \lambda^p \tag{5.21}$$

where λ^p denotes the number of times that configuration p is used and $Lngth^p$ is the length of this configuration (in unit of time slots), obtained from equation (5.9).

5.6.1 The Configuration Representation

Since this minimization should be performed through a joint routing, network coding and TP scheduling, we need to make the routing and network coding constraints (Section 5.4) compatible with the TP scheduling ones (Section 5.5). Therefore, we need to provide a connection between these two sets of constraints.

Once a set of concurrent transmission patterns (which satisfies the required scheduling constraints) is formed we call it a configuration p. This configuration is introduced to the rest of the model by the associated active links in the scheduled transmission patterns. The transmission details in each of the time slots, however, is not useful to the rest of the model. Assume that the h^{th} transmission pattern at node n is selected to be active in configuration p; therefore, $\Upsilon^{\zeta_n^h} = 1$ and the configuration includes all the links involved in this TP. Subsequently, the active links in each configuration can be used to route the packets and we have to make sure that the bandwidth constraints (5.8) are not violated. In these bandwidth constraints, we determine the required activity duration of each link in every configuration (λ^p) to guarantee that all the packets flowing on the link are delivered. Note that the amount of native and coded packets flowing on each link is determined by other constraints in the routing model.

We write the following equations for each of the configurations and all the links in any TP types which can be a coding component or a unicast transmission. $\forall p, \forall \Upsilon^{\zeta_n^h}, \forall i \in \Upsilon^{\zeta_n^h}$:

$$\Upsilon^{\zeta_n^h} = \begin{cases} z^p_{i,\xi_n^q} & \text{if } \zeta_n^h = \xi_n^q, n = t(i) \\ u^p_{i,\xi_n^q} & \text{if } \zeta_n^h = \xi_n^q, n = r(i) \\ t_i^p & \text{Otherwise} \end{cases} \tag{5.22}$$

where in the first case, link i is in a TP which is a coding component with $t(i)$ as the relay node. In other words, link i is carrying a coded packet which is broadcast at $t(i)$. Similarly, the TP is a coding component in the second case. However, the receiver of i is the relay node here, which means a native packet is unicast from the edge node $t(i)$ through i to the relay node. This packet will become encoded later. In the third case where the TP is not a coding component, link i carries a native packet. Here, $r(i)$ either is the final destination of this packet or will simply forward it to the next hop, without combining (encoding) it with any other packet.

5.6.2 Joint Routing, Network Coding and Scheduling

Our model for joint routing, network coding and transmission pattern scheduling problem can briefly be presented as:

$$\min \sum_{p \in \mathcal{P}} Lngth^p.\lambda^p \tag{5.23}$$

such that

$$Constraints : (5.4), (5.6) - (5.20), (5.22)$$

$$Lngth^p, \lambda^p \in R$$

which is a non-linear model. Moreover, the routing problem should be solved over all of the configurations in $\mathcal{P}$ to yield the optimal solution. However, the number of configurations ($|\mathcal{P}|$) grows exponentially with the possible number of transmission patterns involving different nodes in the network. Obviously, enumerating all such configurations is not practical. However, we decompose this problem into smaller subproblems using a Column Generation (CG) approach [26], to: 1) linearize it, and 2) obtain the optimal solution without having to enumerate all the configurations. CG decomposition is a very powerful method for obtaining the optimal solution of large sized linear problems. In CG, the problem is divided into two models: the

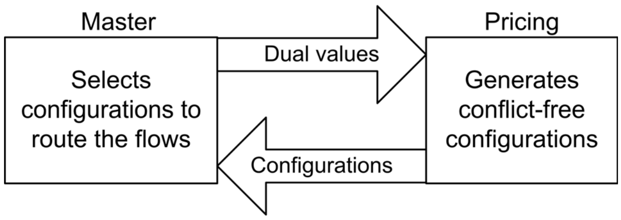


Figure 5.3: Interaction of the Master and the Pricing problems.

restricted Master problem and the Pricing problem [26]. More details in solving the problem by CG technique are explained in the following paragraphs.

5.6.3 Master and Pricing problems

We determine the routing of native and coded packets in the Master problem. The constraints in this model are flow balance conservations (5.4), network coding constraints (5.6)-(5.7) and bandwidth constraints (5.8). The Master and Pricing problems are solved iteratively, as shown in Figure 5.3. The Restricted Master Model is presented as:

$$\min \sum_{p \in \mathcal{P}_0} Lngth^p . \lambda^p \tag{5.24}$$

subject to

$$Constraints : (5.4), (5.6) - (5.8)$$

where $Lngth^p$ is a parameter used by the Master and determined in the Pricing. Hence, (5.24) is a linear program and can be easily solved. In the first iteration of solving (5.24), a set of initial feasible configurations, $\mathcal{P}_0 \subseteq \mathcal{P}$ is available to solve the first instance of the Master sub-problem. In this problem, each configuration in the initial $\mathcal{P}_0$ contains only one active transmission pattern which is only one link in unicast transmission ($t_i^p = 1$), and $|\mathcal{P}_0| = |L|$. The solution of the Master is the best routing over all available configurations ($\mathcal{P}_0$), which is a local optimum. When the Master obtains the local optimal solution, it generates the dual values of the bandwidth constraints, as $\{\sigma^i_{\xi^q_n}, \tau^i_{\xi^q_n}, \nu^i : \; i \in L\}$ which depends on the transmission pattern where i is involved. Then, these dual values are sent to the Pricing where the best link scheduling is determined. Then, the Pricing sub-problem applies these dual values to construct the Pricing objective as:

$$\min_{i\in L}(Lngth^p - \sum_{\forall \xi^q_{t(i)} \ni \forall i} \sigma^i_{\xi^q_{t(i)}} z^p_{i,\xi^q_{t(i)}} - \sum_{\forall \xi^q_{r(i)} \ni \forall i} \tau^i_{\xi^q_{r(i)}} u^p_{i,\xi^q_{r(i)}} - \sum_{\forall i} \nu^i t^p_i) \tag{5.25}$$

where $(Lngth^p - \sum_{\forall \xi^q_{t(i)} \ni \forall i} \sigma^i_{\xi^q_{t(i)}} z^p_{i,\xi^q_{t(i)}} - \sum_{\forall \xi^q_{r(i)} \ni \forall i} \tau^i_{\xi^q_{r(i)}} u^p_{i,\xi^q_{r(i)}} - \sum_{\forall i} \nu^i t^p_i)$ is called the *Reduced Cost*. The constraints of the Pricing are those given in (5.9)-(5.20) and (5.22). The Pricing Model is presented as:

$$\min_{i\in L}(Lngth^p - \sum_{\forall \xi^q_{t(i)} \ni \forall i} \sigma^i_{\xi^q_{t(i)}} z^p_{i,\xi^q_{t(i)}} - \sum_{\forall \xi^q_{r(i)} \ni \forall i} \tau^i_{\xi^q_{r(i)}} u^p_{i,\xi^q_{r(i)}} - \sum_{\forall i} \nu^i t^p_i)$$

subject to (5.26)

$$Constraints : (5.9) - (5.20), (5.22)$$

From now on, in each iteration, a feasible configuration is generated in the Pricing sub-problem to improve the objective of the Master. The value of the Pricing objective is always checked to determine the optimality of the solution. If (5.25) is a non-negative value, the obtained solution in the Master sub-problem is the optimal solution to the main problem. In this case, generating more configurations in the Pricing would not provide further improvement, because the optimal solution has already been found. A proof of optimality of the solution is given in [26]. On the other hand, if (5.25) has a negative value, the Pricing generates another configuration named p and sends it to the Master sub-problem. Then, the Master adds p into the set of previous configurations, $\mathcal{P}_0 \biguplus p \rightarrow \mathcal{P}_0$, and resolves the routing problem. Whenever a solution is produced by the Master, the corresponding dual values are passed to the Pricing and the procedure continues until finding the optimal solution.

5.6.4 CG-based Models for NC and NC+SIC

We have provided our complete CG-based for network coding combined with SIC model in Table 5.2, and the CG-based model using only network coding without successive interference cancellation is illustrated in Table 5.3.

Table 5.2: Column Generation model of NC+SIC

[RM problem] :

$$\min \sum_{p\in\mathcal{P}_0} Lngth^p.\lambda^p$$

subject to

$$\sum_{i\in L_n^+}\sum_{\psi\in\Psi_m}\delta_\psi^i\omega_\psi^m - \sum_{j\in L_n^-}\sum_{\psi\in\Psi_m}\delta_\psi^j\omega_\psi^m = \begin{cases} 0 & \text{if } n\neq s_m, d_m \\ R_m & \text{if } n = s_m \\ -R_m & \text{if } n = d_m \end{cases}$$

$$W_{\xi_n^q} = \min_{\forall(i,j)\in\xi_n^q}\{\sum_{m\in M}\sum_{\psi\in\Psi_m}\delta_\psi^{i,j}\omega_\psi^m - \sum_{\substack{\forall\xi_n^{q'}\in\Xi_n:\\ (i,j)\in\xi_n^{q'},\\ q\neq q'}} W_{\xi_n^{q'}}\}$$

$$W_i = \sum_{m\in M}\sum_{\psi\in\Psi_m}\delta_\psi^i\omega_\psi^m - \sum_{\substack{\forall\xi_{t(i)}^q\in\Xi_{t(i)}:\\ i\in\xi_{t(i)}^q,}} W_{\xi_{t(i)}^q} - \sum_{\substack{\forall\xi_{r(i)}^q\in\Xi_{r(i)}:\\ i\in\xi_{r(i)}^q}} W_{\xi_{r(i)}^q}$$

$$\begin{cases} \sum_{p\in\mathcal{P}} z_{i,\xi_{t(i)}^q}^p c_i\lambda^p \geq W_{\xi_{t(i)}^q} & \forall\xi_{t(i)}^q\in\Xi_{t(i)} \\ \sum_{p\in\mathcal{P}}\sum_{\substack{\forall\xi_{r(i)}^q\in\Xi_{r(i)}:\\ i\in\xi_{r(i)}^q}} u_{i,\xi_{r(i)}^q}^p c_i\lambda^p \geq \sum_{\substack{\forall\xi_{r(i)}^q\in\Xi_{r(i)}:\\ i\in\xi_{r(i)}^q}} W_{\xi_{r(i)}^q} & \\ \sum_{p\in\mathcal{P}} t_i^p c_i\lambda^p \geq W_i & \end{cases}$$

[Pricing problem] :

$$\min_{i\in L}(Lngth^p - \sum_{\forall\xi_{t(i)}^q\ni\forall i}\sigma_{\xi_{t(i)}^q}^i z_{i,\xi_{t(i)}^q}^p - \sum_{\forall\xi_{r(i)}^q\ni\forall i}\tau_{\xi_{r(i)}^q}^i u_{i,\xi_{r(i)}^q}^p - \sum_{\forall i}\nu^i t_i^p)$$

subject to

$$Lngth = \textstyle\sum_{s=1\ldots|\mathcal{S}|} x^s$$

$$x^s \geq \frac{\sum_{i\in L} v_{i,\zeta_n^h}^s}{|L|} \qquad \forall s\in\mathcal{S}$$

$$v_i^s + v_j^s \leq 1 \qquad \forall n\in N: i\in L_n^-,\ j\in L_n^+$$

Table 5.2: (continued)

$$\sum_{i\in L_n^-} v_i^s \leq k_n \qquad \forall n \in N$$

$$PG_{t(i),r(i)} + M_i^s(1-v_i^s) \geq \Gamma(\eta + \sum_{j\neq i, j\in L} PG_{t(j),r(i)} \sum_{\forall \zeta_n^h \ni j} v_{j,\zeta_n^h}^s - \sum_{l\neq i, j\in L: G_{t(i),r(i)} \geq G_{t(l),r(i)}} PG_{t(l),r(i)} \sum_{\forall \zeta_n^h \ni l} v_{l,\zeta_n^h}^s)$$

$$M_i^s \geq \Gamma(\eta + \sum_{n\neq t(i), G_{t(i),r(i)} \geq G_{n,r(i)}} PG_{n,r(i)})$$

$$v_{i,\zeta_n^h}^s \leq v_i^s \quad \forall s, \forall i \in \zeta_n^h$$

$v_{i,\zeta_{r(i)}^h}^s \leq v_j^s \quad \forall s, \forall j : t(j) = t(i), r(j)$: an opportunistic listener in $\zeta_{r(i)}^h$

$v_{i,\zeta_{r(i)}^h}^s = v_{j,\zeta_{r(i)}^h}^s \quad \forall s, \forall i,j \in \zeta_{r(i)}^h$: i,j must be active/inactive concurrently

$$\sum_{\forall \zeta_n^h : \zeta_n^h \text{is a coding component}} \Upsilon^{\zeta_n^h} \leq 1 \qquad \forall n \in N$$

$$\Upsilon^{\zeta_n^h} = \sum_{\forall s\in \mathcal{S}} v_{i,\zeta_n^h}^s \qquad \forall i \in \zeta_n^h \qquad \forall n \in N$$

$$\sum_{\forall i\in \zeta_n^h : t(i)=n} \sum_{\forall s\in \mathcal{S}} v_{i,\zeta_n^h}^s = \Upsilon^{\zeta_n^h}.\kappa^{\zeta_n^h} \qquad \forall n \in N$$

$$\Upsilon^{\zeta_n^h} = \begin{cases} z_{i,\xi_n^q}^p & \text{if } \zeta_n^h = \xi_n^q, n = t(i) \\ u_{i,\xi_n^q}^p & \text{if } \zeta_n^h = \xi_n^q, n = r(i) \\ t_i^p & \text{Otherwise} \end{cases}$$

Table 5.3: Column Generation model of NC

[RM problem] :

$$\min \sum_{p\in\mathcal{P}_0} Lngth^p.\lambda^p$$

subject to

$$\sum_{i\in L_n^+}\sum_{\psi\in\Psi_m}\delta_\psi^i\omega_\psi^m - \sum_{j\in L_n^-}\sum_{\psi\in\Psi_m}\delta_\psi^j\omega_\psi^m = \begin{cases} 0 & \text{if } n\neq s_m, d_m \\ R_m & \text{if } n = s_m \\ -R_m & \text{if } n = d_m \end{cases}$$

$$W_{\xi_n^q} = \min_{\forall(i,j)\in\xi_n^q}\{\sum_{m\in M}\sum_{\psi\in\Psi_m}\delta_\psi^{i,j}\omega_\psi^m - \sum_{\substack{\forall\xi_n^{q'}\in\Xi_n:\\(i,j)\in\xi_n^{q'},\\q\neq q'}} W_{\xi_n^{q'}}\}$$

$$W_i = \sum_{m\in M}\sum_{\psi\in\Psi_m}\delta_\psi^i\omega_\psi^m - \sum_{\substack{\forall\xi_{t(i)}^q\in\Xi_{t(i)}:\\i\in\xi_{t(i)}^q,}} W_{\xi_{t(i)}^q} - \sum_{\substack{\forall\xi_{r(i)}^q\in\Xi_{r(i)}:\\i\in\xi_{r(i)}^q}} W_{\xi_{r(i)}^q}$$

$$\begin{cases} \sum_{p\in\mathcal{P}} z_{i,\xi_{t(i)}^q}^p c_i\lambda^p \geq W_{\xi_{t(i)}^q} \quad \forall\xi_{t(i)}^q\in\Xi_{t(i)} \\ \sum_{p\in\mathcal{P}}\sum_{\substack{\forall\xi_{r(i)}^q\in\Xi_{r(i)}:\\i\in\xi_{r(i)}^q}} u_{i,\xi_{r(i)}^q}^p c_i\lambda^p \geq \sum_{\substack{\forall\xi_{r(i)}^q\in\Xi_{r(i)}:\\i\in\xi_{r(i)}^q}} W_{\xi_{r(i)}^q} \\ \sum_{p\in\mathcal{P}} t_i^p c_i\lambda^p \geq W_i \end{cases}$$

[Pricing problem] :

$$\min_{i\in L}(Lngth^p - \sum_{\forall\xi_{t(i)}^q\ni\forall i}\sigma_{\xi_{t(i)}^q}^i z_{i,\xi_{t(i)}^q}^p - \sum_{\forall\xi_{r(i)}^q\ni\forall i}\tau_{\xi_{r(i)}^q}^i u_{i,\xi_{r(i)}^q}^p - \sum_{\forall i}\nu^i t_i^p)$$

subject to

$$Lngth = \textstyle\sum_{s=1\ldots|\mathcal{S}|} x^s$$

$$x^s \geq \frac{\sum_{i\in L} v_{i,\zeta_n^h}^s}{|L|} \qquad \forall s\in\mathcal{S}$$

$$v_i^s + v_j^s \leq 1 \qquad \forall n\in N: i\in L_n^-,\ j\in L_n^+$$

$$\sum_{i\in L_n^-} v_i^s \leq k_n \qquad \forall n\in N$$

$$PG_{t(i),r(i)} + M_i^s(1-v_i^s) \geq \Gamma(\eta + \sum_{j\neq i, j\in L} PG_{t(j),r(i)}\sum_{\forall\zeta_n^h\ni j} v_{j,\zeta_n^h}^s)$$

$$M_i^s \geq \Gamma(\eta + \sum_{n\neq t(i), G_{t(i),r(i)}\geq G_{n,r(i)}} PG_{n,r(i)})$$

$$v_{i,\zeta_n^h}^s \leq v_i^s \quad \forall s, \forall i\in\zeta_n^h$$

$$v_{i,\zeta_{r(i)}^h}^s \leq v_j^s \quad \forall s, \forall j: t(j)=t(i), r(j): \text{an opportunistic listener in } \zeta_{r(i)}^h$$

$$v_{i,\zeta_{r(i)}^h}^s = v_{j,\zeta_{r(i)}^h}^s \quad \forall s, \forall i,j\in\zeta_{r(i)}^h: i,j \text{ must be active/inactive concurrently}$$

$$\sum_{\forall\zeta_n^h:\zeta_n^h \text{is a coding component}} \Upsilon^{\zeta_n^h} \leq 1 \qquad \forall n\in N$$

$$\Upsilon^{\zeta_n^h} = \sum_{\forall s\in\mathcal{S}} v_{i,\zeta_n^h}^s \qquad \forall i\in\zeta_n^h \qquad \forall n\in N$$

$$\sum_{\forall i\in\zeta_n^h:t(i)=n}\sum_{\forall s\in\mathcal{S}} v_{i,\zeta_n^h}^s = \Upsilon^{\zeta_n^h}.\kappa^{\zeta_n^h} \qquad \forall n\in N$$

Table 5.3: (continued)

$\Upsilon^{\zeta_n^h} = \begin{cases} z^p_{i,\xi^q_n} & \text{if } \zeta^h_n = \xi^q_n, n = t(i) \\ u^p_{i,\xi^q_n} & \text{if } \zeta^h_n = \xi^q_n, n = r(i) \\ t^p_i & \text{Otherwise} \end{cases}$

5.7 Conclusion

In this chapter we provided a decomposed model for optimizing wireless networks using network coding with opportunistic listening. To optimize the scheduling of different coding components with/without opportunistic listening, we introduced a new TP scheduling scheme. In this new scheme, variable length configurations are generated to support conflict free link scheduling. We have also developed an optimization model for the case where network coding is combined with SIC. Through our optimization models, we study the benefits of combining network coding with SIC over usage of network coding through numerical results which are deferred to be shown in Chapter 6.

Chapter 6

Numerical Results of SIC and Network Coding in MWNs

In this chapter, we **first** evaluate the performance of successive interference cancellation combined with network coding. We highlight the benefits of this method in comparison with three other methods where either network coding or successive interference cancellation or none of them is used. We have provided CG-based models for the mentioned methods denoted by NC, MPR and BM which stand for network coding, multi-packet reception and base model (neither network coding nor multi-packet reception capabilities), respectively. These models were explained in Chapters 3, 4 and 5. Next, we investigate the performance of SIC in combination with power control which was explained in Chapter 4.

We provide the numerical results using four examples. We assume networks that are located in an urban environment and the path loss exponent β is equal to 4. In our evaluation, we consider equal transmission power at all nodes, which is $P = 1$mW in Examples I and II. However we vary the transmission power **of all the nodes** in the third example. In example IV, however, we have employed power control technique in which the transmission power of each node is adjusted to an optimal value obtained from the optimization model. In all of the examples, P is measured at $d_0 = 1$m from the transmitter; the SINR threshold Γ is set to 1 and the background noise power is assumed to be equal to 10^{-6}mW. In all of these examples, the feasible coding components are generated off-line; and the Pricing selects them from the list through regenerating them using the scheduling constraints and forms new configurations. These three examples are explained as follows.

6.1 Example I

The network in this example consists of 6 nodes, as shown in Figure 6.1. There are 22 possible coding components which are listed in Table 6.1. The structure of the coding components are shown in Figure 6.2. In this example, we refer to the i^{th} coding component as ξ^i. We assume that six traffic sessions have to be delivered which are shown in Table 6.2. For each session, we consider a single path available for the routing. We have obtained the total required time for delivering all the sessions in each of the NC-MPR, NC, MPR and BM methods which is shown in Table 6.3. Table 6.4 illustrates the details of transmission scheduling for each of the NC-MPR, NC, MPR and BM methods in single-path routing. We provide the set of used configurations p according to the active transmission patterns in them, the configuration usage frequency (λ^p), the length of each configuration in terms of time slots ($Lngth^p$), and the active links in each time slot s within one configuration. In this table, we denote the TP corresponding to the unicast transmission on link i by ς^i, and corresponding to the j^{th} coding component in Table 6.4 by ξ^j. The numerical results for each of the methods are discussed next.

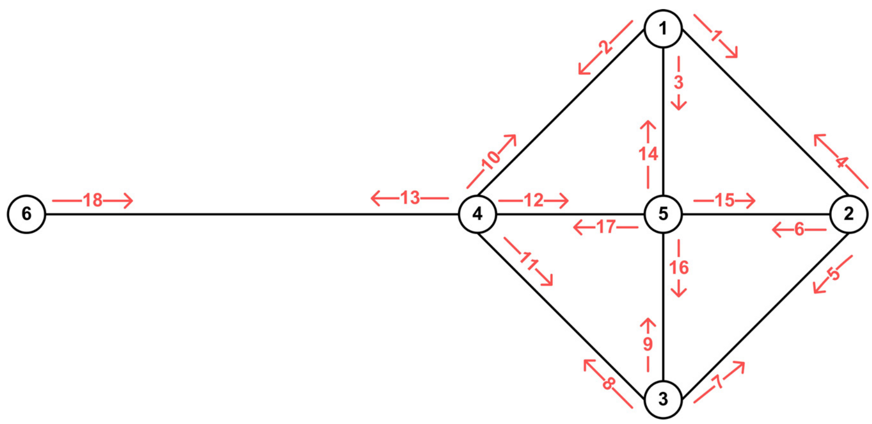


Figure 6.1: Network topology of Example 1

ξ	Type	Relay Node	Edge Nodes	Involved Links
1	Chain	1	2,4	1,2,4,10
2	Chain	2	1,3	1,4,5,7
3	Chain	3	2,4	5,7,8,11
4	Chain	4	1,6	2,10,13,18
5	Chain	4	1,3	2,8,10,11
6	Chain	4	3,6	8,11,13,18
7	Chain	4	5,6	12,13,17,18
8	Bell	4	1,5,6	10,13,17,18
9	Bell	4	3,5,6	11,13,17,18
10	Bell	4	1,5,6	2,12,13,18
11	Bell	4	3,5,6	8,12,13,18
12	Cross	5	1,2,3,4	3,6,9,12, 14,15,16,17
13	P-Cross	5	1,2,3,4	3,6,9,14, 16,17
14	P-Cross	5	1,2,3,4	6,9,12,14, 15,17
15	P-Cross	5	1,2,3,4	3,9,12,14, 15,16
16	P-Cross	5	1,2,3,4	3,6,12,15, 16,17
17	X	5	1,2,3,4	3,6,16,17
18	X	5	1,2,3,4	6,9,14,17
19	X	5	1,2,3,4	9,12,14,15
20	X	5	1,2,3,4	3,12,15,16
21	Chain	5	1,3	3,9,14,16
22	Chain	5	2,4	6,12,15,17

Table 6.1: Coding components in Example I.

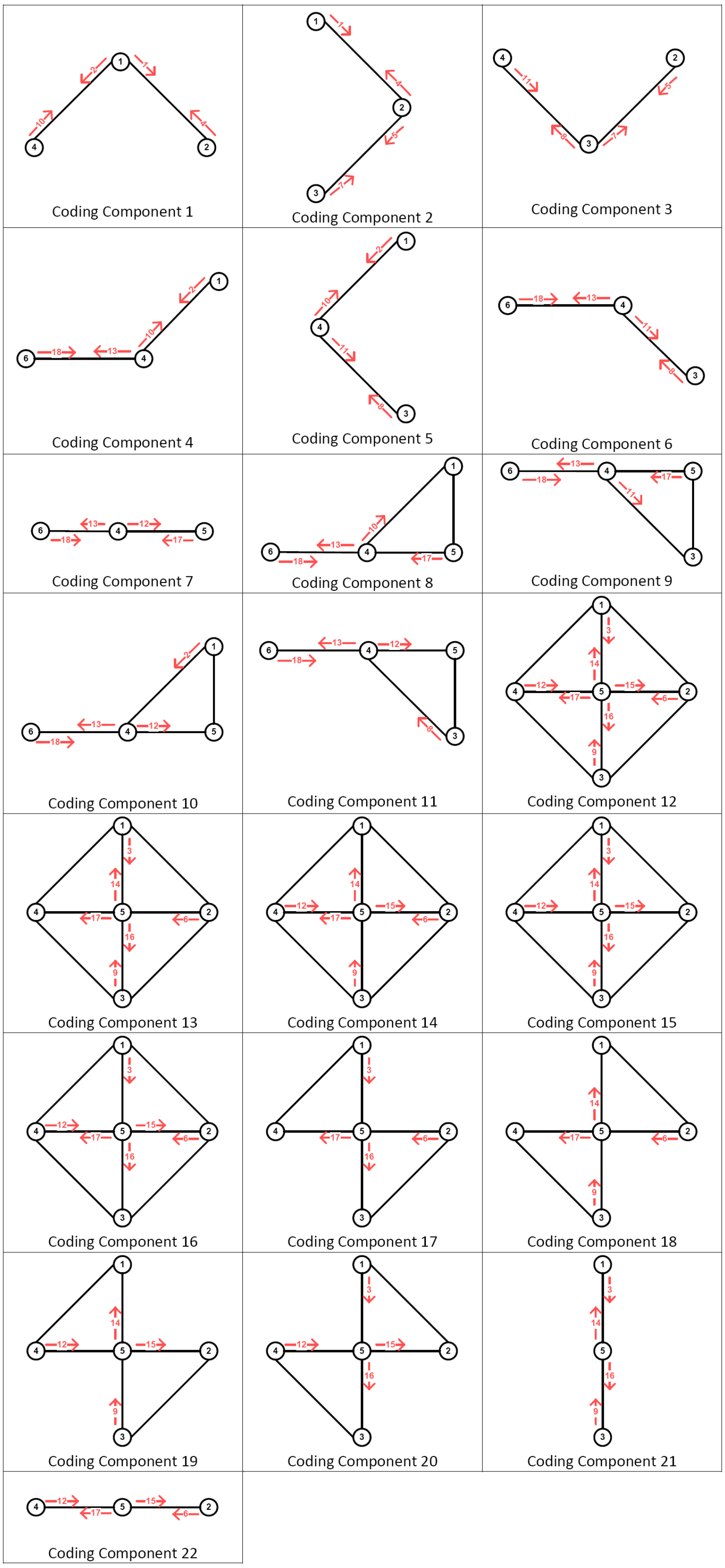


Figure 6.2: The structure of all feasible coding components in Example 1

m	$n_s \xrightarrow{R_m} n_d$	Single Path
1	$1 \xrightarrow{3} 3$	$[3, 16]$
2	$3 \xrightarrow{2} 1$	$[9, 14]$
3	$2 \xrightarrow{1} 4$	$[6, 17]$
4	$4 \xrightarrow{1} 2$	$[12, 15]$
5	$5 \xrightarrow{1} 6$	$[17, 13]$
6	$6 \xrightarrow{2} 2$	$[18, 12, 15]$

Table 6.2: Traffic sessions and their routing paths in Example I.

6.1.1 BM method

We observe that 19 time slots are required to deliver all the packets, hence, transmission links may be scheduled concurrently, if they can be spatially reused. For example none of the required links (to carry the sessions) can be simultaneously active with link 3, while links 13 and 14 are concurrent in configuration p_8 because they do not cause strong interference on each other. However, note that configurations of transmission patterns in the Pricing are generated based on the information about routing of the sessions in the RM. This information is passed from the RM to the Pricing through dual values of the bandwidth constraints. Therefore, we notice that the model also uses configuration p_4 where link 14 is individually active; since there is no need to have link 13 active as well.

6.1.2 MPR method

In this method, by allowing the nodes to receive concurrent packets, the spatial reuse is improved and more number of links can simultaneously be active. For example, each set of links $\{3, 9, 12\}$ and $\{17, 18\}$ are all active in one time slot in configurations p_6 and p_8, respectively. Therefore, the overall delivery time reduces to 13 time slots which shows a $\frac{(19-13)\times 100}{19} = 30.8\%$ gain compared to the BM method.

6.1.3 NC method

Unlike the BM and the MPR methods where all the configurations consist of only one time slot, in this method, we allow each configuration to consist of different number of time slots due to the variable length of the used transmission patterns (unicast transmissions or coding components with or without opportunistic listening). These configurations can include only one TP which is either a single unicast transmission in one time slot (e.g. p_1) or one coding component consisting of several time slots. For example, ξ^{15} which is a Partial Cross component, needs a total of 4 time slots in p_2; where in the first three time slots s_1, s_2 and s_3 packets are unicast from the edge nodes 1, 3 and 4 to the relay node 5; and 5 broadcasts the coded packet to nodes 1, 2 and 3 in s_4. The configurations in the NC method can also include several transmission patterns of different unicast transmissions and/or coding components. The concurrent transmission patterns occur when all of them can have their links scheduled concurrently and conflict-free in any time slot of the configuration. In this example, we observe that p_3 includes two transmission patterns, a Cross component (ξ^{12}) along with a pure unicast transmission on link 18 (ς^{18}) which is concurrent with link 6 of ξ^{12} in s_2. Through broadcasting a coded packet, which contains several native packets, to several receiving nodes, NC method reduces the transmission overhead and requires only 13 slots to deliver all the demands which shows a 30.8% gain over the BM method.

Another obvious benefit of enabling network coding at the nodes is the power efficiency. Our numerical results in Table 6.3 illustrate that the 21 transmissions required in the BM and MPR methods reduces to 6 in the case of NC method, where half of them are unicast, and the remaining are broadcast, transmissions. At this point, one important engineering insight is that by using network coding, we can save up to $\frac{(21-6)\times 100}{21} = 71\%$ of the transmission power; since we consider equal transmission power in the unicast and the broadcast transmissions. This observation

Method	Delivery Time	$\lvert\mathcal{P}_0\rvert$	# of used p	# of Broadcasts	# of Unicasts
BM	19	27	9	-	21
MPR	13	28	7	-	21
NC	13	34	4	3	3
NC-MPR	9	40	4	3	3

Table 6.3: Numerical results in Example I.

reveals one of the most significant benefits of the network coding technique. We note that in practice, the power consumption to perform network coding and decoding is negligible to the transmission power [89].

6.1.4 NC-MPR method

By using this method, the scheduling time reduces to 9 time slots which is 52.6% better than the BM method. In this example, we notice that the structure of the transmission patterns in the used configurations in the NC-MPR is similar to the NC method. However, due to SIC properties in the NC-MPR, the length of each coding component is shorter than the same coding component in the NC method, which yields a shorter overall scheduling time. For example, while p_4 in both NC and NC-MPR methods include the same transmission patterns ς^{13} and ξ^{20}, it is one time slot shorter in the NC-MPR model, because nodes 1 and 4 (links 3 and 12) can unicast their native packets to node 5 and save one time slot due to SIC properties.

This example shows that the combination of SIC and network coding not only provides a greater improvement in reducing the delivery time compared to each of the NC, MPR or BM methods but also provides a better power efficiency similar to the NC method.

Method	Active transmission patterns in each p	$Lngth^p$	λ^p	Active links in each time slot
BM	p_1: $\{\varsigma^3\}$	1	3	s_1:$\{3\}$
	p_2: $\{\varsigma^2\}$	1	2	s_1:$\{9\}$
	p_3: $\{\varsigma^{12}\}$	1	3	s_1:$\{12\}$
	p_4: $\{\varsigma^{14}\}$	1	1	s_1:$\{14\}$
	p_5: $\{\varsigma^{15}\}$	1	3	s_1:$\{15\}$
	p_6: $\{\varsigma^{16}\}$	1	3	s_1:$\{16\}$
	p_7: $\{\varsigma^{17}\}$	1	2	s_1:$\{17\}$
	p_8: $\{\varsigma^{13}, \varsigma^{14}\}$	1	1	s_1:$\{13, 14\}$
	p_9: $\{\varsigma^{6}, \varsigma^{18}\}$	1	1	s_1:$\{6, 18\}$
MPR	p_1: $\{\varsigma^4\}$	1	1	s_1:$\{12\}$
	p_2: $\{\varsigma^{15}\}$	1	3	s_1:$\{14\}$
	p_3: $\{\varsigma^{16}\}$	1	3	s_1:$\{15\}$
	p_4: $\{\varsigma^{17}\}$	1	1	s_1:$\{16\}$
	p_5: $\{\varsigma^3, \varsigma^6, \varsigma^9, \varsigma^{12}\}$	1	2	s_1:$\{3, 6, 9, 12\}$
	p_6: $\{\varsigma^4, \varsigma^{17}, \varsigma^{18}\}$	1	1	s_1:$\{4, 17, 18\}$
	p_7: $\{\varsigma^{13}, \varsigma^{14}\}$	1	1	s_1:$\{13, 14\}$
NC	p_1: $\{\varsigma^{17}\}$	1	1	s_1:$\{17\}$
	p_2: $\{\xi^{15}\}$	4	1	s_1:$\{3\}$
				s_2:$\{9\}$
				s_3:$\{12\}$
				s_4:$\{14, 15, 16\}$
	p_3: $\{\varsigma^{18}, \xi^{12}\}$	5	1	s_1:$\{3\}$
				s_2:$\{6, 18\}$
				s_3:$\{9\}$
				s_4:$\{12\}$
				s_5:$\{14, 15, 16, 17\}$
	p_4: $\{\varsigma^{13}, \xi^{20}\}$	3	1	s_1:$\{3\}$
				s_2:$\{12\}$
				s_3:$\{13, 15, 16\}$
NC-MPR	p_1: $\{\varsigma^{17}\}$	1	1	s_1:$\{17\}$
	p_2: $\{\xi^{12}\}$	3	1	s_1:$\{3, 9\}$
				s_2:$\{6, 12\}$
				s_3:$\{14, 15, 16, 17\}$
	p_3: $\{\varsigma^{18}, \xi^{15}\}$	3	1	s_1:$\{3, 9, 18\}$
				s_2:$\{12\}$
				s_3:$\{14, 15, 16\}$
	p_4: $\{\varsigma^{13}, \xi^{20}\}$	2	1	s_1:$\{3, 12\}$
				s_2:$\{13, 15, 16\}$

Table 6.4: Details of the used configurations in Example I.

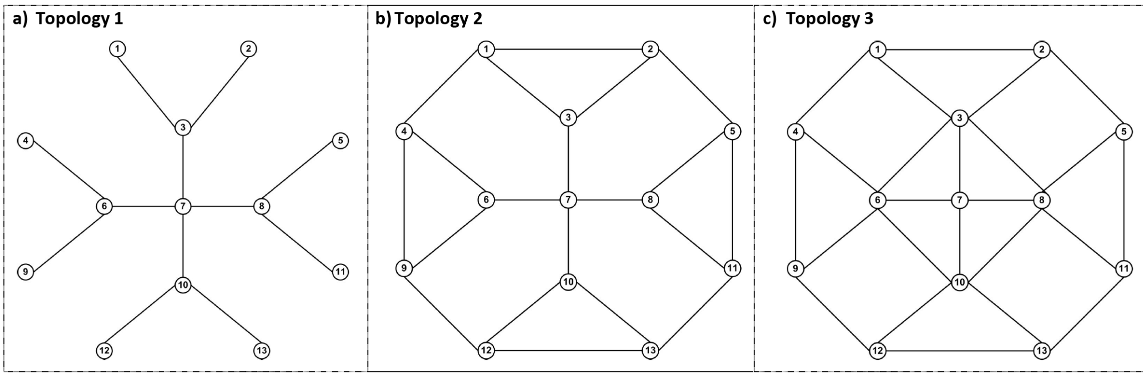


Figure 6.3: Network topologies in Example II: a)Top1: 24 links, b)Top2: 40 links, c)Top3: 48 links.

6.2 Example II

In this example, we consider a network consisting of 13 nodes in three different topologies of low, medium and high connectivity, illustrated in Figure 6.3, and named as Top1, Top2 and Top3. In each topology, we assume a fixed transmission power of 1 mW, and the links are defined at the switching center of the network to allow whether two nodes can communicate or not. We study the performance of all the four methods BM, MPR, NC and NC-MPR, in these three networks where two scenarios of light and medium traffic sessions are considered. We assume a 20 random unit-packet sessions as the light traffic and a 60 random unit-packet sessions as the medium one. To increase the accuracy of our numerical results, we have repeated each run of the models for 3 different random sessions in these cases. We evaluate the performance of our models in Top1 with a single path routing (no extra path can be provided) and in Top2 and Top3, with single, two and three paths routing.

The numerical results of this example which are shown in Figures 6.4 and 6.5, affirm that in all of the considered scenarios, the performance of the NC-MPR method is superior, followed by MPR and NC methods. These results reveal that indeed we can significantly improve the performance of network coding or SIC through combining them with each other. We present the relative gain of NC-MPR, NC and

MPR method over the BM in Table 6.5. We also illustrate the relative gain of the NC-MPR over the NC and MPR in Table 6.5. These gains are higher in the case of NC-MPR/NC which shows that SIC contributes better to improve the performance of NC-MPR method. Some other engineering insights about our numerical results are discussed as follows.

6.2.1 Impact of connectivity

By comparing the average of overall required time slots for the 1-path routing in any of Top1, Top2 and Top3 in Figures 6.4 and 6.5, we realize that this scheduling time is the least in Top3 for all the methods. Because in Top3 where there are more transmission links available, the single available shortest path is on average shorter (in terms of number of hops) than Top2 and Top1.

The NC method shows a significantly higher gain over the BM in Top1 due to the same reason. In Top1, the single available path for the routing is longer. Moreover, the sessions are intersecting more frequently due to less number of available links. These two reasons cause more opportunities for the packets to intersect, be coded and broadcast.

6.2.2 Impact of path selection

The results show that in the case of Top2 and Top3 in Figures 6.4 and 6.5, the delivery time reduces for all the methods when more than one path is available for the routing. However, the improvement obtained by adding a third path is negligible (for their particular instance).

We notice that in the case of NC method, increasing the number of paths leads to fewer network coding opportunities, because the sessions will rarely intersect on their routes. Therefore, this method shows the least gain over the BM in the case of Top3 where 3 paths are available.

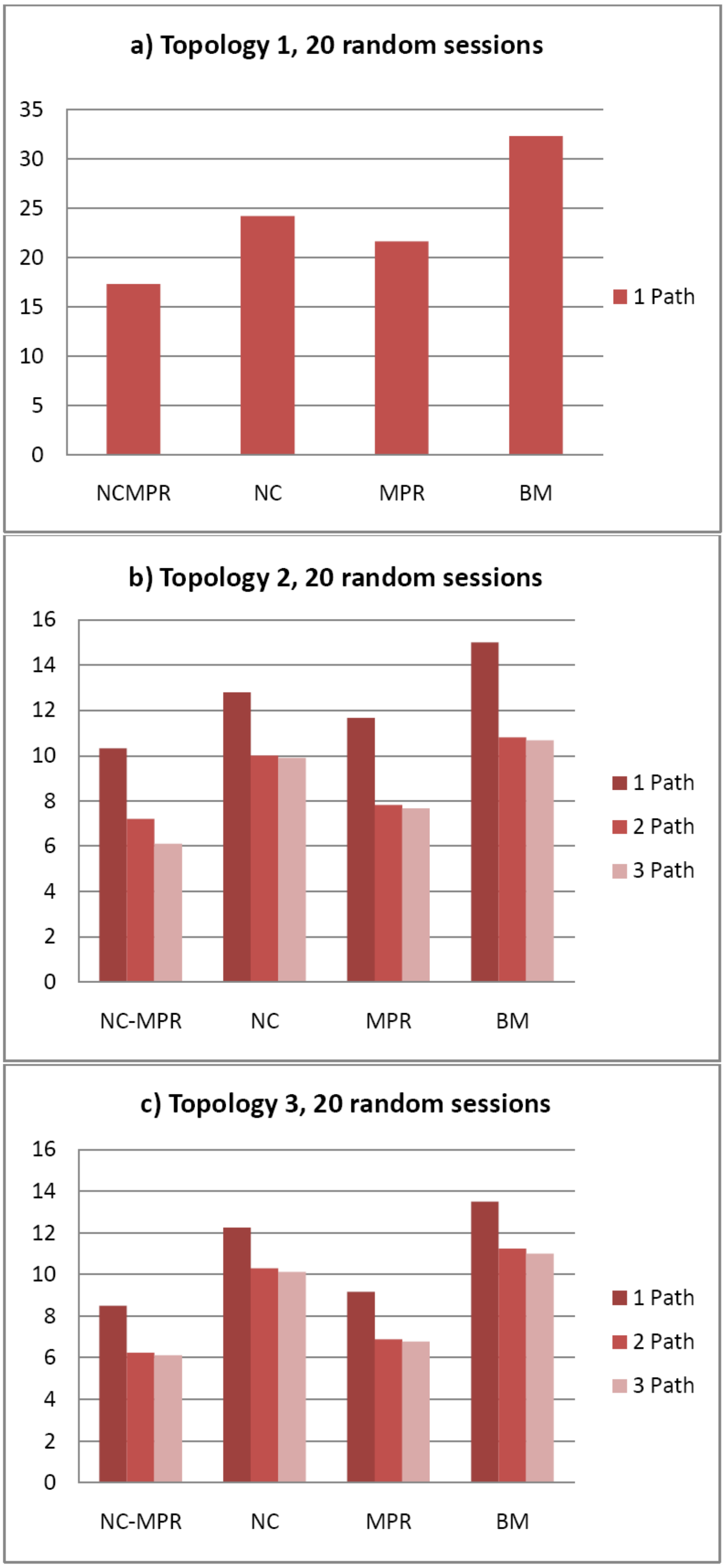


Figure 6.4: The averaged scheduling time to deliver 20 similar random sessions in Top1, Top2 and Top3, where at Top2 and Top3 1 to 3 paths are available for each traffic session.

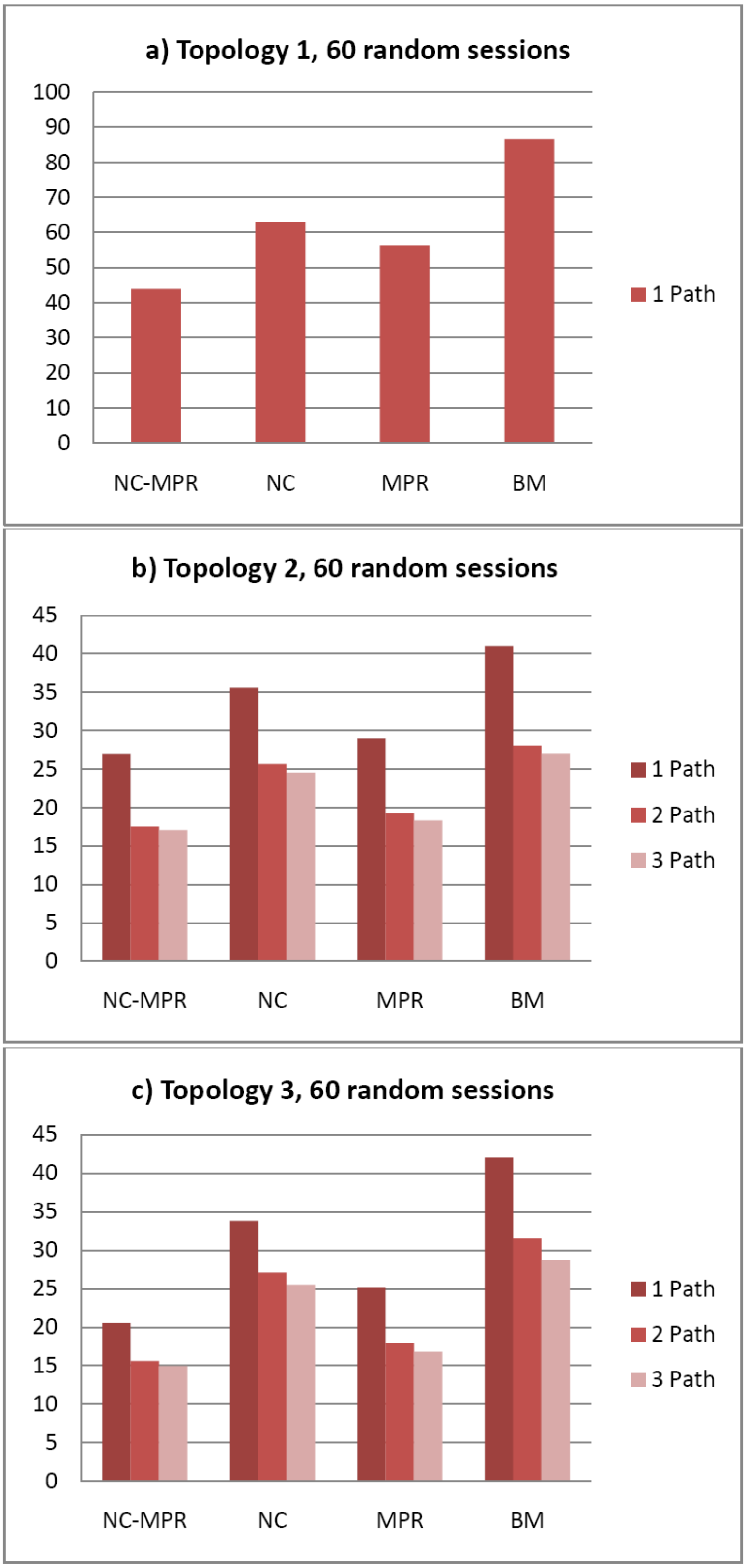


Figure 6.5: The averaged scheduling time to deliver 60 similar random sessions in Top1, Top2 and Top3, where at Top2 and Top3 1 to 3 paths are available for each traffic session.

	Packets	NC-MPR / BM		NC / BM		MPR / BM		NC-MPR / NC		NC-MPR / MPR	
		20	60	20	60	20	60	20	60	20	60
Top1	1-Path	46	49	25	27	33	35	28	30	20	21
Top2	1-Path	31	34	14	13	22	29	19	24	11	7
	2-Path	33	37	7	8	27	31	28	32	8	9
	3-Path	36	37	7	9	28	32	38	30	20	7
Top3	1-Path	37	51	9	19	32	40	30	39	7	18
	2-Path	44	50	8	14	38	43	39	42	9	13
	3-Path	44	48	8	11	38	43	40	41	10	11

Table 6.5: Relative gains (%) in Example II

However, in the case of NC-MPR method, the gain does not reduce due to the fact that different sessions are not required to be routed on the furthest path from each other for avoiding the interference; because SIC enables the nodes to exploit the interference.

6.2.3 Impact of load

Our numerical results reveal that NC-MPR, NC and MPR methods provide a higher gain over the BM when the network load is heavy. In the case of NC method, the network enjoys using the most of available coding components to route 60 sessions that better highlights the benefits of this method over the BM.

The MPR method also shows a higher gain because in routing more sessions (60 vs 20), we require more number of transmissions which increases the level of accumulated interference level. Therefore, the MPR model which is capable of exploiting the interference rather than avoiding it, shows superiority over the BM which uses a limited spatial reuse to deliver the demands.

The NC-MPR method performs the best because it is capable of both using more coding components (i.e., reducing the number of transmissions) and exploiting the interference.

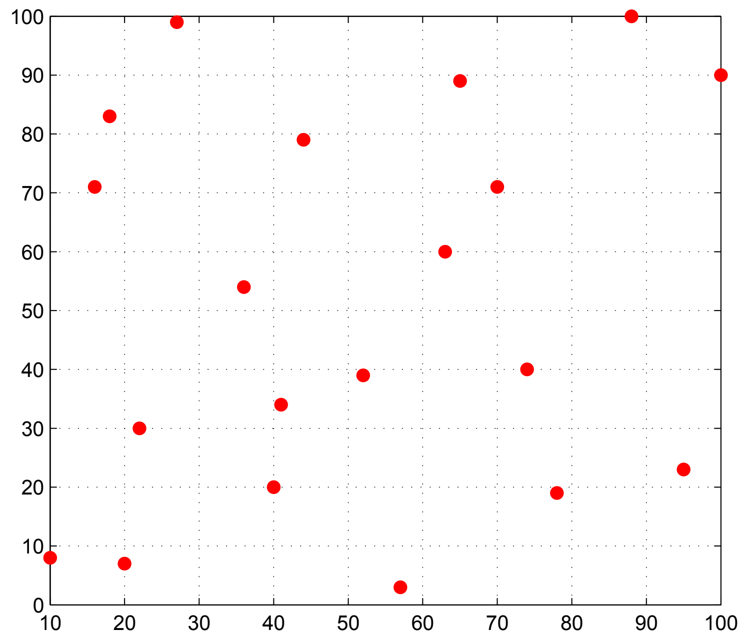


Figure 6.6: Placement of the nodes in Example III.

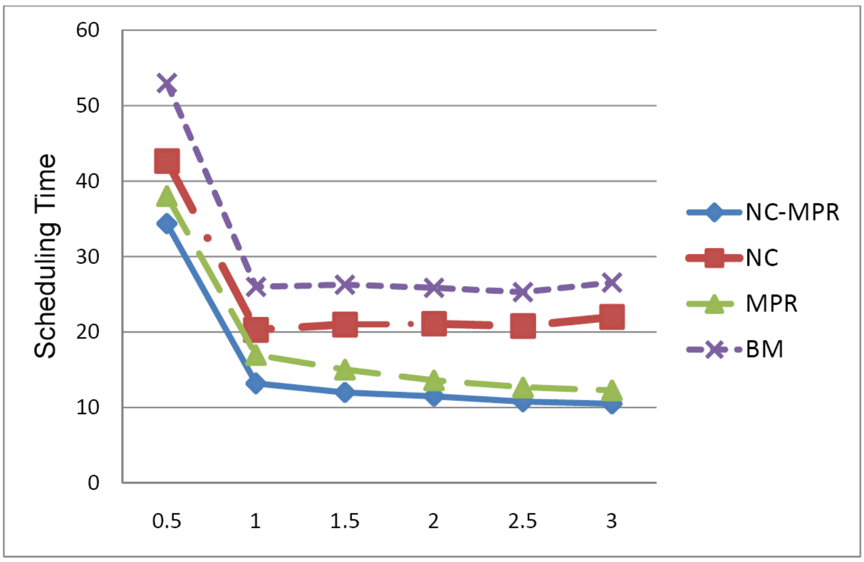


Figure 6.7: Delivery time (Seconds) vs transmission power (Miliwatts) in Example III.

6.3 Example III

We consider a network consisting of 20 nodes which are uniformly distributed over an area of $100m \times 100m$, shown in Figure 6.6. We have considered a medium-size load of 40 random sessions and evaluated the scheduling time by using each of the NC-MPR, NC, MPR and BM methods, where the transmission power varies. Our numerical results are shown in Figure 6.7. First, we obtain the numerical results for $P = 0.5mW$ which is the minimum required power to keep the network connected. Then we increase the power up to $3mW$ with a step of 0.5mW. Our numerical results show that in all of the methods, increasing the power from $0.5mW$ to $1mW$ significantly reduces the scheduling time. Because higher power brings more transmission links which increases the connectivity in the network.

However, we notice that the total delivery time remains almost constant in the NC and BM methods for $P > 1mW$. Although NC shows generally a better performance over BM, in both of these methods, the increased connectivity and hence more coding opportunities (only in NC method), and thus reduced transmission overhead will attempt to conquer the higher level of accumulated interference caused by increased power.

However, in the MPR and NC-MPR we observe that the scheduling time gradually decreases when the transmission power increases. Because SIC is capable of exploiting interference in these methods. Here, we note that if the number of interfering packets exceeds the multi-packet reception capability (k_n in equation (4.7)), MPR and NC-MPR methods will perform similar to the NC and BM, though with a higher relative gain. In this example, we obtain $k_n = 19$ at $P = 0.5mW$ which increases to 22 at $P = 3mW$. Therefore, k_n is very close to the nodes number which are the potential sources of the interference.

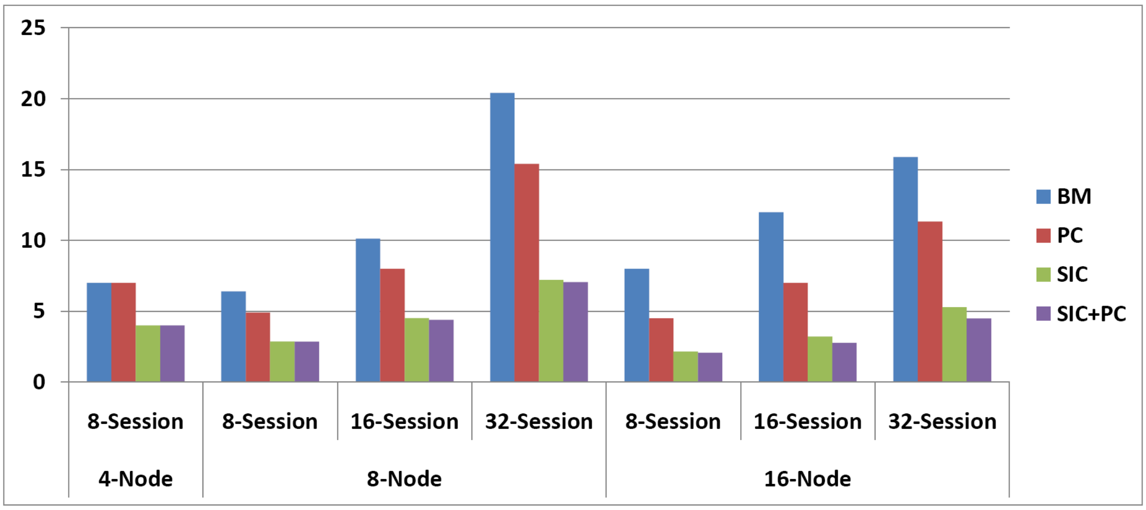


Figure 6.8: Scheduling time (Seconds) for BM, PC, SIC and SIC+PC models in three different networks.

6.4 Example IV

In this example, we consider 3 network topologies consisting of 4, 8 and 16 nodes which are uniformly distributed over an area of $100m \times 100m$. We have evaluated and compared the performance of four models, namely BM, PC (power control), SIC and SIC+PC. Figure 6.8 illustrates our numerical results for the minimum scheduling time for each case network instance and for different traffic load scenarios.

Figure 6.9 depicts the performance gains achieved by the joint routing and scheduling design model when SIC is deployed over the base model both with and without power control. Clearly, the figure shows substantial performance gains achieved when successive interference cancellation technology is used; although power control is an effective technique for managing the interference (especially in dense networks), SIC remains superior in terms of handling interference and improving the spectrum spatial reuse.

Next, we study the performance gain of the design model when network nodes are empowered with successive interference cancellation capabilities as well as power

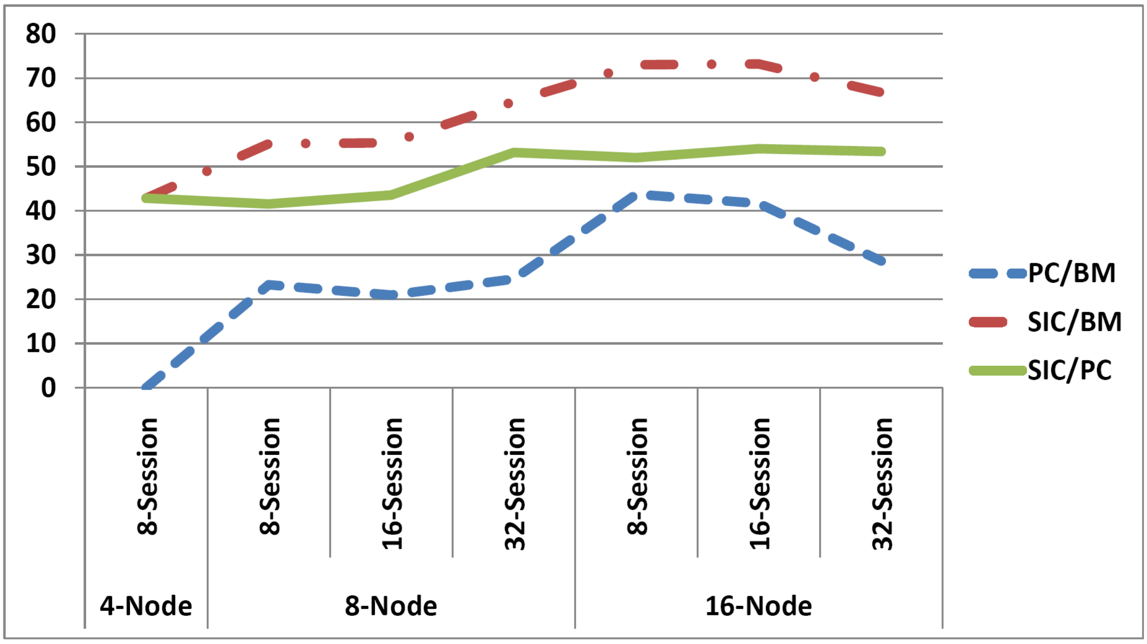


Figure 6.9: Reduction in scheduling time achieved by each of successive interference cancellation and power control techniques and their comparison.

control capabilities as another tool for interference management (the scheme is denoted as SIC+PC). The results are shown in Figure 6.10 where SIC+PC is compared with the base model (BM), with a power control aware model (PC) and with another model employing only successive interference cancellation (SIC). The objective of this study is to understand the performance benefits when both SIC and PC are concurrently deployed. As the figure shows, substantial gains are achieved over BM and PC models and a relatively better performance can be seen over SIC, especially as the network gets denser and higher traffic loads. Indeed, when the network is denser and more sessions are in the network, power control yields added value through its interference management capabilities and hence both interference cancellation and interference avoidance show improved network performance.

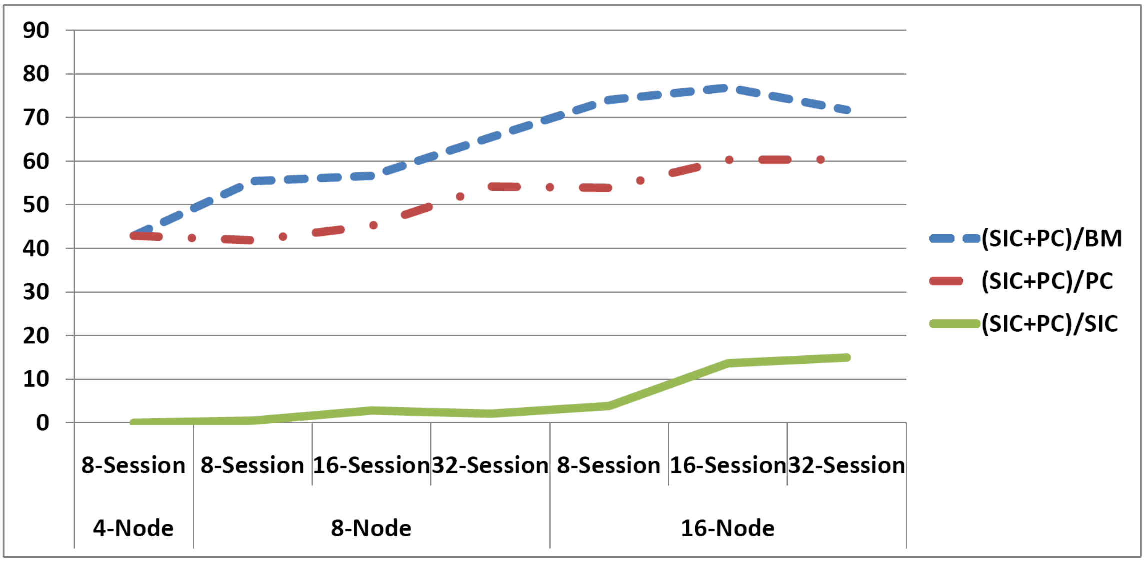


Figure 6.10: Reduction in scheduling time achieved by SIC+PC model over the BM, PC and SIC models.

6.5 Conclusion

In this chapter we evaluated the numerical results and compared the performance of four different joint routing and scheduling models, namely, BM, MPR, NC and NC-MPR. We showed that NC combined with SIC provides more than 50% reduction in scheduling time in a medium-size dense network. However, this reduction is around 40% and 19% when only SIC or NC are applied, respectively. Moreover, we investigated the impact of connectivity, path selection, loads and varying the transmission power on each of these models.

Next, we studied the benefits of SIC and power control and their combinations in different network scenarios. Our results showed that SIC always outperforms power control at least by 40%. We also demonstrated that combining power control with SIC provides almost 15% reduction in scheduling time when the network is dense and heavily loaded.

Chapter 7

Cross-Layer Optimization with Smart Antennas

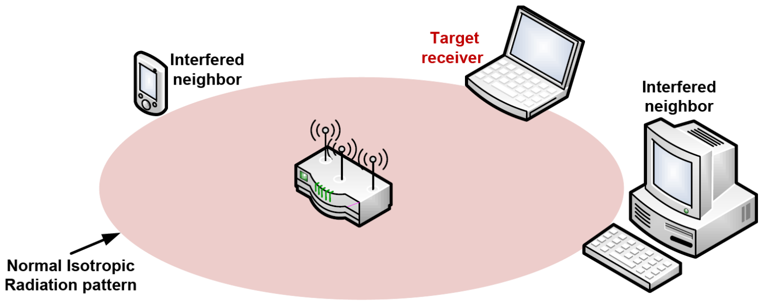
In this chapter, we present a cross-layer problem formulation which incorporates multi-path routing and link layer scheduling (Time Division Multiple Access) in multi-hop wireless networks with smart antenna nodes. We model this combinatorial complex problem as Mixed Integer Linear Program (MILP) and present a decomposition method based on Column Generation (CG) for solving it.

Our objective is to minimize the system activation time to satisfy certain traffic demands. The system activation time refers to the schedule length during which the links in the network should be active for supporting a certain traffic demand. In other words, given a certain traffic load (e.g., Mbits), our system activation time refers to the period of time (seconds) the network would take to deliver all the sessions. It should be noted that the lower bound on the system activation time is equivalent to the upper bound of the network throughput. To the best of our knowledge, this is the first decomposition framework for the design problem and obtaining exact solutions of wireless networks deploying smart antennas and addressing the benefits of the three technologies (BF, SDM, SDMA) in a cross-layer approach.

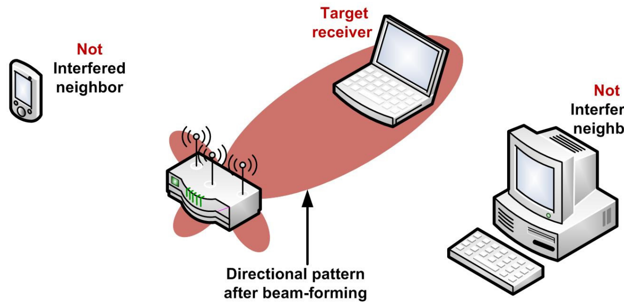
Another significant contribution is extending our design framework to consider the case of heterogeneous network. For the purpose of this thesis, heterogeneous network refers to a multi-antenna network with various number of antennas at each node. The proposed model of optimization is a multi-criteria MILP that jointly minimizes the system activation time and the cost of deployment. We minimize the cost of deployment by obtaining the minimum number of antennas required at each node.

7.1 Multiple Antennas Technology

The use of multiple antennas has been exploited as an effective solution to combat the effect of multipath fading in wireless communications [6], [37], [42]. Multipath fading occurs due to movement or presence of different objects in the environment which makes the signal scattered as it travels. In a rich scattering environment, like indoors or urban areas where there is no Line-of-Sight, the signal maybe strongly corrupted by interference. A solution to this problem can be achieved by utilizing smart or multiple antenna techniques. Smart antenna techniques include beam-forming, which provides directive gain, and Multiple-Input Multiple-Output (MIMO) methods having diversity and multiplexing gains. Directive gain combats the interference and increases the spatial reuse, diversity gain increases robustness to fading and improves the performance and multiplexing gain increases the transmission data rate. Multiple antennas in a network scenario can be used to provide multiple access gain through Spatial Division Multiple Access (SDMA) method, as well. It has been shown that there is a tradeoff between directive, diversity, multiplexing and multiple access gains in multi-antenna multi-user systems [116]. Three various multi-antenna techniques are described in the following sections.



(a)



(b)

Figure 7.1: (a)Omnidirectional transmission (b)Beam-forming at the transmitter side to provide directive gain.

7.1.1 Antenna Array Beam-forming Technique

In this technique, an array of multiple antennas in combination with appropriate signal processing is utilized to beam-form the signal [53] and provide directive gain. Beam-forming (BF) at the sender side implies sending the same signal from different antennas, where each antenna has an adjustable weight that determines the amplitude and the phase of the signal transmitted by that element. Similarly, this technique can be applied at the receiver side by adjusting different weights for each antenna element. Hence, with this technology, a transmitter can send the signal to a desired receiver and nullify it for another neighboring one. Figure 7.1 demonstrates an example of omnidirectional transmission versus transmission beam-forming. Alternatively, a receiver can suppress undesired signals (from interfering transmitters) and just receive the target signal. Such technique enabled by this technology is commonly referred to as interference suppression that allows for parallel transmissions

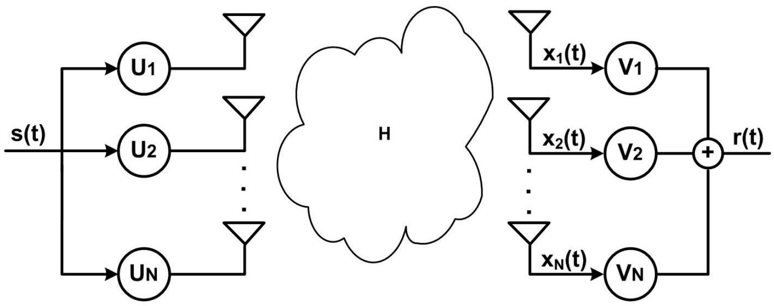


Figure 7.2: Multiple antennas at the transmitter and receiver sides of a system.

on neighboring links to be active concurrently. More number of concurrent active links result in a better spectrum spatial reuse which leads to capacity improvement in the network [126].

Figure 7.2 shows a link between two nodes endowed by N antennas each. H is a matrix with each of the entries representing the channel coefficients between any pair of antennas; therefore $h_{i,j}$ shows the channel coefficient between the i^{th} antenna at the transmitter and the j^{th} antenna at the receiver side. $U = [u_1, u_2...]$ and $V = [v_1, v_2...]$ are the weight vectors at both transmitter and receiver, respectively. The length of these vectors is equal to the number of antennas (i.e. N). The original signal $s(t)$ is to be transmitted from different antennas. After propagation, various replicas of the signal are received at different antennas, each is represented as:

$$x_j(t) = s(t) \sum_{i=1}^{N} u_i h_{i,j}$$

Then, the received signal $r(t)$ can be written as follows:

$$r(t) = \sum_{j=1}^{N} v_j x_j(t)$$

After some linear manipulations, these two equations can be combined and written in the form of:

$$r(t) = s(t) U^T H V$$

Note that knowing the channel coefficients is required for beam-forming the signal.

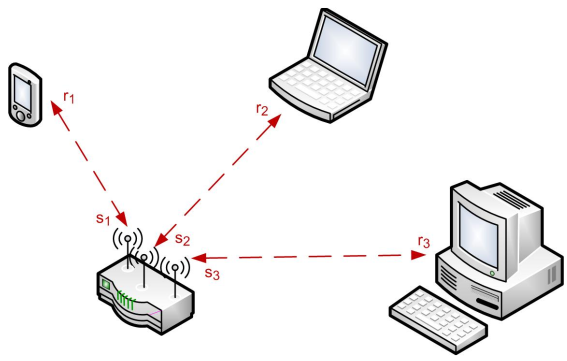


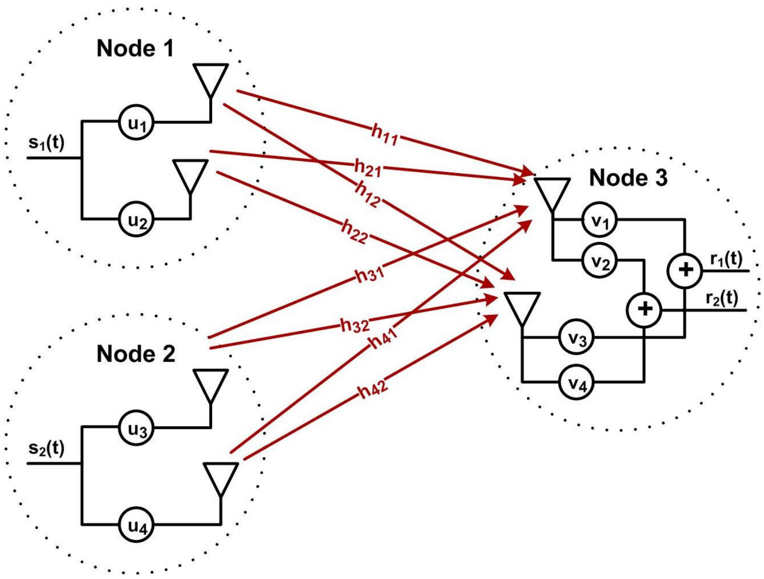
Figure 7.3: Multiple Access Spatial Division (SDMA)

These coefficients can be estimated by different methods such as sending pilot symbols [44]. It has been shown that, a more reliable Channel State Information (CSI) yields a higher throughput in the network [19]. By knowing the CSI and properly designing each of the matrices U and V, a signal can be suppressed or received with a desired gain at the receiver side. If $U^T HV = 1$ for a link, it means that the receiver node gets the transmitted signal perfectly. On the other hand, if $U^T HV = 0$, the corresponding node is not receiving any information about $s(t)$.

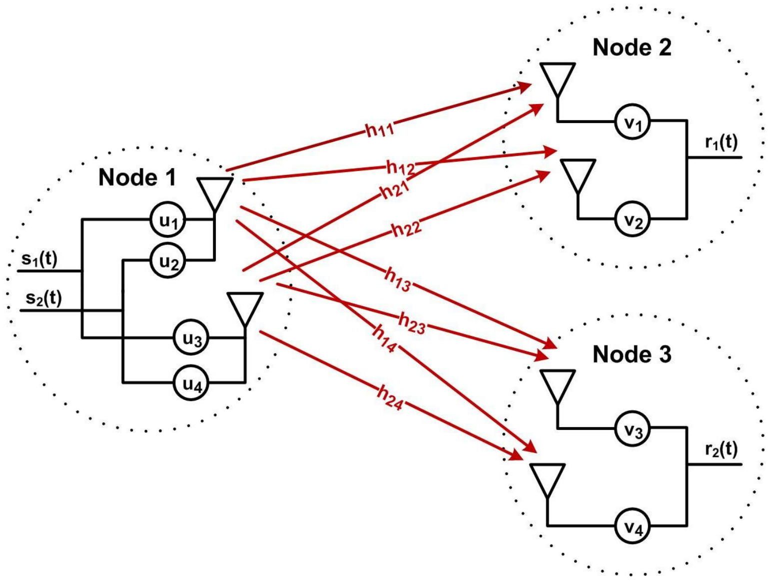
7.1.2 Spatial Division Multiple Access Technique

Figure 7.3 illustrates another multi-antenna technique in multi-user systems where the antenna elements at each node are exploited to communicate with different nodes concurrently. This technology is called Spatial Division Multiple Access (SDMA).

As shown in Figure 7.4, in this method also, each element of multi-antenna has an adjustable weight (u_is and v_is) and signal processing techniques are employed to adjust the antenna weights based on the available CSI, i.e. h_{ij}s. Hence, a node in the receiving mode can decode several data streams sent from various nodes simultaneously through using different antenna elements, which is referred to as Multiple Packet Reception (MPR). Figure 7.4(a) demonstrates an example of MPR.



(a) Multiple Packet Reception (MPR).



(b) Multiple Packet Transmission (MPT).

Figure 7.4: Spatial Division Multiple Access (SDMA) techniques: MPR and MPT

To determine the antenna weights, the following set of equation is solved.

$$\begin{bmatrix} u_1 & u_2 & 0 & 0 \\ 0 & 0 & u_3 & u_4 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{31} & h_{32} \\ h_{41} & h_{42} \end{bmatrix} \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This set contains 4 equations and 8 variables, $u_1 \cdots u_4$ and $v_1 \cdots v_4$. By assuming some values for 4 of these variables, the other four are obtained from the equations. Similarly, a node endowed with multiple antennas can send few data streams to individual nodes as well (Figure 7.4(b)), which is called Multiple Packet Transmission (MPT). Similar to MPR case, for obtaining the antenna weights in Figure 7.4(b), the following system must be solved.

$$\begin{bmatrix} u_1 & u_3 \\ u_2 & u_4 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \end{bmatrix} \begin{bmatrix} v_1 & 0 \\ v_2 & 0 \\ 0 & v_3 \\ 0 & v_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

SDMA provides multiple access gain that allows for transporting more packets in the network and consequently increases the throughput.

7.1.3 Spatial Division Multiplexing Technique

Spatial Division Multiplexing (SDM) is one of MIMO techniques that provides higher capacity for wireless communication channels. Layered Space-Time codes, also known as BLAST, are a set of Space-Time codes that brings the promised high data rates for the MIMO links [36]. In this scheme, if the system consists of N_T transmitter and N_R receiver antennas, depending on the type of the fading (fast or slow) in the channel, a maximum spatial multiplexing gain of $K = \min(N_T, N_R)$ can

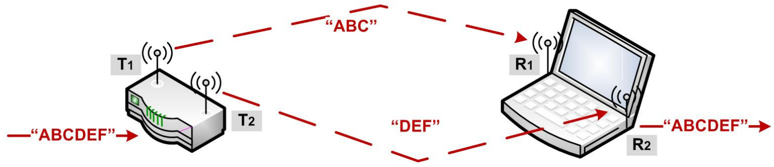


Figure 7.5: An example of Spatial Division Multiplexing (SDM) technique, communication between two double-antenna nodes.

be achieved. This gain is equal to the number of independent data streams transmitted from different antennas simultaneously. At the receiver side, the propagated signals are collected by the receiver antennas. Then, by employing powerful signal processing techniques like Zero-forcing and Minimum Mean Square Estimation combined with Successive Interference Cancelation (MMSE-SIC) [115], the transmitted signals are estimated. Sending and receiving K independent streams of signal at each time slot, yields a K-fold transmission data rate.

As an example, in Figure 7.5 the capacity of the link between two multi-antenna devices is doubled. First, the data "ABCDEF" is divided into two packets of data streams. Packets "ABC" and "DEF" are transmitted from antennas T_1 and T_2 where there is at least half-wavelength of space between them. Due to spatial diversity, these packets reach the destination through different paths and they are decoded at receiver's antennas R_1 and R_2, respectively. Later, the estimation of data streams are combined to form the estimation of the original data.

7.2 Problem statement

Consider a centralized WMN consisting of a number of static nodes, or access points, each equipped with multiple antennas. To achieve the maximum potential throughput of the network, a joint routing and scheduling is desired, where different multi-antenna techniques are applied and the throughput enhancement is studied. We present a mathematical optimization framework for the network design and formulate mathematically the joint routing and scheduling problem with the objective of minimizing the system activation time when the traffic load in the network does not change. It should be noted that minimizing the system activation time for transporting all demands in this steady-state network is equivalent to maximizing the throughput of the system. The scheduling in this network is based on Time Division Multiple Access (TDMA). In a TDMA-based MAC layer, time is divided into equal duration slots. At each time slot, a set of links can be active together without violating the requirement for successful communication. We assume the protocol model is used for modeling the interference in the network [47]. This model is explained in Section 3.1.2. In this work, multiple antennas installed at each node in the network are used to nullify the interference; either a sender nullifies its signal at the receiver of a neighboring link or a receiver suppresses the signal coming from the interfering transmitter. We consider three different models, all employing multiple antennas, as follows:

- BF Model: In this model all the antenna elements installed at each node are employed to suppress the existing interference through beam-forming (BF).
- BF+SDMA Model: In this model the antenna elements at each node are used either to cooperatively suppress the interference (BF) or to access multiple nodes (SDMA).
- BF+SDMA+SDM Model: In this model the antenna elements at each node are employed to provide interference suppression (BF), access multiple nodes (SDMA) or provide higher capacity for the links (SDM).

Figure 7.6 illustrates an example of a network consisting of 6 nodes and 20 links. In this example, the interference range and the transmitter range are assumed to be equal (i.e. $\delta = 0$) and each link is indexed by a number in the directed graph.

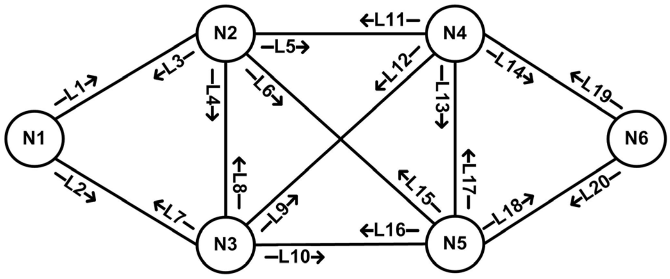


Figure 7.6: Illustrative Example

Suppose that the capacity of the links is normalized to one and we need to transport a demand of 2 units ($R = 2$) from N_1 to N_6. In the following paragraphs, the basic case without any multi-antenna technique and the above three models for the network in Figure 7.6 are explained. In each model, the optimal system activation time and the link schedule for routing the demand is obtained and demonstrated. We refer a configuration to a set of links which can be active concurrently in the same time slot.

Case I: Base Model

In this model, we consider nodes without BF, SDMA or SDM capabilities. This case is equivalent to a scenario where each node is endowed with a single antenna. Table 7.1 shows three configurations. At first, N_1 transmits the data to N_2 through L_1 where L_1 has to be active for 2 seconds. Then the demand is routed through links L_5 and L_{14} to reach N_6. Each of L_1, L_5 and L_{14} have one unit capacity and thus they must be active for 2 seconds to deliver the 2 unit demands. Thus, in the base model, only one link can be active at each configuration and no concurrent transmission is possible. Therefore, the system has to be active for a minimum of 6 seconds to deliver the 2 unit demands.

Table 7.1: Base model with single antenna

Configuration	Active Links	Activation Time (sec)
1	L_1	2
2	L_5	2
3	L_{14}	2
Total System Activation Time		6

Case II: BF Model

Suppose the number of antennas at each node is increased to 2 and beam-forming is performed to suppress the interference at/from neighboring links. For example, although the transmission on link L_{10} interferes with the reception on L_1, either N_2 can suppress this interference or N_3 can null its signal at N_2 by employing multi-antennas beam-forming. In this case, all demands are delivered in 2 seconds and two configurations are active each for 1 second; this yields a 3 times throughput improvement of that of case I. Table 7.2 shows that there can be up to 3 concurrent active links in each configuration.

Table 7.2: BF model with 2 antennas

Configuration	Active Links	Activation Time (sec)
1	L_1, L_{10}, L_{14}	1
2	L_2, L_5, L_{18}	1
Total System Activation Time		2

Case III: BF+SDMA Model

Here, we increase the number of antennas at each node to 3 and apply both beam-forming to nullify the interference and SDMA for multi-packet transmission and reception. We observe that, the system activation time is reduced to 1.5 seconds as shown in Table 7.3. As illustrated, in each configuration a node can transmit or receive on multiple links due to multiple access feature. For example, N_1 is transmitting multiple-packets on both L_1 and L_2 in the first configuration. The throughput in this model is 4 times better than case I since we have 4 simultaneous active links in each of the configurations.

Table 7.3: BF+SDMA model with 3 antennas

Configuration	Active Links	Activation Time (sec)
1	L_1, L_2, L_{14}, L_{18}	1
2	L_5, L_6, L_9, L_{10}	0.5
Total System Activation Time		1.5

Case IV: BF+SDMA+SDM Model

In this case, in addition to beam-forming and multiple access techniques, we employ spatial division multiplexing capability to the system to allow multiple data streams on each link. Assume 3 antennas are installed at each node. The minimum system activation time in this case is 1.3 seconds and the required configurations are shown in Table 7.4. In this table, the number of streams on each link is written in front of that link, in a bracket. As demonstrated in this table, in each configuration, a node can be active in more than one link connected to that node or it may be active in multi-streams on one link. Here, the throughput is 4.6154 times bigger than the single antenna case.

Table 7.4: BF+SDMA+SDM model with 3 antennas

Configuration	Active Links (Number of Data Streams)	Activation Time (sec)
1	$L_1(2), L_2(1), L_{14}(2)$	0.1
2	$L_1(1), L_2(1), L_{14}(1), L_{18}(2)$	0.4
3	$L_5(1), L_6(1), L_9(1), L_{10}(1)$	0.5
4	$L_1(2), L_2(1), L_{18}(2)$	0.1
5	$L_1(1), L_2(2), L_{14}(2)$	0.2
Total System Activation Time		1.3

However, if the number of antennas is increased to 4 as shown in Table 7.5, the overall system activation time is reduced to 0.9 seconds. In this model, we have up to 4 simultaneous active links in a configuration, where in each one, there is more than one active data stream.

Table 7.5: BF+SDMA+SDM model with 4 antennas

Configuration	Active Links (Number of Data Streams)	Activation Time (sec)
1	$L_1(2), L_2(2), L_{14}(2), L_{18}(2)$	0.5
2	$L_5(1), L_6(2), L_9(2)$	0.2
3	$L_6(2), L_9(2), L_{10}(1)$	0.2
Total System Activation Time		0.9

7.3 Mathematical Model

A multi-hop network can be modeled as a directed graph $G = (N, L)$ where N is the set of all nodes equipped with multiple antennas and L is the set of all feasible transmission links in the network. Therefore, $L = \{i,\ d_{t(i),r(i)} \leq T_{t(i)}, t(i) \in N, r(i) \in N, t(i) \neq r(i)\}$ where $t(i)$ and $r(i)$ are the transmitter and the receiver of link i, respectively; $T_{t(i)}$ is the transmission range of node $t(i)$. Here, p formally denotes one configuration. A binary variable that indicates whether a link is active in p or not is defined as:

$$v_i^p = \begin{cases} 1 & \text{if link } i \text{ is active in } p \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $p = \{v_i^p,\ \forall i \in L\}$. Next, we present the set of constraints required for formulating the design problem.

7.3.1 Radio Constraints

Let L_n be a subset of L that includes all the links connected to node n. $L_n^+ = \{i \in L : t(i) = n\}$ is the set of links whose transmitter is node n and $L_n^- = \{i \in L : r(i) = n\}$ is the set of all links whose receiver is node n. Let $L_n = \{L_n^+ \cup L_n^-\}$. The radio constraints that must be satisfied for the proper link scheduling in the network, are explained for each model as follows.

BF: Assuming a single radio per node, a node can either transmit or receive (but not both) in one configuration. Formally, this is written as:

$$\sum_{i \in L_n} v_i^p \leq 1 \qquad \forall n \in N \tag{7.1}$$

BF+SDMA: In this case also, only one radio is installed at each of the nodes, however, due to multiple access properties, each node is capable of being active in more than one link at each configuration. Nonetheless, all of the active links connected to a node have to be either in transmission or reception mode.

$$v_i^p + v_j^p \leq 1 \qquad \forall n \in N : i \in L_n^-,\ j \in L_n^+ \tag{7.2}$$

BF+SDMA+SDM: Similar to the BF+SDMA model, the single radio constraint is provided as:

$$v_i^p + v_j^p \leq 1 \qquad \forall n \in N : i \in L_n^-,\ j \in L_n^+ \tag{7.3}$$

Due to SDM capabilities, another set of constraints has to be considered in addition to what is presented for BF+SDMA model; we introduce a new integer variable z_i^p to denote the number of active streams on link i in configuration p. We also denote the number of allowed concurrent streams in the vicinity of node n by α_n and β_n, known as effective transmit and effective receive Degrees of Freedom (DoF), respectively. These degrees of freedom can be less than or equal to the number of antennas at the node [50]. When the system is functioning in SDM mode, a number of data streams can be transported independently between each pair of transmitter and receiver antennas. However, there is a maximum limit on the number of active streams on each link which is equal to the number of DoFs and has to be considered in radio constraints.

$$\begin{cases} z_i^p \leq \alpha_{t(i)}\ v_i^p \\ z_i^p \leq \beta_{r(i)}\ v_i^p \end{cases} \qquad \forall i \in L \tag{7.4}$$

The set of constraints in (7.4) ensures that no stream is to be transmitted over an inactive link.

7.3.2 Interference Constraints

The possible number of simultaneously active links in the network are limited by the level of interference in the network. For link scheduling, we make sure that two interfering links are not active concurrently. Define $\mathbf{C}$ as a set including all pairs of interfering links without any common nodes in between. At this point, two new sets are defined as: $C_i^+ = \{j \in L : (i,j) \in \mathbf{C}\}$, which is the set of all links where transmitting on node $t(i)$ interferes with the reception of the desired signal from node $t(j)$ at node $r(j)$; and $C_i^- = \{j \in L : (j,i) \in \mathbf{C}\}$, which is the set of all links such that transmitting on those links corrupts the reception on link i. The interference constraints for each of the models are explained below.

BF: Assume i and j are two active links and $(i,j) \in \mathbf{C}$. A full cooperation among nodes is assumed to avoid interference; either $t(i)$ is responsible to suppress its signal at $r(j)$, or $r(j)$ nullifies the signal coming from $t(i)$. With these assumptions, two new binary variables are defined as:

$$\gamma_{i,j}^p = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are active links in } p \text{ and } t(i) \text{ nullifies its signal at } r(j) \\ 0 & \text{otherwise.} \end{cases}$$

$$\mu_{i,j}^p = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are active links in } p \text{ and } r(j) \text{ suppresses interference of } t(i) \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, for each pair of interfering links (i,j) the following set of constraints should be satisfied to have a completely cooperative interference nullification.

$$\begin{cases} 1 + \sum_{k \in C_i^+} \gamma_{i,k}^p \leq \alpha_{t(i)} \\ 1 + \sum_{k \in C_j^-} \mu_{k,j}^p \leq \beta_{r(j)} \end{cases} \qquad \forall (i,j) \in \mathbf{C} \tag{7.5}$$

The group of constraints given in (7.6) ensure that there exist enough DoFs at each node to send/receive the data as well as to suppress the interference. They guarantee that a couple of interfering links can be active simultaneously only if, either the transmitter or the receiver of the interfering signal is capable of suppressing it.

$$v_i^p + v_j^p \leq \gamma_{i,j}^p + \mu_{i,j}^p + 1 \qquad \forall (i,j) \in \mathbf{C} \tag{7.6}$$

To ensure that variables γ and μ corresponding to an inactive link in configuration p are set to zero in the model and no DoF is assigned to them, we consider the following constraints:

$$\begin{cases} \gamma_{i,k}^p \leq v_i^p \\ \mu_{i,k}^p \leq v_i^p \end{cases} \qquad \forall (i,j) \in \mathbf{C} \tag{7.7}$$

Finally, the group of constraints in (7.8) enforce cooperation. For each two interfering links, if the transmitter of one is beam-forming the signal to cancel the interference, then the receiver of the other is prevented from using its DoFs for nullifying the interference coming from that specific transmitter. These constraints are however not considered in the model presented in [51].

$$\gamma_{i,j}^p + \mu_{i,j}^p \leq 1 \qquad \forall (i,j) \in \mathbf{C} \tag{7.8}$$

BF+SDMA: In this case, to have a completely cooperative interference nullification and SDMA at each node, the set of constraints in (7.9) should be satisfied.

$$\begin{cases} \sum_{i \in L_n^+} v_i^p + \sum_{k \in C_i^+} \gamma_{i,k}^p \leq \alpha_n \\ \sum_{j \in L_n^-} v_i^p + \sum_{k \in C_j^-} \mu_{k,j}^p \leq \beta_n \end{cases} \qquad \forall n \in N,\ \forall (i,j) \in \mathbf{C} \tag{7.9}$$

As the formula shows, due to SDMA capability, a node can be active in more than one link connected to it. Other interference constraints given below are similar to BF model, previously described.

$$v_i^p + v_j^p \leq \gamma_{i,j}^p + \mu_{i,j}^p + 1 \qquad \forall (i,j) \in \mathbf{C} \tag{7.10}$$

$$\begin{cases} \gamma_{i,k}^p \leq v_i^p \\ \mu_{i,k}^p \leq v_i^p \end{cases} \qquad \forall (i,j) \in \mathbf{C} \tag{7.11}$$

$$\gamma_{i,j}^p + \mu_{i,j}^p \leq 1 \qquad \forall (i,j) \in \mathbf{C} \tag{7.12}$$

BF+SDMA+SDM: In this case, when links are particularly active in SDM mode, no pair of interfering streams can be active simultaneously, unless the system provides interference suppression. If two streams on links i and j are interfering, either $t(i)$ is responsible to provide interference suppression at $r(j)$, or $r(j)$ suppresses the arriving signal from $t(i)$. Hence, we define two new integer variables: $\theta^p_{i,j}$, is the number of DoFs assigned by $t(i)$ to nullify its signal at $r(j)$ when there are active streams on both links i and j in configuration p; and $\phi^p_{i,j}$, which is the number of DoFs assigned by $r(j)$ to suppress $t(i)$'s signal at $r(j)$, when both links i and j are active in configuration p. Therefore, the constraints given in (7.13) should be satisfied to have a completely cooperative interference suppression capability where multiple streams are allowed on active links.

$$\begin{cases} \sum_{i \in L_n^+} z_i^p + \sum_{k \in C_i^+} \theta^p_{i,k} \leq \alpha_n \\ \sum_{j \in L_n^-} z_i^p + \sum_{k \in C_j^-} \phi^p_{k,j} \leq \beta_n \end{cases} \qquad \forall n \in N, \ \forall (i,j) \in \mathbf{C} \qquad (7.13)$$

The maximum number of streams on each active link is determined by considering the number of streams on interfering links, as shown in (7.14).

$$\begin{cases} z_i^p \leq \theta^p_{i,j} + \phi^p_{i,j} + min(\alpha_{t(i)}, \beta_{r(i)})(1 - v_j^p) \\ z_j^p \leq \phi^p_{i,j} + \theta^p_{i,j} + min(\alpha_{t(j)}, \beta_{r(j)})(1 - v_i^p) \end{cases} \qquad \forall (i,j) \in \mathbf{C} \qquad (7.14)$$

The set of inequalities in (7.15) make sure that no DOFs is assigned for suppressing the interference when the link is inactive.

$$\begin{cases} \theta^p_{i,j} \leq \alpha_{t(i)} \ v_i^p \\ \theta^p_{i,j} \leq \beta_{r(j)} \ v_j^p \\ \phi^p_{i,j} \leq \alpha_{t(i)} \ \ v_i^p \\ \phi^p_{i,j} \leq \beta_{r(j)} \ v_j^p \end{cases} \qquad \forall (i,j) \in \mathbf{C} \qquad (7.15)$$

The last interference constraints for BF+SDMA+SDM model given in (7.16) show a perfect cooperation between transmitters and receivers to prevent assigning extra

DoFs to nullify the interference of the same data stream.

$$\theta_{i,j}^{p} + \phi_{i,j}^{p} \leq z_{i}^{p} \qquad \forall (i,j) \in \mathbf{C} \tag{7.16}$$

7.3.3 Flow Conservation Constraints

Consider a set of M end-to-end traffic load in the network G. The m^{th} session is denoted by $S_m = \{(s_m, d_m, R_m) : s_m \in N, d_m \in N, R_m > 0, m = 1, ...M\}$. Therefore, in the m^{th} session, the source node s_m sends the commodity R_m to the destination node d_m. Denote f_i^m as the amount of commodity or the flow for the session m passing through link i. The flow-balance constraints make sure that for each of the relaying nodes, the sum of all incoming flow is equal to the sum of the outgoing flow in each session of traffic.

$$\sum_{i \in L_n^+} f_i^m = \sum_{j \in L_n^-} f_j^m \qquad \forall n \in N - \{s_m, d_m\} \tag{7.17}$$

For the source of each session, the sum of the outgoing flows should be equal to the amount of the traffic. Similarly, for the destination of a session, the sum of incoming flows is equal to the amount of the traffic. The mathematical representation of flow conservations for each session m is as follows.

$$\sum_{i \in L_{s_m}^+} f_i^m = \sum_{j \in L_{d_m}^-} f_j^m = R_m \qquad \forall\ s_m, \forall\ d_m \tag{7.18}$$

Moreover, in each session, the sum of the incoming flows to the source, as well as the sum of the outgoing flows from the destination should be zero.

$$\sum_{i \in L_{s_m}^-} f_i^m = \sum_{j \in L_{d_m}^+} f_j^m = 0 \qquad \forall\ s_m, \forall\ d_m \tag{7.19}$$

7.3.4 Bandwidth Constraints

Due to the limit on the available capacity for each link in the network, bandwidth constraints are introduced when performing routing. Denote the fraction of time that a configuration p is active, by λ_p, and denote the capacity of link i, by c_i. Let $\mathcal{P}$ be set of all possible configurations and denote the number of such configurations by $\bar{P}$. Then, the bandwidth constraints is written as:

$$\sum_{p=1}^{\bar{P}} \lambda_p c_i v_i^p - \sum_{m=1}^{M} f_i^m \geq 0 \qquad \forall i \in L$$

These constraints guarantee that the sum of the flows passing over a link belonging to different sessions does not exceed the capacity of that link.

BF and **BF+SDMA:** By normalizing the capacity, $c_i = 1$, the bandwidth constraint above is changed to the following constraint.

$$\sum_{p=1}^{\bar{P}} \lambda_p v_i^p - \sum_{m=1}^{M} f_i^m \geq 0 \qquad \forall i \in L \tag{7.20}$$

BF+SDMA+SDM: Denote the capacity of Single-Input Single-Output (SISO) link i in the bandwidth constraint, by c_i^{SISO}. When z_i^p independent streams are transported through a MIMO link, the capacity is multiplied by the integer number z_i^p, i.e. $c_i^{\text{MIMO}} = z_i^p \ c_i^{\text{SISO}}$. By normalizing the capacity of SISO channel, $c_i^{\text{SISO}} = 1$, bandwidth constraints for this case are given as:

$$\sum_{p=1}^{\bar{P}} \lambda_p z_i^p - \sum_{m=1}^{M} f_i^m \geq 0 \qquad \forall i \in L \tag{7.21}$$

7.4 Objective and Methodology

Our objective is to maximize the throughput of the network by minimizing the overall system activation time, which is represented as:

$$\min \sum_{p=1}^{\bar{P}} \lambda_p \tag{7.22}$$

where $\bar{P}$ is the total number of configurations. As noted before, the minimization is done through a joint routing and scheduling model, considering all the radio, interference, flow conservation and bandwidth constraints given in (7.1)-(7.21) for different models. In this optimization, all the configurations in $\mathcal{P}$ should be generated and the routing problem should be solved over all of them to yield the optimal solution. However, the number of the configurations grows exponentially with the size of the network, number of antennas per node and other parameters. Obviously, enumerating all such configurations is not practical. To obtain the optimal solution without enumerating all configurations, the problem is decomposed into smaller subproblems using a Column Generation (CG) approach [26]. CG decomposition is a very powerful method to get the optimal solution of large sized linear problems. This approach has various applications in economy, management and different other fields. There are also some research in communication area using CG [13, 17, 63, 67, 138]. In CG, the problem is divided into two parts: Master sub-problem and Pricing sub-problem. More details in solving the problem by CG technique are explained in the following paragraphs.

7.4.1 Master and Pricing sub-problems

The routing of demands is determined in the Master sub-problem. The objective of the Master sub-problem is the same as the objective of the original problem and the constraints are flow conservations (7.17)-(7.19) and bandwidth constraints (7.20) for BF or BF+SDMA model, and (7.21) for BF+SDMA+SDM model. In the first iteration, a set of initial feasible configurations, $\mathcal{P}_0 \subseteq \mathcal{P}$ is available to solve the first instance of the Master sub-problem. It should be noted that v_i^p is not a variable in the Master. The solution of the Master is the best routing over all possible configurations ($\mathcal{P}_0$), which is a local optimum. When the Master obtains the local optimal solution, it generates the dual values of the bandwidth constraints, as $\{\bar{y}_i,\ i \in L\}$ and sends them to the Pricing to find the best link scheduling. Then the Pricing sub-problem applies these dual values to construct the Pricing objective

Table 7.6: Column Generation model of BF case

[Master sub-problem] :

$$\min \sum_{p\in\mathcal{P}_0} \lambda_p$$

subject to

$$\sum_{i\in L_n^+} f_i^m = \sum_{j\in L_n^-} f_j^m \qquad \forall n \in N - \{s_m, d_m\}$$

$$\sum_{i\in L_{s_m}^+} f_i^m = \sum_{j\in L_{d_m}^-} f_j^m = R_m \qquad \forall s_m, d_m$$

$$\sum_{i\in L_{s_m}^-} f_i^m = \sum_{j\in L_{d_m}^+} f_j^m = 0 \qquad \forall s_m, d_m$$

$$\sum_{p=1}^{\bar{P}} \lambda_p v_i^p - \sum_{m=1}^{M} f_i^m \geq 0 \qquad \forall i \in L$$

[Pricing sub-problem] :

$$\min(1 - \sum_{i\in L} \bar{y}_i v_i^p)$$

subject to

$$\sum_{i\in L_n} v_i^p \leq 1 \qquad \forall n \in N$$

$$\begin{cases} 1 + \sum_{k\in C_i^+} \gamma_{i,k}^p \leq \alpha_{t(i)} \\ 1 + \sum_{k\in C_j^-} \mu_{k,j}^p \leq \beta_{r(j)} \end{cases} \qquad \forall (i,j) \in \mathbf{C}$$

$$v_i^p + v_j^p \leq \gamma_{i,j}^p + \mu_{i,j}^p + 1 \qquad \forall (i,j) \in \mathbf{C}$$

$$\begin{cases} \gamma_{i,k}^p \leq v_i^p \\ \mu_{i,k}^p \leq v_i^p \end{cases} \qquad \forall (i,j) \in \mathbf{C}$$

$$\gamma_{i,j}^p + \mu_{i,j}^p \leq 1 \qquad \forall (i,j) \in \mathbf{C}$$

as:

$$\min(1 - \sum_{i\in L} \bar{y}_i v_i^p) \tag{7.23}$$

where $(1 - \sum_{i\in L} \bar{y}_i v_i^p)$ is called the reduced cost. The constraints of the Pricing are those given in (7.1)-(7.16).

From now on, in each iteration, a feasible configuration is generated in the Pricing sub-problem. The value of the Pricing objective is always checked to determine the optimality of the solution. If (7.23) is a non-negative value, the obtained solution in the Master sub-problem is the optimal solution to the main problem. In this case, generating more configurations in the Pricing would not provide further improvement because the optimal solution has already been found. A proof of optimality of the solution is given in [26]. On the other hand, if (7.23) has a negative value, the Pricing generates another configuration named p and sends it to the Master sub-problem. Then, the Master adds p into the set of previous configurations, $\mathcal{P}_0 \biguplus p \rightarrow \mathcal{P}_0$, and resolves the routing problem. Whenever a solution is produced by the Master, the corresponding dual values are passed to the Pricing and the procedure continues until finding the optimal solution.

The CG model of BF, BF+SDMA and BF+SDMA+SDM are presented in Tables 7.6, 7.7 and 7.8, respectively.

Table 7.7: Column Generation model of BF+SDMA case

[Master sub-problem] :

$$\min \sum_{p \in \mathcal{P}_0} \lambda_p$$

subject to

$$\sum_{i \in L_n^+} f_i^m = \sum_{j \in L_n^-} f_j^m \qquad \forall n \in N - \{s_m, d_m\}$$

$$\sum_{i \in L_{s_m}^+} f_i^m = \sum_{j \in L_{d_m}^-} f_j^m = R_m \qquad \forall s_m, d_m$$

$$\sum_{i \in L_{s_m}^-} f_i^m = \sum_{j \in L_{d_m}^+} f_j^m = 0 \qquad \forall s_m, d_m$$

$$\sum_{p=1}^{\bar{P}} \lambda_p v_i^p - \sum_{m=1}^{M} f_i^m \geq 0 \qquad \forall i \in L$$

[Pricing sub-problem] :

$$\min(1 - \sum_{i \in L} \bar{y}_i v_i^p)$$

subject to

$$v_i^p + v_j^p \leq 1 \qquad \forall n \in N : i \in L_n^-,\ j \in L_n^+$$

$$\begin{cases} \sum_{i \in L_n^+} v_i^p + \sum_{k \in C_i^+} \gamma_{i,k}^p \leq \alpha_n \\ \sum_{j \in L_n^-} v_i^p + \sum_{k \in C_j^-} \mu_{k,j}^p \leq \beta_n \end{cases} \qquad \forall n \in N, \quad \forall (i,j) \in \mathbf{C}$$

$$v_i^p + v_j^p \leq \gamma_{i,j}^p + \mu_{i,j}^p + 1 \qquad \forall (i,j) \in \mathbf{C}$$

$$\begin{cases} \gamma_{i,k}^p \leq v_i^p \\ \mu_{k,i}^p \leq v_i^p \end{cases} \qquad \forall (i,j) \in \mathbf{C}$$

$$\gamma_{i,j}^p + \mu_{i,j}^p \leq 1 \qquad \forall (i,j) \in \mathbf{C}$$

Table 7.8: Column Generation model of BF+SDMA+SDM case

[Master sub-problem] :

$$\min \sum_{p \in \mathcal{P}_0} \lambda_p$$

subject to

$$\begin{aligned}
&\textstyle\sum_{i\in L_n^+} f_i^m = \sum_{j\in L_n^-} f_j^m && \forall n \in N - \{s_m, d_m\}\\
&\textstyle\sum_{i\in L_{s_m}^+} f_i^m = \sum_{j\in L_{d_m}^-} f_j^m = R_m && \forall s_m, d_m\\
&\textstyle\sum_{i\in L_{s_m}^-} f_i^m = \sum_{j\in L_{d_m}^+} f_j^m = 0 && \forall s_m, d_m\\
&\textstyle\sum_{p=1}^{\bar{P}} \lambda_p z_i^p - \sum_{m=1}^{M} f_i^m \geq 0 && \forall i \in L
\end{aligned}$$

[Pricing sub-problem] :

$$\min(1 - \sum_{i \in L} \bar{y}_i z_i^p)$$

subject to

$$\begin{aligned}
&v_i^p + v_j^p \leq 1 && \forall n \in N : i \in L_n^-,\ j \in L_n^+\\
&\begin{cases} z_i^p \leq \alpha_{t(i)}\ v_i^p \\ z_i^p \leq \beta_{r(i)}\ v_i^p \end{cases} && \forall i \in L\\
&\begin{cases} \sum_{i\in L_n^+} z_i^p + \sum_{k\in C_i^+} \theta_{i,k}^p \leq \alpha_n \\ \sum_{j\in L_n^-} z_i^p + \sum_{k\in C_j^-} \phi_{k,j}^p \leq \beta_n \end{cases} && \forall n \in N\\
&\begin{cases} z_i^p \leq \theta_{i,j}^p + \phi_{i,j}^p + min(\alpha_{t(i)}, \beta_{r(i)})(1 - v_j^p) \\ z_j^p \leq \phi_{i,j}^p + \theta_{i,j}^p + min(\alpha_{t(j)}, \beta_{r(j)})(1 - v_i^p) \end{cases} && \forall (i,j) \in \mathbf{C}\\
&\begin{cases} \theta_{i,j}^p \leq \alpha_{t(i)}\ v_i^p \\ \theta_{i,j}^p \leq \beta_{r(j)}\ v_j^p \\ \phi_{i,j}^p \leq \alpha_{t(i)}\ v_i^p \\ \phi_{i,j}^p \leq \beta_{r(j)}\ v_j^p \end{cases} && \forall (i,j) \in \mathbf{C}\\
&\theta_{i,j}^p + \phi_{i,j}^p \leq z_i^p && \forall (i,j) \in \mathbf{C}
\end{aligned}$$

7.5 The Case of Heterogeneous Network

The framework we presented in the previous section aims at minimizing the system activation time in a multi-antenna wireless mesh network where beam-forming (BF), spatial division multiple access (SDMA) and spatial division multiplexing (SDM) techniques are employed. The optimal solution for a homogeneous network in cases of different number of antennas were obtained through column generation approach and is demonstrated in Chapter 8. Now, let us investigate what happens if the network is heterogeneous.

Consider the problem of designing a network of access points with the objective of maximizing the throughput and minimizing the cost of deployment. In this case, the optimization problem is multi-criteria with two objectives: minimizing the system activation time, ($\min \sum_{p \in \mathcal{P}} \lambda_p$) and minimizing the total number of antennas installed at each node, ($\min \sum_{\forall n \in N} \alpha_n$).

In the case where the number of antennas is not predetermined, we introduce the set of *antenna constraints* as:

$$\begin{cases} \alpha_n = \beta_n \\ 1 \leq \alpha_n \leq DoF_{max} \\ 1 \leq \beta_n \leq DoF_{max} \end{cases} \qquad \forall n \in \mathbf{N} \tag{7.24}$$

where DoF_{max} depends on the maximum number of antennas available in manufactured access points. Moreover, constraints (7.4), (7.14) and (7.15) become non-linear. Dealing with multi-criteria optimization has been explained in section 2.2.4 and linearization of the constraints is explained next.

7.5.1 Linearizing the Constraints

To linearize constraints (7.4), (7.14) and (7.15) which contain multiplication of two variables, one binary (v) and the other integer (α and β), we introduce two new variables, $\alpha v(i,j) = \alpha_{t(i)} v_j^p$ and $\beta v(i,j) = \beta_{r(i)} v_j^p$, along with six set of constraints, as follows.

$$\begin{cases} \alpha v(i,j) \leq \alpha_{t(i)} + (1 - v_j^p)M \\ \alpha v(i,j) \geq \alpha_{t(i)} - (1 - v_j^p)M \\ \alpha v(i,j) \leq M v_j^p \end{cases} \quad \forall (i,j) \in \mathbf{C}, \text{M is a big number} \tag{7.25}$$

$$\begin{cases} \beta v(i,j) \leq \beta_{r(i)} + (1 - v_j^p)M \\ \beta v(i,j) \geq \beta_{r(i)} - (1 - v_j^p)M \\ \beta v(i,j) \leq M v_j^p \end{cases} \quad \forall (i,j) \in \mathbf{C}, \text{M is a big number} \tag{7.26}$$

Then, equation (7.4) is written as:

$$\begin{cases} z_i^p \leq \alpha v(i,j) \\ z_i^p \leq \beta v(i,j) \end{cases} \quad \forall i \in L \tag{7.27}$$

Equation (7.14) changes to:

$$\begin{cases} z_i^p \leq \theta_{i,j}^p + \phi_{i,j}^p + min(\alpha_{t(i)} - \alpha v(i,j)), (\beta_{r(i)} - \beta v(i,j)) \\ z_i^p \leq \theta_{i,j}^p + \phi_{i,j}^p + min(\alpha_{t(j)} - \alpha v(j,i)), (\beta_{rj)} - \beta v(j,i)) \end{cases} \quad \forall (i,j) \in \mathbf{C} \tag{7.28}$$

And (7.15) will be modified to:

$$\begin{cases} \theta_{i,j}^p \leq \alpha v(i,j) \\ \theta_{i,j}^p \leq \beta v(i,j) \\ \phi_{i,j}^p \leq \alpha v(i,j) \\ \phi_{i,j}^p \leq \beta v(i,j) \end{cases} \quad \forall (i,j) \in \mathbf{C} \tag{7.29}$$

7.5.2 Linear Multi-criteria Column Generation Model

In the CG model of heterogeneous network, the Master sub-problem remains with a single objective of minimizing the system activation time and we include the second objective, i.e. minimizing the number of antennas, in the Pricing sub-problem. Therefore, the Pricing is a multi-criteria optimization; minimizing the reduced cost

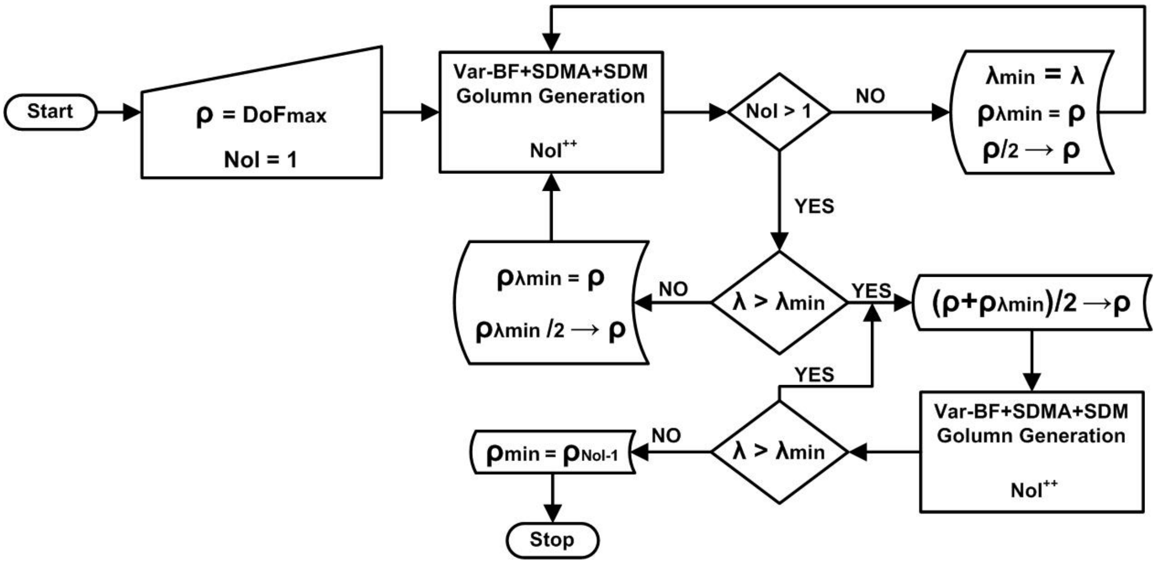


Figure 7.7: The flowchart of multi-criteria optimization model

is kept as the main objective and the other objective function (which is minimizing the total number of antennas in the network) is considered as a constraint:

$$\sum_{\forall n \in N} \alpha_n \leq \text{number of nodes} * \rho \tag{7.30}$$

By varying the value of ρ ($1 \leq \rho \leq DoF_{max}$), optimal solutions over different parts of the feasible region can be obtained. Table 7.9 shows the decomposition of our multi-criteria optimization problem namely HTRGN-BF+SDMA+SDM model and Figure 7.7 demonstrates our proposed algorithm for obtaining the minimum system activation time while minimizing the number of antennas.

Table 7.9: Column Generation Model of HTRGN-BF+SDMA+SDM

[Master sub-problem] :

$$\min \lambda = \sum_{p \in \mathcal{P}_0} \lambda_p$$

subject to

$\sum_{i \in L_n^+} f_i^m = \sum_{j \in L_n^-} f_j^m$	$\forall n \in N - \{s_m, d_m\}$
$\sum_{i \in L_{s_m}^+} f_i^m = \sum_{j \in L_{d_m}^-} f_j^m = R_m$	$\forall s_m, d_m$
$\sum_{i \in L_{s_m}^-} f_i^m = \sum_{j \in L_{d_m}^+} f_j^m = 0$	$\forall s_m, d_m$
$\sum_{p=1}^{\bar{P}} \lambda_p z_i^p - \sum_{m=1}^{M} f_i^m \geq 0$	$\forall i \in L$

[Pricing sub-problem] :

$$\min(1 - \sum_{i \in L} \bar{y}_i z_i^p)$$

subject to

$\sum_{\forall n \in N} \alpha_n \leq \text{number of nodes} * \rho$	$\rho \leq DoF_{max}$
$\alpha_n = \beta_n$	$\forall n \in N$
$v_i^p + v_j^p \leq 1$	$\forall n \in N : i \in L_n^-,\ j \in L_n^+$
$\begin{cases} \alpha v(i,j) \leq \alpha_{t(i)} + (1 - v_j^p)M \\ \alpha v(i,j) \geq \alpha_{t(i)} - (1 - v_j^p)M \\ \alpha v(i,j) \leq M v_j^p \end{cases}$	$\forall (i,j) \in \mathbf{C}, \quad \text{M}= 10^6$
$\begin{cases} \beta v(i,j) \leq \beta_{r(i)} + (1 - v_j^p)M \\ \beta v(i,j) \geq \beta_{r(i)} - (1 - v_j^p)M \\ \beta v(i,j) \leq M v_j^p \end{cases}$	$\forall (i,j) \in \mathbf{C}, \quad \text{M}= 10^6$
$\begin{cases} z_i^p \leq \alpha v(i,j) \\ z_i^p \leq \beta v(i,j) \end{cases}$	$\forall i \in L$
$\begin{cases} \sum_{i \in L_n^+} z_i^p + \sum_{k \in C_i^+} \theta_{i,k}^p \leq \alpha_n \\ \sum_{j \in L_n^-} z_i^p + \sum_{k \in C_j^-} \phi_{k,j}^p \leq \beta_n \end{cases}$	$\forall n \in N$
$\begin{cases} z_i^p \leq \theta_{i,j}^p + \phi_{i,j}^p + min(\alpha_{t(i)} - \alpha v(i,j)), (\beta_{r(i)} - \beta v(i,j)) \\ z_i^p \leq \theta_{i,j}^p + \phi_{i,j}^p + min(\alpha_{t(j)} - \alpha v(j,i)), (\beta_{rj)} - \beta v(j,i)) \end{cases}$	$\forall (i,j) \in \mathbf{C}$
$\begin{cases} \theta_{i,j}^p \leq \alpha v(i,j) \\ \theta_{i,j}^p \leq \beta v(i,j) \\ \phi_{i,j}^p \leq \alpha v(i,j) \\ \phi_{i,j}^p \leq \beta v(i,j) \end{cases}$	$\forall (i,j) \in \mathbf{C}$
$\theta_{i,j}^p + \phi_{i,j}^p \leq z_i^p$	$\forall (i,j) \in \mathbf{C}$

In the first iteration of solving the problem, $NoI = 1$ and ρ is set to DoF_{max}. Then, the minimum system activation time, λ_{min}, is obtained and saved. λ_{min} shows the individual optimal value of the system activation time in the pay off table. In the second iteration, ρ is reduced to $\frac{\rho}{2}$ and λ is re-computed. As long as $\lambda = \lambda_{min}$, the algorithm saves the value of ρ as $\rho_{\lambda_{min}}$ and continues decreasing ρ to its half in the next iteration. However, if the optimal system activation time in one iteration is found more than λ_{min}, the algorithm starts increasing ρ to $\frac{\rho+\rho_{min}}{2}$. The program terminates whenever the system activation time of λ_{min} is obtained.

7.6 Conclusions

In this chapter, at first a mathematical framework for cross-layer design of Wireless Mesh Networks was provided where multiple antennas at each node are employed to suppress the interference from neighboring transmissions (BF), to access multiple nodes simultaneously (SDMA) and to transmit with higher data rates (SDM). By integrating different combinations of multi-antenna techniques in physical layer with various constraints from MAC and network layers, three MILP models were presented. The objective of these models was to minimize the system activation time. Due to the large number of constraints and variables in these models, the problem was decomposed using Column Generation approach which made each model more scalable.

In the second part, an optimization model for a heterogeneous network was introduced to minimize the deployment cost. The provided mathematical model was a multi-criteria optimization with two objectives of providing the minimum system activation time and the minimum antenna elements per node where BF, SDMA and SDM techniques are available. We defer our numerical results to be presented in Chapter 8.

Chapter 8

Numerical Results of BF,SDMA and SDM

In this chapter, the numerical results of the cross-layer optimization are provided by implementing the MILP models and solving them by using CPLEX Concert Technology adapted for C^{++} [30]. One objective of this section is to evaluate the performance of BF, BF+SDMA and BF+SDMA+SDM models with respect to the number of antennas, transmission ranges, node densities and traffic load. We also investigate the scalability of these models. Moreover, the performance of HTRGN-BF+SDMA+SDM model will be evaluated for various traffic load. To accomplish these objectives we have considered two scenarios, the first scenario is over a $100m \times 100m$ area, and the other one is over a $250m \times 250m$ area. For both scenarios, we assume the interference range coefficient, δ, to be 0.3 as discussed in Section 7.2. To obtain the numerical results in each scenario, the minimum system activation time is attained. The minimum system activation time for transporting all the traffic load in the network corresponds to the maximum throughput of the system.

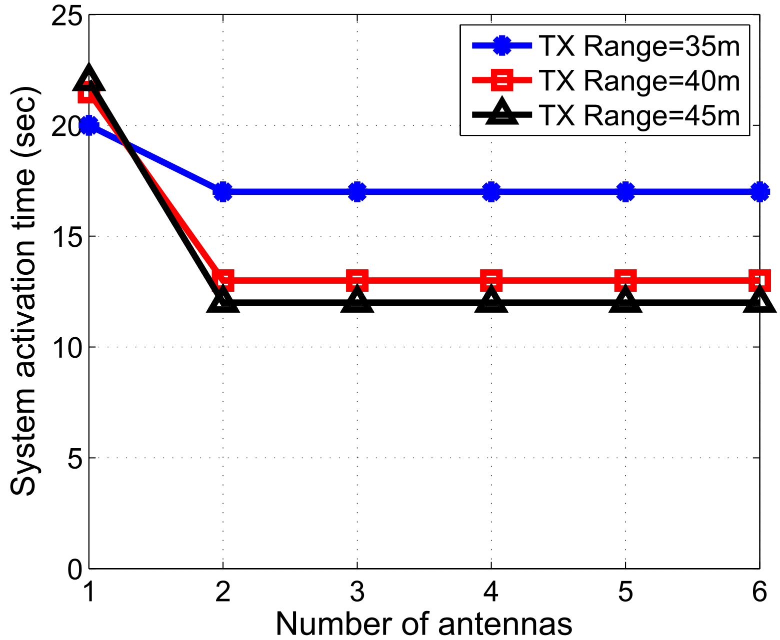


Figure 8.1: The effect of transmission range on the system activation time in a network consisting of 10 nodes with multi-antenna in BF model where 15 sessions are delivered.

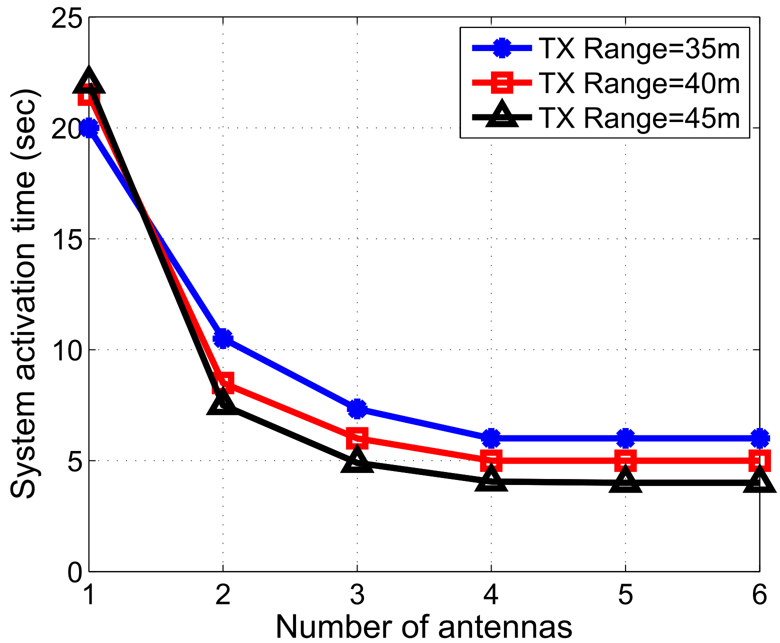


Figure 8.2: The effect of transmission range on the system activation time in a network consisting of 10 nodes with multi-antenna in BF+SDMA model where 15 sessions are delivered.

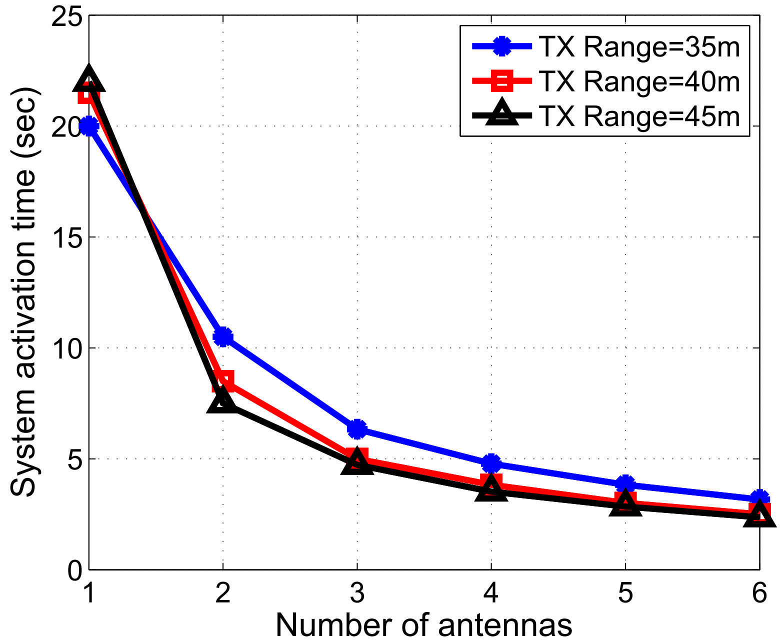


Figure 8.3: The effect of transmission range on the system activation time in a network consisting of 10 nodes with multi-antenna in BF+SDMA+SDM model where 15 sessions are delivered.

8.1 Number of Antennas

To show the effect of number of antennas, we performed some experiments over a network consisting of 10 nodes uniformly distributed over a $100m \times 100m$ area, where 15 unit-demand sessions are routed. The source and destination nodes of each session are chosen randomly. Figures 8.1 to 8.3 show that, generally employing multiple antennas improve the throughput of the network in BF, BF+SMDA and BF+SDMA+SDM models. In BF model (Figure 8.1), we do not achieve better enhancement with more than 2 antennas because the network is not heavily dense and the interference in the network can be suppressed if each node has two antennas. In BF+SDMA model, additional antenna elements are operating in multiple access mode to improve the overall throughput; however, when the number of antennas is more than 4, no improvement is observed (Figure 8.2). Employing SDM technique (i.e. BF+SDMA+SDM model) yields higher throughput as more antenna elements are added. Note that, there is no limit on improving the throughput by increasing the number of antennas. The reason is that the capacity of wireless links is increased and therefore, the traffic is routed with higher transmission rates.

8.2 Transmission Range

The effect of transmission range is studied in Figures 8.1 to 8.3 where 10 nodes are distributed over the area of $100m \times 100m$ and 15 random traffic load are delivered. These figures illustrate that in all three models, increasing the transmission range from $35m$ to $45m$ yields a decrease in the system activation time since the traffic load are now routed through shorter paths (smaller number of hops). The reduction is more significant for BF, in comparison with other models. In BF+SDMA+SDM model, however, the reduction is negligible because when high SDM gain is observed, the effect of number of hops in delivering the traffic load becomes negligible.

8.3 Node Density

To study the effect of node density on the network, we performed our experiments in two different configurations. The first configuration is a network consisting of 20 nodes over a $100m \times 100m$ area and the second one is a 15-node network over the same area. The performance of BF, BF+SDMA and BF+SDMA+SDM models are studied where 10 unit-demand sessions are routed and the transmission range is assumed to be $30m$. Figure 8.4 demonstrates the fact that in general, by increasing the node density in the network the system performance is improved since there is more accessibility for delivering the traffic load in the network.

Moreover, Figure 8.4 shows that, in a 20-node network, the reduction of system activation time is 47%, 78% and 86.9% for the BF, BF+SDMA, BF+SDMA+SDM models, respectively, as we increase the number of antenna elements from 1 to 6.

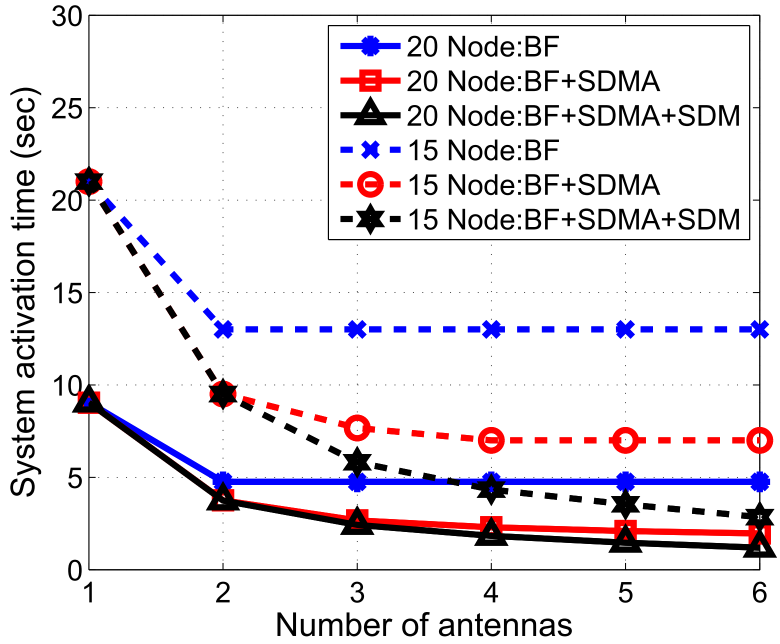


Figure 8.4: The effect of number of nodes on the activation time of the system where transmission range is 30m and 10 sessions are routed.

Nevertheless, this gain is decreased to 38%, 66.7% and 86.5% respectively, in the case of a 15-node network. This reduction in gain is due to the dependency of directive gain (BF) and multiple access gain (SDMA) on the number of nodes. However, for the BF+SDMA+SDM model, the impact of varying the number of nodes on reducing the system activation time is minor because SDM technique improves the performance of the network by increasing the data rate over each link.

Furthermore, Figure 8.4 depicts that by endowing the nodes with multi-antenna techniques, we can cover the service area with fewer access points. For example, 15 nodes endowed with 2 antennas each in the BF+SDMA model, can provide the same throughput as 20 single-antenna nodes (system activation time of 9 seconds). This is happening because interference suppression and simultaneous access to multiple nodes lead to more concurrent active links and increase the spectrum spatial reuse. Therefore, in the BF+SDMA model, the same level of accessibility can be achieved although the network is more sparse. Alternatively, by increasing the number of antennas per node to 4 in the BF+SDMA+SDM model in a 15-node network, we can achieve the same throughput of a 20-node network in BF model with 2 antennas per node (a 4-second activation time for both systems).

Based on the above findings, we distinguish two alternatives for designing wireless networks with guaranteed end-to-end throughput. The first one is to deploy more access points (APs) with less resources and simpler technology (e.g., single antenna or multi-antenna with only BF or BF+SDMA capabilities). The second alternative is to deploy fewer APs with powerful multi-antenna technology like SDM. Therefore, network designers have to investigate the cost trade-off before deployment.

8.4 Traffic load

The effect of the traffic load on the performance of different models is discussed through performing some experiments on a network consisting of 50 nodes with transmission range of $35m$; these nodes are uniformly distributed over a $250m \times 250m$ area. The activation time of the system for 50, 40, 30, 20, 10 random unit-demand sessions in BF, BF+SDMA and BF+SDMA+SDM models are demonstrated in Figures 8.5 to 8.7, respectively. These figures show that in all models, by increasing the number of sessions, the activation time of the system is increased. Moreover, Figure 8.7 illustrates that by providing multiplexing gain (SDM), as more sessions are routed in the network, the performance is enhanced as long as there are enough number of antenna elements. This is true because all the sessions can be delivered through high capacity links.

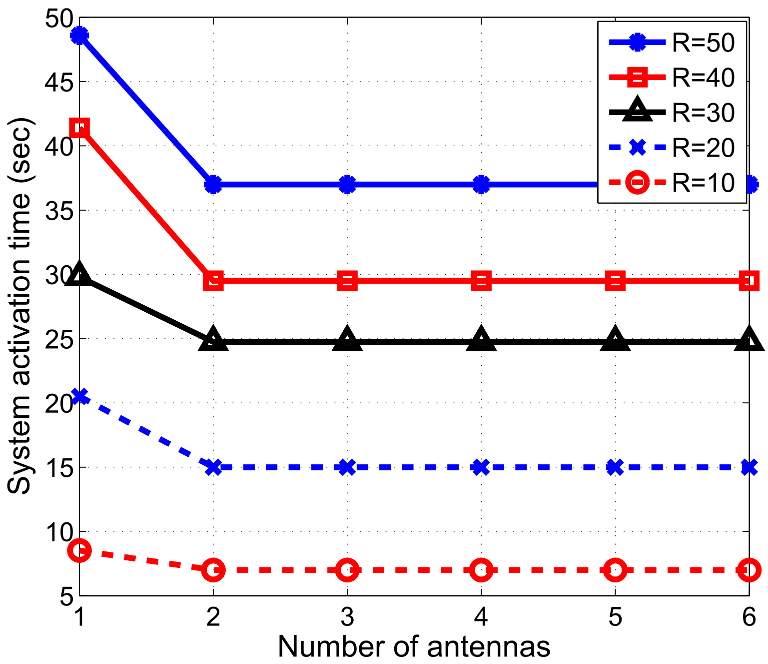


Figure 8.5: The effect of number of sessions on the system activation time where BF model is applied to a network consisting of 50 nodes with $35m$ transmission range in an area of $250m \times 250m$.

8.5 Scalability of the Models

The scalability of BF, BF+SDMA and BF+SDMA+SDM models are studied through comparing the computation time (CPU) for obtaining the optimal solution in two cases: a 15-node network over the $100m \times 100m$ area and a 50-node network over the $250m \times 250m$ area; as 10 sessions are delivered. Table 8.1 demonstrates the computation time for each model in each case. Our models have been run on nodes of a cluster where there are four processing cores per node (model HP Proliant DL145G2) with the available memory of 16 GB. The table shows that in BF and BF+SDMA model, the increase in CPU time is negligible when the number of nodes increases. However, the CPU time grows faster in BF+SDMA+SDM model because the number of active streams on each link can vary from 1 to 6 and consequently, the number of generated configurations (the size of $\mathcal{P}_0$) to obtain the optimal solution

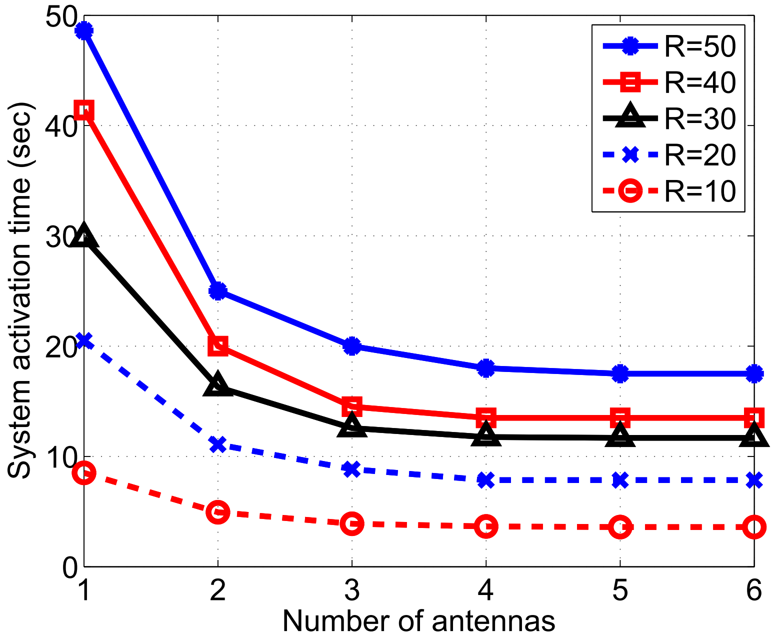


Figure 8.6: The effect of number of sessions on the system activation time where BF+SDMA model is applied to a network consisting of 50 nodes with $35m$ transmission range in an area of $250m \times 250m$.

is increased.

We note that the computation time is mostly determined by the Branch and Bound method [52] which is required for obtaining Integer Linear Programming (ILP) solution in the Pricing sub-problem. However, the Branch and Bound method does not always behave evenly and therefore leads to unexpected computation time in some instances. For example in BF+SDMA model, we observe that in the 15-node network with 3-antenna instance, 66 configurations are generated in a shorter computation time compared to that of the same instance with 4 antennas where 57 configurations are generated.

8.6 Design of Heterogeneous Networks

In this section, we investigate the design of heterogeneous wireless networks which include non-homogenous multi-antenna nodes. We assume we will provide Internet

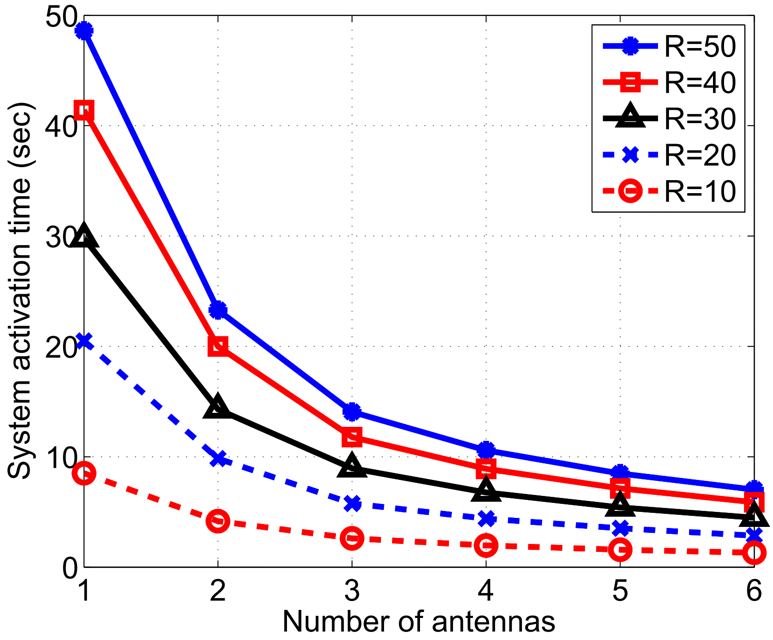


Figure 8.7: The effect of number of sessions on the system activation time where BF+SDMA+SDM model is applied to a network consisting of 50 nodes with 35m transmission range in an area of 250m × 250m.

access to an area where the node locations and the traffic loads are known. We study the optimal network design in the case of uniform and non-uniform distribution of traffic sessions. For instance, the non-uniform traffic occurs when the APs are deployed in a neighborhood with a non-uniform population. Thus, some parts of the network experience heavy loads while in other parts the load may be light. We consider HTRGN-BF+SDMA+SDM model and we obtain the required minimum number of antennas at each node which achieves the minimum system activation time. Indeed, finding a solution for such a heterogeneous network would be more cost effective than equipping all nodes with equal pre-determined number of antennas. In evaluating the performance of HTRGN-BF+SDMA+SDM model, DoF_{max} is assumed equal to 4. In practice, AP manufacturers barely produce APs with more than 4 antenna elements. We have considered two network examples and evaluated their performance as follows.

Table 8.1: Comparison of computation time for a 15-node network and a 50-node network when 10 sessions are delivered.

Multi-antenna technique	Number of antennas	15-node network		50-node network	
		CPU time	size of $\mathcal{P}_0$	CPU time	size of $\mathcal{P}_0$
BF model	1	13s	63	14s	229
	2	11s	43	16s	196
	3	14s	43	11s	197
	4	14s	43	9s	189
	5	10s	43	10s	188
	6	12s	43	15s	185
BF+SDMA model	1	11s	63	9s	227
	2	11s	97	11s	417
	3	6s	66	57s	432
	4	12s	57	27s	291
	5	12s	49	1m12s	239
	6	12s	48	7s	220
BF+SDMA+SDM model	1	7s	75	3m22s	489
	2	13s	113	2h53m23s	859
	3	1m28s	187	2h35m17s	1164
	4	1m36s	237	2h55m27s	1763
	5	2m38s	247	7h38m45s	2103
	6	1m30s	236	1h46m43s	1265

The first example is a 20-node network, uniformly distributed over a $100m \times 100m$ area which is shown in Figure 8.8. We obtain the optimal solutions to the joint routing and scheduling for delivering unit-demand sessions $1, 2$ and 3 (Figure 8.8). The routing paths are illustrated in Figure 8.9 and the numerical results for this scenario are presented in Table 8.2. The table depicts the number of antennas each node should be equipped with to achieve the optimal performance; further, in this scenario, most nodes $(0-4, 8-19)$ have similar number of antennas.

The numerical results for delivering unit-demand sessions $1, 2 \ldots 6$ (Figure 8.8) are presented in Table 8.2. Figure 8.10 demonstrates the corresponding routing paths for these sessions. When comparing Figures 8.9 and 8.10, we observe that in different loading scenarios, our model provides various routing configurations for a particular session to minimize the number of antennas and the system activation time.

Next, we compare these results with a case that the same number of traffic

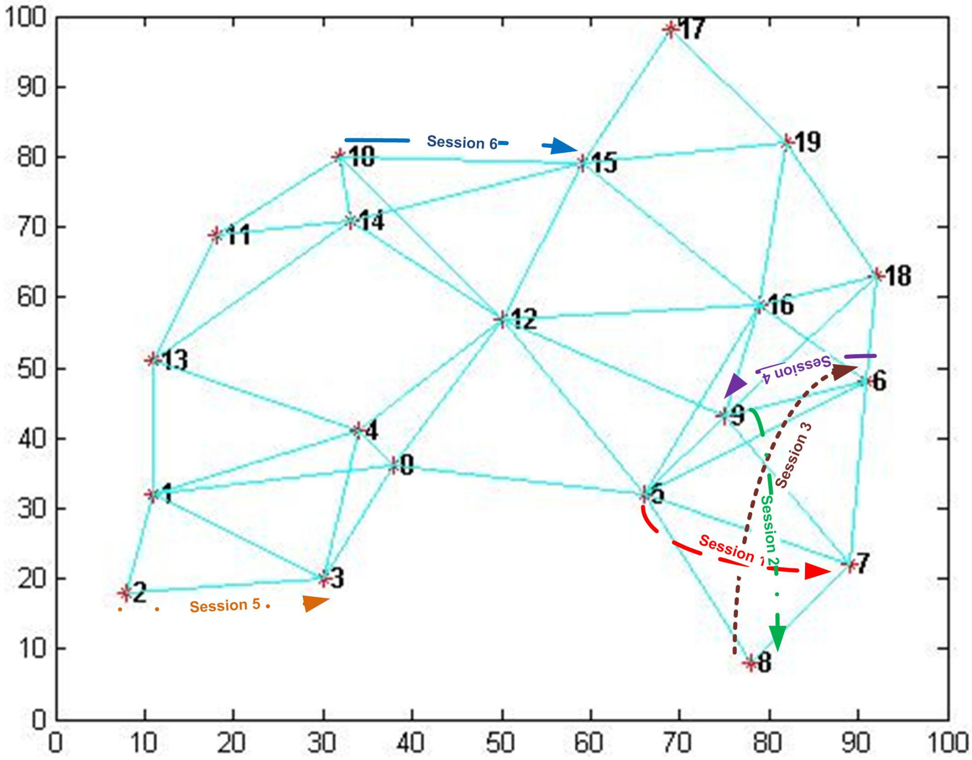


Figure 8.8: A 20-node network uniformly distributed over a $100m \times 100m$ area with transmission range of $30m$. Session 1: $5 \to 7$, Session 2: $9 \to 8$, Session 3: $8 \to 6$, Session 4: $6 \to 9$, Session 5: $2 \to 3$, Session 6: $10 \to 15$.

sessions, i.e. 6 is uniformly distributed over the same network and we obtain the average of the results for 10 different random traffic sessions as shown in last row of Table 8.2. We observe that in non-uniform distributed traffics, nodes with higher number of antennas are barely located in the light traffic areas and thus, there is almost 42% reduction in the total number of antennas which is significantly less than of the uniform traffic case. However, λ_{min} of the non-uniform traffic is larger than that of the uniform traffic sessions; because longer duration of time is required to deliver the heavy traffic through the capacity-limited links in the populated region. In this case, the traffic is routed through the paths outside the populated area (which is a longer route with more number of hops); or alternatively, is scheduled for longer

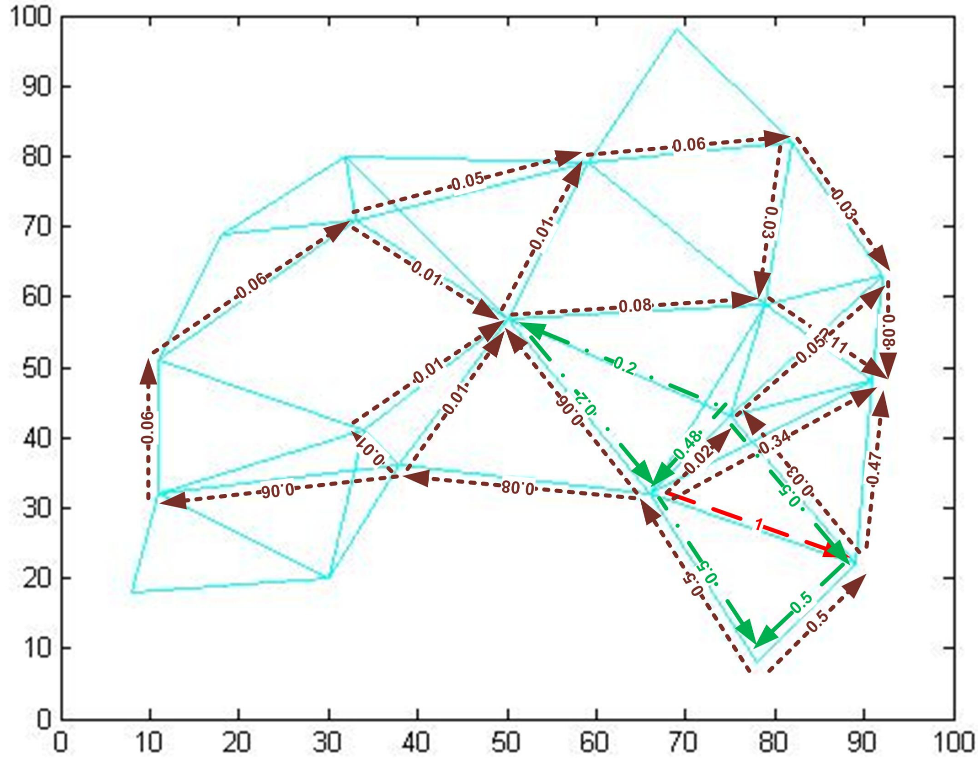


Figure 8.9: Delivering sessions 1 to 3.

period of times to be delivered over the capacity limited links of the shorter paths in the populated area.

In the second example, we consider a network of 50 nodes uniformly distributed over a $250m \times 250m$ area as shown in Figure 8.11. We assume a transmission range of $35m$ at each node. The numerical results of the model are presented for two traffic distributions. In the first one, we assume the source and the destination of traffic loads are randomly distributed throughout the network area. Table 8.3 shows that by applying HTRGN-BF+SDMA+SDM design method, the number of antennas can be reduced up to 56.25% per node on average for 10, 20 and 30 unit-demand sessions. However, the reduction of required antenna elements is decreased to 38.5% when the traffic load is increased to 40 and 50 unit-demand sessions.

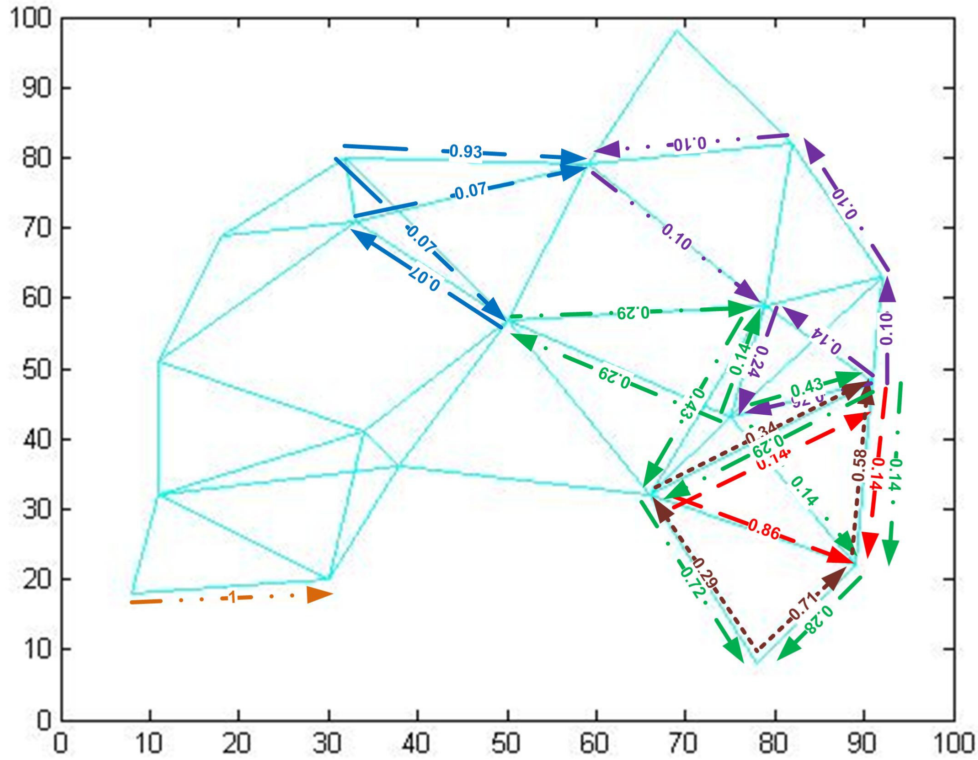


Figure 8.10: Delivering sessions 1-6.

In the second traffic distribution, we consider the network of Figure 8.11 with non-uniform 30 unit traffic; particularly 67% of the sessions are scattered over region A and the rest is distributed across the remaining area. Figure 8.12 shows the minimum number of antennas per node obtained by the design model. Nodes with 3 or 4 antennas are mostly located in and around region A. These nodes will be involved for routing the heavy traffic of region A as well as to combat the exceeded interference due to the large number of active links around this area.

Next, we also investigate the impact of transmission range on the obtained optimal solution where the traffic loads remain unchanged. Figure 8.13 shows that by increasing the transmission range to $50m$, the amount of interference in the network increases and as a result, the total number of antennas increases to suppress

Table 8.2: The numerical results of different traffic sessions in the first example.

Number of antennas per node	4	3	2	1	Total antenna number	ρ_{min}	λ_{min}
3 non-uniform as in Figure 8.8	nodes: 5,7	-	-	nodes: 0-4,8-19	32	1.6	0.5
6 non-uniform as in Figure 8.8	nodes: 5,7	node: 15	nodes: 2,3,10,12	nodes: 0,1,4,6,8, 9,11,13-19	35	1.75	0.5
6 uniform (Ave. of 10 runs)	———				60	3	0.431

the interference. However, when the transmission range becomes $55m$ (Figure 8.14), both λ_{min} and ρ_{min} are improved; and this is due to the fact that the traffic loads are now routed over paths with fewer number of hops. As a result, fewer nodes are getting involved in transmissions/receptions and thus, fewer links will be active. Hence, less (intra-path and inter-path) interference is generated in the network and there will be no need for extra antennas to suppress the interference. These results imply that there is an optimal transmission range (and hence transmission power) for the nodes in the network which may be obtained through proper power control. Indeed, including power control into our optimization model remains as future work.

As a summary to our discussions, for light and uniformly distributed traffic loads (Table 8.3), fewer antennas per node are required. When the traffic is heavier and still uniformly distributed, the number of antennas per node is averagely increased where all nodes are almost equipped with same number of antennas. However, for non-uniformly distributed traffic sessions, the number of antennas per node significantly varies in different parts of the network. As illustrated in Table 8.2 and Figures 8.12, 8.13 and 8.14, in highly loaded areas, the number of antennas is larger. However, this situation has an impact on how sessions are routed; some sessions are routed through multi-path and longer routes to keep the number of antennas at the nodes small.

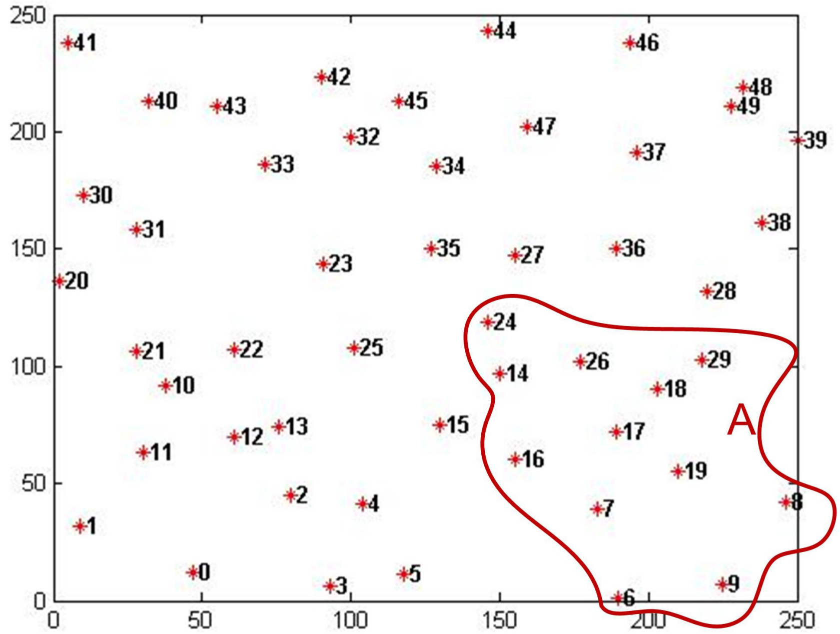


Figure 8.11: A 50-node network uniformly distributed over a $250m \times 250m$.

We have observed that the number of antennas per nodes varies depending on the traffic load and the distribution of traffic in the network. Our numerical results, indeed show the efficacy of HTRGN-BF+SDMA+SDM optimization model in determining the optimal number of antennas per nodes while achieving the minimum system activation time. The minimum system activation time corresponds to the highest throughput. This optimal resource allocation and maximizing the network throughput is practically very useful for all network settings.

Table 8.3: The average minimum number of antennas per node, ρ_{min}, for achieving the minimum system activation time, λ_{min}, in a 50-node network over $250m \times 250m$ area, with transmission range of $35m$ where different traffic load are delivered.

Number of demands	10	20	30	40	50
λ_{min}	1.75	2.67	3.5	4.2	5.17
ρ_{min}	1.75	1.75	1.75	2.5	2.5
Percentage of reduction in number of antennas	56.25%	56.25%	56.25%	38.5%	38.5%

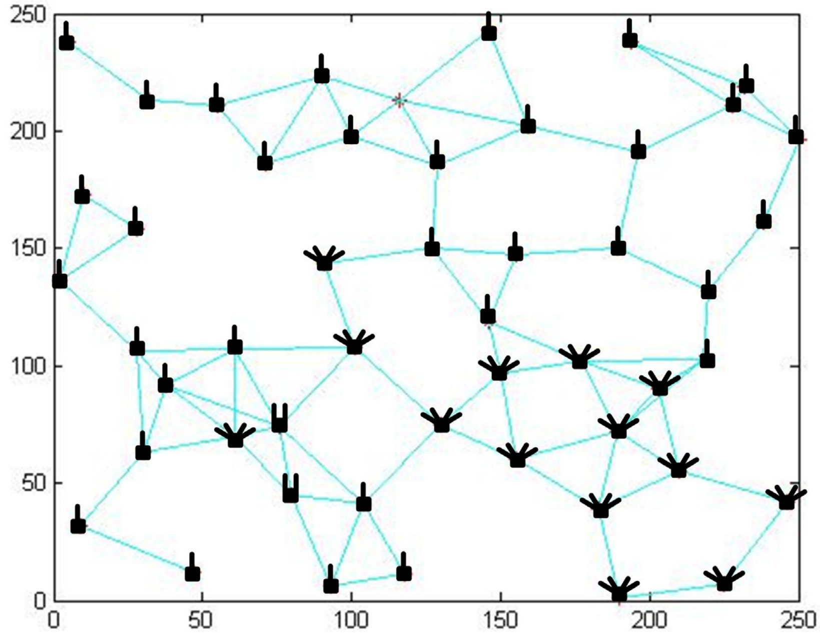


Figure 8.12: 50-node network with non-uniform traffic loads and transmission range of 45m. $\lambda_{min} = 1.17$ seconds, $\rho_{min} = 1.875$ and total number of antennas=94.

8.7 Conclusions

In this chapter, we evaluated the numerical results for joint routing and scheduling in multihop wireless networks using differnt multi-antenna technologies, namely, BF, BF+SDMA and BF+SDMA+SDM. We discussed the effect of transmission range, number of antennas, traffic sessions and number of nodes in different scenarios of network for each model. The numerical results expressed that less number of multi-antenna nodes can yield the same throughput as more single-antenna nodes. In addition, the results showed the benefits of multi-antenna techniques and in particular SDM for improving the network throughput; up to 86.9% reduction in system activation time was achieved by 6-antenna elements in BF+SDMA+SDM model. Moreover, the provided numerical results for designing heterogenous networks indicated that applying our proposed model in the network planning process would considerably reduce the cost of wireless mesh networks deployment.

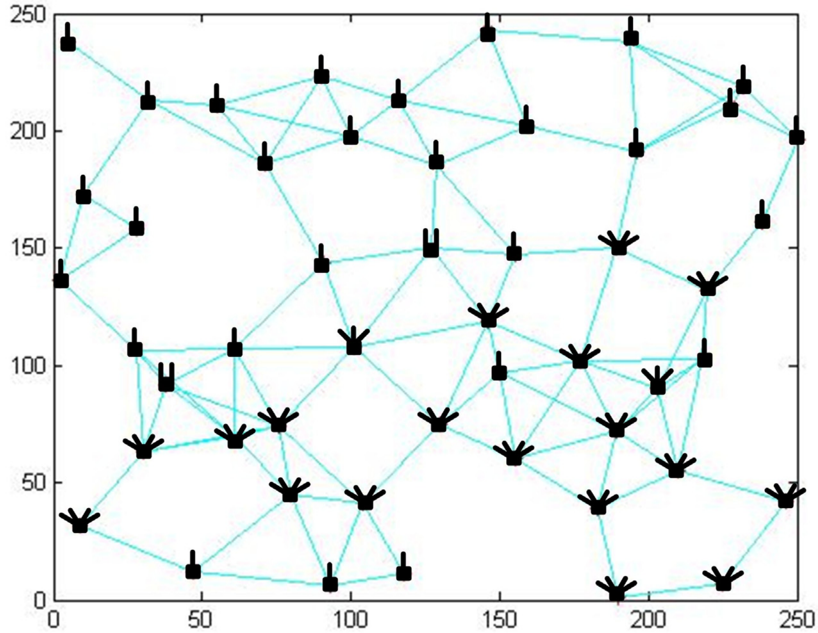


Figure 8.13: 50-node network with non-uniform traffic loads and transmission range of 50m. $\lambda_{min} = 1.17$ seconds, $\rho_{min} = 2.5$ and total number of antennas=113.

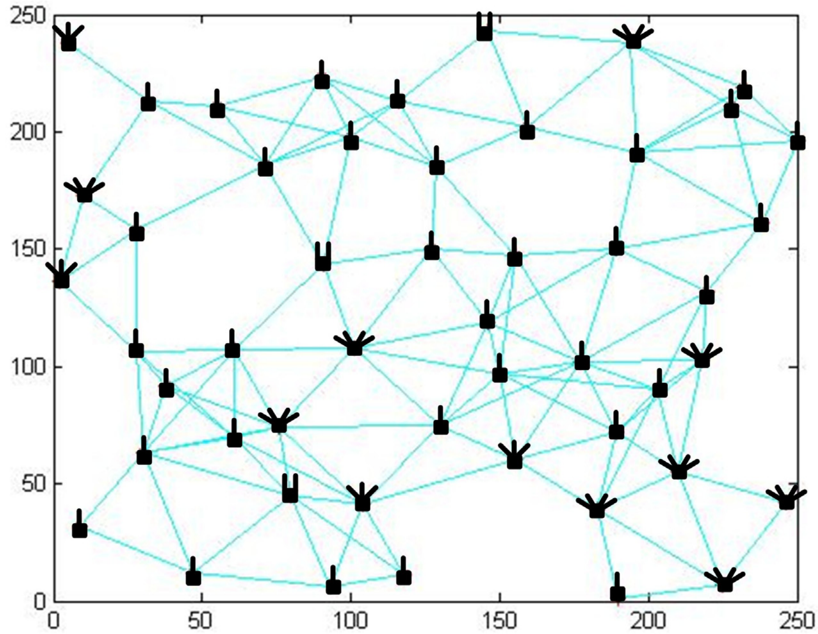


Figure 8.14: 50-node network with non-uniform traffic loads and transmission range of 55m. $\lambda_{min} = 1$ second, $\rho_{min} = 1.75$ and total number of antennas=88.

Chapter 9

Conclusion and Future Work

In this chapter, we overview the conclusions of our research. Then, we identify some of the future directions of this research.

9.1 Conclusions

This thesis presented several cross-layer design methods for studying the benefits of advanced physical and network layer techniques on the performance of multihop wireless networks. Namely, we considered wireless networks where nodes endowed with multi-antennas as well as nodes with multi-packet reception and network coding capabilities. We derived mathematical models for solving the cross-layer design problems and used large scale optimization tools (such as column generation) for solving them.

First, we developed a decomposed model for joint routing and scheduling when successive interference cancellation is employed at the receiving nodes to cancel undesired interference and decoding multiple packets at the same time. Then, we extended our model to another one in which transmission power control is also employed at transmitting nodes. We showed that combining SIC and power control at the physical layer improves the performance of dense networks by 75% comparing to the basic model where no advance technique is used. This remarkable improvement is obtained by cancelling the interference using SIC and reducing the amount of interference through power control. We also showed that SIC is a more powerful technique than power control (40%-50% better) and in sparse networks it is sufficient to employ SIC. [134]

We considered moreover to study the achievable performance gains attained by combining multi-packet reception at the physical layer with network coding at a higher layer. We formulated this combinatorially complex problem of routing and scheduling, under the realistic interference model, as a non-linear optimization problem and we used column generation technique to linearize it and decompose it into two sub-problems. Namely, in the first sub-problem (Restricted Master), we solve the network coding aware routing and in the second sub-problem (Pricing), we solve the interference-free scheduling of links as well as coding components. Our numerical evaluation affirms our expectation for a remarkable performance improvement; that is network coding combined with SIC shows more than 50% reduction in the scheduling period for a medium-size dense network. However, this reduction is around 40% and 19% when only SIC or network coding are applied, respectively [133].

Further, we presented a cross-layer design approach for a wireless network with nodes using multiple antennas and developed a mathematical optimization model for solving it. We used multiple antennas to suppress the interference from neighboring transmissions (BF), to access multiple nodes simultaneously (SDMA) and to transmit with higher data rates (SDM). By integrating different combinations of multi-antenna techniques in physical layer with various constraints from MAC and network layers, three Mixed Integer Linear Programming (MILP) models were presented. The objective in each of these models is to minimize the system activation time. Due to the large number of constraints and variables in these models, the problem was decomposed using Column Generation approach which made each model more scalable. The effect of transmission range, number of antennas, traffic sessions and number of nodes were discussed in different scenarios of network for each model. The numerical results expressed that less number of multi-antenna nodes can yield the same throughput as more single-antenna nodes. In addition, the results showed the benefits of multi-antenna techniques and in particular SDM for improving the network throughput; up to 86.9% reduction in system activation time was achieved by 6-antenna elements in BF+SDMA+SDM model [131, 132].

Finally, we considered a heterogeneous network with different number of antennas per each node and we presented a design approach for minimizing the deployment cost. The provided mathematical model was a multi-criteria optimization with two objectives; namely, jointly minimizing the system activation time and the number of antenna elements per node where BF, SDMA and SDM techniques are available. The provided numerical results indicated that applying the proposed model in the network planning process would considerably reduce the cost of wireless mesh networks deployment [130].

9.2 Future Work

Several future research directions can be pursued and below we present some details for possible extensions:

9.2.1 Designing Protocols

- It has been clarified that careful considerations in higher layer protocols is needed to improve the throughput of a wireless network equipped various physical layer techniques. While various MAC protocols have been proposed for such networks, the impact of multiple antennas on routing and network layer performance have not received much attention. There is only little work on routing protocols for multi-antenna networks which are mostly based on directional transmissions [22, 101, 111] where connectivity is the main concern in the routing protocol. However, the mentioned gains provided by various multi-antenna techniques can be employed to yield higher end-to-end throughput and guarantee diverse QoS as in [40]. The QoS is defined as either high reliability or minimal delay depending on different applications. Moreover, opportunistic routing which is first introduced in [12] proposes a good solution to enhance the performance of the networks with unicast transmissions. This protocol is an integration of MAC and routing in which, the nearest node to the destination is selected to forward the message. In the literature there is only one work which has considered multiple antennas to directionally transmit the signal for opportunistic routing [81].

Therefore, one possible extension of this research is to study and design improved routing protocols while considering cross-layer interactions where various multi-antenna techniques are employed. Several routing protocols such as opportunistic routing can be employed to evaluate the throughput of the multi-antenna networks.

- In this dissertation, we illustrated the benefits of employing multi-packet reception, network coding and their combinations in wireless networks. However, there is little work on designing the proper protocols for such schemes [9, 28, 55, 102]. Therefore, a MAC protocol in which both of these techniques can be exploited would bring higher throughput and better fairness into the wireless network.

9.2.2 Distributed Optimization Methods

While a central optimization is suitable for finding the capacity limit of wireless networks, a more realistic and scalable throughput analysis can be developed through distributed optimization. In this method, controlling the activity of the nodes is done in a distributed manner. Here, distributed methods for the optimization such as Lagrangian decomposition methods [73, 95] can be another direction to extend this research.

9.2.3 Interference Cancellation in Cooperative Networks

Although providing multiple antennas at the transmitter and receiver side of a communication link is very advantageous, such deployment is not feasible in all scenarios. There is a size limitation in many mobile devices so that the certain required space between the antenna elements is not affordable. In some other applications which are power limited, such as wireless sensor networks, a node may not be able to provide the power for the internal processers' circuitry. Furthermore, multi-antennas deployment demands for more than one RF front-ends that yields higher complexity in the hardware and hence leads to higher expenses. However, by allowing nearby single-antenna nodes to transmit cooperatively in the network, a virtual multiple antenna arrays can be formed that achieves the benefits of multi-antenna transmission without requiring installation of multiple antenna elements at each node. Therefore, cooperative networks are a solution to take advantages of multi-antenna systems, i.e. higher data rates and reliability without needing multiple antennas at each node. This topic has been extensively investigated in the literature [57, 70, 71, 91, 103, 104]. The cooperation can be provided through amplify-and-forward [72], decode-and forward [103] and coded cooperation [56, 57].

Nevertheless, synchronizing the transmissions of the nodes and managing the interference in such networks is very critical for achieving the desired capacity.

Therefore, developing and employing distributed synchronization schemes such as [43, 113] which provide interference nullifying and alignment through a cross-layer design in any of the cooperative network methods would be an interesting future direction of this thesis.

Bibliography

[1] 802.11n challenge. Technical report, Layland Consulting and Joanie M. Wexler & Associates and Webtorials, March 2009.

[2] R. Ahlswede, Ning Cai, S.-Y.R. Li, and R.W. Yeung. Network information flow. *Information Theory, IEEE Transactions on*, 46(4):1204 –1216, jul 2000.

[3] P.D. Alexander, M.C. Reed, J.A. Asenstorfer, and C.B. Schlegel. Iterative multiuser interference reduction: turbo cdma. *Communications, IEEE Transactions on*, 47(7):1008 –1014, jul 1999.

[4] H.M. Almasaeid and A.E. Kamal. Assisted-multicast scheduling in wireless cognitive mesh networks. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1 –5, may 2010.

[5] J.G. Andrews. Interference cancellation for cellular systems: a contemporary overview. *Wireless Communications, IEEE*, 12(2):19 – 29, april 2005.

[6] Emre Telat Ar and I. Emre Telatar. Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications*, 10:585–595, 1999.

[7] A. Argyriou. Cross-layer and cooperative opportunistic network coding in wireless ad hoc networks. *Vehicular Technology, IEEE Transactions on*, 59(2):803 –812, feb. 2010.

[8] T. Hahn B. Mumey, Jian Tang. Algorithmic aspects of communications in multihop wireless networks with mimo links. In *IEEE Proceedings ICC*, pages 1–6, May 2010.

[9] S. Barghi, H. Jafarkhani, and H. Yousefi'zadeh. Mimo-assisted mpr-aware mac design for asynchronous wlans. *Networking, IEEE/ACM Transactions on*, 19(6):1652 –1665, dec. 2011.

[10] R. Bhatia and Li Li. Throughput optimization of wireless mesh networks with mimo links. *INFOCOM 2007. IEEE*, pages 2326–2330, May 2007.

[11] Ezio Biglieri, Robert Calderbank, Anthony Constantinides, Andrea Goldsmith, and H. Vincent Poor. *MIMO Wireless Communications*. Cambridge University Press, 2007.

[12] Sanjit Biswas and Robert Morris. Exor: opportunistic multi-hop routing for wireless networks. In *SIGCOMM '05: Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications*, volume 35, pages 133–144, New York, NY, USA, October 2005. ACM Press.

[13] Patrik Björklund, Peter Värbrand, and Di Yuan. A column generation method for spatial tdma scheduling in ad hoc networks. *Ad Hoc Networks*, 2(4):405 – 418, 2004.

[14] J. Blomer and N. Jindal. Transmission capacity of wireless ad hoc networks: Successive interference cancellation vs. joint detection. In *Communications, 2009. ICC '09. IEEE International Conference on*, pages 1 –5, june 2009.

[15] Yih-Hao Lin Bongyong Song and R.L. Cruz. Weighted max-min fair beamforming, power control, and scheduling for a miso downlink. *IEEE Transactions on Wireless Communications*, 7(2):464–469, 2008.

[16] Stephen Boyd and Jacob Mattingley. Branch and bound methods, 2003.

[17] Min Cao, Xiaodong Wang, Seung-Jun Kim, and M. Madihian. Multi-hop wireless backhaul networks: a cross-layer design paradigm. *Selected Areas in Communications, IEEE Journal on*, 25(4):738–748, May 2007.

[18] G.D. Celik, G. Zussman, W.F. Khan, and E. Modiano. Mac for networks with multipacket reception capability and spatially distributed nodes. *Mobile Computing, IEEE Transactions on*, 9(2):226 –240, feb. 2010.

[19] Biao Chen and M.J. Gans. Mimo communications in ad hoc networks. In *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, volume 4, pages 2434–2438 Vol. 4, 2005.

[20] H. Chen, F. Yu, H.C.B. Chan, and V.C.M. Leung. A novel multiple access scheme over multi-packet reception channels for wireless multimedia networks. *Wireless Communications, IEEE Transactions on*, 6(4):1501 –1511, april 2007.

[21] Wei Chen, K.B. Letaief, and Zhigang Cao. A cross layer method for interference cancellation and network coding in wireless networks. In *Communications, 2006. ICC '06. IEEE International Conference on*, volume 8, pages 3693 –3698, june 2006.

[22] Bow-Nan Cheng, M. Yuksel, and S. Kalyanaraman. Orthogonal rendezvous routing protocol for wireless mesh networks. *Networking, IEEE/ACM Transactions on*, 17(2):542–555, April 2009.

[23] Mung Chiang, S.H. Low, A.R. Calderbank, and J.C. Doyle. Layering as optimization decomposition: A mathematical theory of network architectures. *Proceedings of the IEEE*, 95(1):255 –312, jan. 2007.

[24] R.R. Choudhury and N. Vaidya. Mac-layer capture: A problem in wireless mesh networks using beamforming antennas. *Sensor, Mesh and Ad Hoc Communications and Networks, 2007. SECON '07. 4th Annual IEEE Communications Society Conference on*, pages 401–410, June 2007.

[25] Shan Chu and Xin Wang. Opportunistic and cooperative spatial multiplexing in mimo ad hoc networks. In *MobiHoc '08*, pages 63–72. ACM, 2008.

[26] V. Chvátal. *Linear Programming.* Series of Books in the Mathematical Sciences. W.H. Freeman, 1983.

[27] J. Cloud, L.M. Zeger, and M. Medard. Mac centered cooperation - synergistic design of network coding, multi-packet reception, and improved fairness to increase network throughput. *Selected Areas in Communications, IEEE Journal on*, 30(2):341 –349, february 2012.

[28] Jason Cloud, Linda Zeger, and Muriel Médard. Co-designing multi-packet reception, network coding, and mac using a simple predictive model. *CoRR*, abs/1101.5779, 2011.

[29] V. Corvino, L. Giupponi, A. Perez Neira, V. Tralli, and R. Verdone. Cross-layer radio resource allocation: The journey so far and the road ahead. In *Cross Layer Design, 2009. IWCLD '09. Second International Workshop on*, pages 1–6, 2009.

[30] CPLEX. *Using the CPLEX Callable Library.* CPLEX Optimization Inc., 9.1.3 edition, 2005.

[31] J. Crichigno, M.Y. Wu, and W. Shu. Throughput optimization in wireless networks with multi-packet reception and directional antennas. In *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*, pages 1 –6, april 2009.

[32] Abhik K. Das, Sriram Vishwanath, Syed Ali Jafar, and Athina Markopoulou. Network coding for multiple unicasts: An interference alignment approach. *CoRR*, abs/1008.0235, 2010.

[33] E. W. Dijkstra. A note on two problems in connexion with graphs. *NUMERISCHE MATHEMATIK*, 1(1):269–271, 1959.

[34] Matthias Ehrgott. *Multicriteria Optimization.* Springer, 2nd edition, 2005.

[35] J. El-Najjar, H.M.K. AlAzemi, and C. Assi. On the interplay between spatial reuse and network coding in wireless networks. *Wireless Communications, IEEE Transactions on*, 10(2):560 –569, february 2011.

[36] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, pages 41–59, 1996.

[37] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wirel. Pers. Commun.*, 6(3):311–335, 1998.

[38] J. J. Garcia-Luna-Aceves, Hamid R. Sadjadpour, and Zheng Wang. Challenges: Towards truly scalable ad hoc networks. MobiCom '07. ACM, 2007.

[39] A. Ghosh, D.R. Wolter, J.G. Andrews, and R. Chen. Broadband wireless access with wimax/802.16: current performance benchmarks and future potential. *Communications Magazine, IEEE*, 43(2):129 –136, feb. 2005.

[40] A. Gkelias, F. Boccardi, C.H. Liu, and K.K. Leung. Mimo routing with qos provisioning. In *Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on*, pages 46–50, May 2008.

[41] A. Gkelias and K. K. Leung. *Wireless Mesh Networks Architecture and Protocols*, chapter 11, pages 277–307. Springer, 2008.

[42] Andrea Goldsmith. *Wireless Communications*. Cambridge Univ. Press, 2006.

[43] Shyamnath Gollakota, Samuel David Perli, and Dina Katabi. Interference alignment and cancellation. *SIGCOMM Comput. Commun. Rev.*, 39(4):159–170, 2009.

[44] A. Grant. Joint decoding and channel estimation for linear mimo channels. *Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE*, 3:1009–1012 vol.3, 2000.

[45] Hongyu Gu and Chenyang Yang. Throughput analysis of peer-to-peer uwb asynchronous cdma networks. In *Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th*, pages 1 –5, sept. 2006.

[46] V. C. Gungor, E. Natalizio, P. Pace, and S. Avallone. *Wireless Mesh Networks Architecture and Protocols*, chapter 1. Springer, 2008.

[47] P. Gupta and P.R. Kumar. The capacity of wireless networks. *Information Theory, IEEE Transactions on*, 46(2):388–404, Mar 2000.

[48] Piyush Gupta, Student Member, and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46:388–404, 2000.

[49] Daniel Halperin, Thomas Anderson, and David Wetherall. Taking the sting out of carrier sense: interference cancellation for wireless lans. In *In Proceedings of the 14th ACM international conference on Mobile computing and networking*, 2008.

[50] B. Hamdaoui and P. Ramanathan. A cross-layer admission control framework for wireless ad-hoc networks using multiple antennas. *Wireless Communications, IEEE Trans. on*, 6(11):4014–4024, Nov. 2007.

[51] Bechir Hamdaoui and Kang G. Shin. Characterization and analysis of multi-hop wireless mimo network throughput. In *MobiHoc '07*. ACM, 2007.

[52] Frederick S. Hillier and Gerald J. Lieberman. *Introduction to Operations Research*. McGraw-Hill, 2005.

[53] Honglin Hu, Jijun Luo, and Xiaodong Zhang. *Wireless Mesh Networking, Architectures, Protocols and Standards*, chapter 11. Auerbach Publication, 2007.

[54] M. Hu and J. Zhang. Mimo ad hoc networks: Medium access control, saturation throughput and optimal hop distance. *Communications and Networks*, (Special Issue on Mobile Ad Hoc networks):317–330, December 2004.

[55] Wei Lan Huang, K. Letaief, and Ying Jun Zhang. Cross-layer multi-packet reception based medium access control and resource allocation for space-time coded mimo/ofdm. *Wireless Communications, IEEE Transactions on*, 7(9):3372 –3384, september 2008.

[56] T.E. Hunter and A. Nosratinia. Cooperation diversity through coding. *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, page 220, 2002.

[57] T.E. Hunter and A. Nosratinia. Diversity through coded cooperation. *Wireless Communications, IEEE Transactions on*, 5(2):283 – 289, feb. 2006.

[58] A. Iyer, C. Rosenberg, and A. Karnik. What is the right model for wireless channel interference? *Wireless Communications, IEEE Transactions on*, 8(5):2662 –2671, may 2009.

[59] Rias Muhamed Jeffrey G Andrews, Arunabha Ghosh. *Fundamentals of WiMAX: understanding broadband wireless networking*. Prentice Hall PTR, 2007.

[60] Y. Thomas Hou Jia Liu, Yi Shi. A tractable and accurate cross-layer model for multi-hop mimo networks. In *IEEE Proceedings INFOCOM*, pages 2348–2356, 2010.

[61] Canming Jiang, Yi Shi, Y. Thomas Hou, Sastry Kompella Wenjing Lou, and Scott F. Midkiff. Squeezing the most out of interference: An optimization framework for joint interference exploitation and avoidance. In *INFOCOM, 2012 Proceedings IEEE*, page 9, april 2012.

[62] Björn Johansson, Pablo Soldati, and Mikael Johansson. Mathematical decomposition techniques for distributed cross-layer optimization of data networks. *IEEE Journal on Selected Areas in Communications*, pages 1535–1547, 2006.

[63] M. Johansson and L. Xiao. Cross-layer optimization of wireless networks using nonlinear column generation. *Wireless Communications, IEEE Transactions on*, 5(2):435–445, Feb. 2006.

[64] Nathaniel M. Jones, Brooke Shrader, and Eytan Modiano. optimal routing and scheduling for a simple network coding scheme. In *INFOCOM, 2012 Proceedings IEEE*, page 9, april 2012.

[65] Sachin Katti, Shyamnath Gollakota, and Dina Katabi. Embracing wireless interference: analog network coding. *SIGCOMM Comput. Commun. Rev.*, 37(4):397–408, August 2007.

[66] Sachin Katti, Hariharan Rahul, Wenjun Hu, Dina Katabi, Muriel Médard, and Jon Crowcroft. Xors in the air: practical wireless network coding. *IEEE/ACM Trans. Netw.*, 16(3):497–510, June 2008.

[67] Seung-Jun Kim, Xiaodong Wang, and M. Madihian. Cross-layer design of wireless multihop backhaul networks with multiantenna beamforming. *Mobile Computing, IEEE Transactions on*, 6(11):1259–1269, Nov. 2007.

[68] S. Kompella, J.E. Wieselthier, and A. Ephremides. Revisiting the optimal scheduling problem. In *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*, pages 492 –497, march 2008.

[69] U.C. Kozat, I. Koutsopoulos, and L. Tassiulas. A framework for cross-layer design of energy-efficient communication with qos provisioning in multi-hop wireless networks. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 2, pages 1446 – 1456 vol.2, march 2004.

[70] J. Nicholas Laneman. Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Trans. Inform. Theory*, 49:2415–2525, 2003.

[71] J. Nicholas Laneman, David N. C. Tse, and Gregory W. Wornell. Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Trans. Inform. Theory*, 50:3062–3080, 2004.

[72] J.N. Laneman, G.W. Wornell, and D.N.C. Tse. An efficient protocol for realizing cooperative diversity in wireless networks. page 294, 2001.

[73] Markus Leitner and Günther R. Raidl. Lagrangian decomposition, metaheuristics, and hybrid approaches for the design of the last mile in fiber optic networks. In *Hybrid Metaheuristics 2008, volume 5296 of LNCS*, pages 158–174. Springer, 2008.

[74] Kai Li and Xiaodong Wang. Cross-layer design of wireless mesh networks with network coding. *Mobile Computing, IEEE Transactions on*, 7(11):1363 –1373, nov. 2008.

[75] Xin Li, Yimin Zhang, and Moeness G. Amin. Node throughput analysis of decentralized wireless networks using multibeam antennas in multipath environments. *Wirel. Pers. Commun.*, 50(4):447–456, September 2009.

[76] Xiaojun Lin, N.B. Shroff, and R. Srikant. A tutorial on cross-layer optimization in wireless networks. *Selected Areas in Communications, IEEE Journal on*, 24(8):1452 –1463, aug. 2006.

[77] Yih-Hao Lin, T. Javidi, R.L. Cruz, and L.B. Milstein. Distributed link scheduling, power control and routing for multi-hop wireless mimo networks. *ACSSC '06.*, pages 122–126, 2006.

[78] Jia Liu, Y.T. Hou, Yi Shi, and H. Sherali. Cross-layer optimization for mimo-based wireless ad hoc networks: Routing, power allocation, and bandwidth allocation. *Selec.Areas in Comm.,IEEE Jour. on*, Aug. 2008.

[79] Jia-Liang Lu, P. Vandenhove, Wei Shu, and Min-You Wu. Enhancing throughput in wireless multi-hop network with multiple packet reception. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1 –5, june 2011.

[80] Marco E. Lübbecke and Jacques Desrosiers. Selected topics in column generation. *Oper. Res.*, 53(6):1007–1023, November 2005.

[81] Chun-Pong Luk, Wing-Cheong Lau, and On-Ching Yue. An analysis of opportunistic routing in wireless mesh network. In *Communications, 2008. ICC '08. IEEE International Conference on*, pages 2877–2883, 2008.

[82] Chunbo Luo, Yu Gong, and Fu-Chun Zheng. Interference cancellation in two-path successive relay system with network coding. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, pages 465 –469, sept. 2010.

[83] Jun Luo, C. Rosenberg, and A. Girard. Engineering wireless mesh networks: Joint scheduling, routing, power control, and rate adaptation. *Networking, IEEE/ACM Transactions on*, 18(5):1387 –1400, oct. 2010.

[84] Shaohe Lv, Xiaodong Wang, and Xingming Zhou. Scheduling under sinr model in ad hoc networks with successive interference cancellation. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1 –5, dec. 2010.

[85] Shaohe Lv, Weihua Zhuang, Xiaodong Wang, Chi Liu, and Xingming Zhou. Maximizing capacity in the sinr model in wireless networks with successive interference cancellation. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1 –6, june 2011.

[86] Shaohe Lv, Weihua Zhuang, Xiaodong Wang, and Xingming Zhou. Link scheduling in wireless networks with successive interference cancellation. *Computer Networks*, 55(13):2929 – 2941, 2011.

[87] Shaohe Lv, Weihua Zhuang, Xiaodong Wang, and Xingming Zhou. Scheduling in wireless ad hoc networks with successive interference cancellation. In *INFOCOM, 2011 Proceedings IEEE*, pages 1287 –1295, april 2011.

[88] George Mavrotas. Effective implementation of the ϵ-constraint method in multi-objective mathematical programming problems. *Applied Mathematics and Computation*, 213(2):455–465, 2009.

[89] Guowang Miao, Nageen Himayat, Ye (Geoffrey) Li, and Ananthram Swami. Cross-layer optimization for energy-efficient wireless communications: a survey. *Wirel. Commun. Mob. Comput.*, 9(4):529–542, April 2009.

[90] P. Mitran, C. Rosenberg, and S. Shabdanov. Throughput optimization in wireless multihop networks with successive interference cancellation. In *Wireless Telecommunications Symposium (WTS), 2011*, pages 1 –7, april 2011.

[91] Patrick Mitran, Hideki Ochiai, and Vahid Tarokh. Space-time diversity enhancements using collaborative communications. *IEEE Trans. Inform. Theory*, 51:2041–2057, 2005.

[92] B. Mumey, Jian Tang, and T. Hahn. Joint stream control and scheduling in multihop wireless networks with mimo links. In *Commun., 2008. ICC '08. IEEE International Conferance on*, pages 2921–2925, May 2008.

[93] J.C. Mundarath, Pt'. Ramanathan, and B.D. Van Veen. Exploiting spatial multiplexing and reuse in multi-antenna wireless ad hoc networks. *Ad Hoc Networks*, 7(2):281 – 293, 2009.

[94] R. Niati, A.H. Banihashemi, and T. Kunz. Throughput and energy optimization in wireless networks: Joint mac scheduling and network coding. *Vehicular Technology, IEEE Transactions on*, 61(3):1372 –1382, march 2012.

[95] Daniel P. Palomar and Mung Chiang. A tutorial on decomposition methods for network utility maximization. *IEEE J. Sel. Areas Commun*, 24:1439–1451, 2006.

[96] Kyung-Joon Park, LaeYoung Kim, and Jennifer C. Hou. Adaptive physical carrier sense in topology-controlled wireless networks. *IEEE Transactions on Mobile Computing*, 9:87–97, 2010.

[97] P. Patel and J. Holtzman. Analysis of a simple successive interference cancellation scheme in a ds/cdma system. *Selected Areas in Communications, IEEE Journal on*, 12(5):796 –807, jun 1994.

[98] Man-On Pun, Weiyan Ge, Dong Zheng, Junshan Zhang, and H. Vincent Poor. Distributed opportunistic scheduling for mimo ad-hoc networks. In *Proceedings of IEEE International Conference on Communications, ICC 2008, Beijing, China, 19-23 May 2008*, pages 3689–3693. IEEE, 2008.

[99] Dajun Qian, Dong Zheng, Junshan Zhang, and Ness Shroff. Csma-based distributed scheduling in multi-hop mimo networks under sinr model. In *Proceedings of the 29th conference on Information communications*, INFOCOM'10, pages 2865–2873, Piscataway, NJ, USA, 2010. IEEE Press.

[100] A. Rezaee, L. Zeger, and M. Medard. Multi packet reception and network coding. In *MILITARY COMMUNICATIONS CONFERENCE, 2010 - MILCOM 2010*, pages 1393 –1398, 31 2010-nov. 3 2010.

[101] Siuli Roy, Dola Saha, S. Bandyopadhyay, Tetsuro Ueda, and Shinsuke Tanaka. A network-aware mac and routing protocol for effective load balancing in ad hoc wireless networks with directional antenna. In *MobiHoc '03: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, pages 88–97, New York, NY, USA, 2003. ACM.

[102] S. Sen, N. Santhapuri, R. Choudhury, and S. Nelakuditi. Successive interference cancellation: Carving out mac layer opportunities. *Mobile Computing, IEEE Transactions on*, PP(99):1, 2012.

[103] Andrew Sendonaris, Elza Erkip, and Behnaam Aazhang. User cooperation diversity - part i: System description. *IEEE Trans. Commun*, 51:1927–1938, 2003.

[104] Andrew Sendonaris, Elza Erkip, and Behnaam Aazhang. User cooperation diversity - part ii: Implementation aspects and performance analysis. *IEEE Transactions on Communications*, 51:1939–1948, 2003.

[105] S. Sengupta, S. Rayanchu, and S. Banerjee. An analysis of wireless network coding for unicast sessions: The case for coding-aware routing. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 1028 –1036, may 2007.

[106] Jun-Bae Seo and V.C.M. Leung. Design and analysis of cross-layer contention resolution algorithms for multi-packet reception slotted aloha systems. *Wireless Communications, IEEE Transactions on*, 10(3):825 –833, march 2011.

[107] S. Shabdanov, C. Rosenberg, and P. Mitran. Joint routing, scheduling, and network coding for wireless multihop networks. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2011 International Symposium on*, pages 33 –40, may 2011.

[108] Samat Shabdanov, Patrick Mitran, and Catherine Rosenberg. Cross-layer optimization using advanced physical layer techniques in wireless mesh networks. *IEEE Transactions on Wireless Communications*, page 10, to appear.

[109] Xin Wang Shan Chu. Adaptive exploitation of cooperative relay for high performance communications in mimo ad hoc networks. In *IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, pages 21–30, 2010.

[110] Yi Shi, Y. Thomas Hou, Jia Liu, and Sastry Kompella. How to correctly use the protocol interference model for multi-hop wireless networks. In *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '09, pages 239–248, New York, NY, USA, 2009. ACM.

[111] Weisheng Si and S. Selvakennedy. A position-based deployment and routing approach for directional wireless mesh networks. In *Computer Communications and Networks, 2008. ICCCN '08. Proceedings of 17th International Conference on*, pages 1–8, Aug. 2008.

[112] V. Srivastava and M. Motani. Cross-layer design: a survey and the road ahead. *Communications Magazine, IEEE*, 43(12):112 – 119, dec. 2005.

[113] Kun Tan, He Liu, Ji Fang, Wei Wang, Jiansong Zhang, Mi Chen, and Geoffrey M. Voelker. Sam: enabling practical spatial multiple access in wireless lan. In *MobiCom '09: Proceedings of the 15th annual international conference on Mobile computing and networking*, pages 49–60. ACM, 2009.

[114] Lang Tong, Qing Zhao, and G. Mergen. Multipacket reception in random access wireless networks: from signal processing to optimal medium access control. *Communications Magazine, IEEE*, 39(11):108 –112, nov 2001.

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

[116] D.N.C. Tse, P. Viswanath, and Lizhong Zheng. Diversity-multiplexing tradeoff in multiple-access channels. *IEEE Trans. on Inform. Theory*, Sept. 2004.

[117] R. Van Nee. Breaking the gigabit-per-second barrier with 802.11ac. *Wireless Communications, IEEE*, 18(2):4, april 2011.

[118] M.K. Varanasi. Group detection for synchronous gaussian code-division multiple-access channels. *Information Theory, IEEE Transactions on*, 41(4):1083 –1096, jul 1995.

[119] M.K. Varanasi and B. Aazhang. Multistage detection in asynchronous code-division multiple-access communications. *Communications, IEEE Transactions on*, 38(4):509 –519, apr 1990.

[120] Jun Wang, Peng Du, Weijia Jia, Liusheng Huang, and Huan Li. Joint bandwidth allocation, element assignment and scheduling for wireless mesh networks with mimo links. *Computer Commun.*, 31(7):1372 – 1384, 2008. Special Issue: Resource Management and routing in Wireless Mesh Networks.

[121] Shenghui Wang and Zhenzhao Liu. Interference known network coding aware routing metric in wireless network. In *Wireless Communications, Networking and Mobile Computing (WiCOM), 2011 7th International Conference on*, pages 1 –4, sept. 2011.

[122] Xin Wang and J. J. Garcia-Luna-Aceves. Embracing interference in ad hoc networks using joint routing and scheduling with multiple packet reception. *Ad Hoc Netw.*, 7(2):460–471, 2009.

[123] Xin Wang and J. J. Garcia-Luna-Aceves. Embracing interference in ad hoc networks using joint routing and scheduling with multiple packet reception. *Ad Hoc Netw.*, 7(2):460–471, March 2009.

[124] Zheng Wang, Hamid Sadjadpour, and Jose Joaquin Garcia-Luna-Aceves. The capacity and energy efficiency of wireless ad hoc networks with multi-packet reception. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '08, pages 179–188, New York, NY, USA, 2008. ACM.

[125] S.P. Weber, J.G. Andrews, Xiangying Yang, and G. de Veciana. Transmission capacity of wireless ad hoc networks with successive interference cancellation. *Information Theory, IEEE Transactions on*, 53(8):2799 –2814, aug. 2007.

[126] J.H. Winters. Smart antenna techniques and their application to wireless ad hoc networks. *Wireless Communications, IEEE*, 13(4):77–83, Aug. 2006.

[127] Laurence A. Wolsey and George L. Nemhauser. *Integer and Combinatorial Optimization*. Wiley, 1999.

[128] Yang Xiao. Ieee 802.11n: enhancements for higher throughput in wireless lans. *Wireless Communications, IEEE*, 12(6):82 – 91, dec. 2005.

[129] Jian Tang R.S. Wolff Yajing Xu, Shen Wan. Interference aware routing and scheduling in wireless backhaul networks with smart antennas. In *IEEE Prceedings SECON*, pages 1–9, 2009.

[130] Mina Yazdanpanah, Chadi Assi, and Yousef Shayan. Cross-layer optimization for wireless mesh networks with smart antennas. *Computer Communications*, 34(16):1894–1911, 2011.

[131] Mina Yazdanpanah, Chadi Assi, and Yousef R. Shayan. Cross-layer optimization for wireless mesh networks with multi-antenna beam-forming. In *WCNC*, pages 1–6. IEEE, 2010.

[132] Mina Yazdanpanah, Chadi Assi, and Yousef R. Shayan. Optimal joint routing and scheduling in wireless mesh networks with smart antennas. In *WOWMOM*, pages 1–7, 2010.

[133] Mina Yazdanpanah, Samir Sebbah, Chadi Assi, and Yousef Shayan. How much gain does network coding combined with successive interference cancellation bring to a wireless network? *submitted to IEEE/ACM Transaction on Networking*, 2012.

[134] Mina Yazdanpanah, Samir Sebbah, Chadi Assi, and Yousef Shayan. Is power control beneficial in wireless multihop networks with successive interference cancellation? *to be submitted*, 2012.

[135] Jin Y. Yen. Finding the k shortest loopless paths in a network. *Management Science*, 17(11):pp. 712–716, 1971.

[136] L. You, P. Wu, Z. Pan, H. Hu, J. Song, and M. Song. Cross-layer optimization of wireless multi-hop networks with network coding. In *Wireless Communications Signal Processing, 2009. WCSP 2009. International Conference on*, pages 1 –5, nov. 2009.

[137] Jun Yuan, Zongpeng Li, Wei Yu, and Baochun Li. A cross-layer optimization framework for multihop multicast in wireless mesh networks. *Selected Areas in Communications, IEEE Journal on*, 24(11):2092 –2103, nov. 2006.

[138] Jihui Zhang, Haitao Wu, Qian Zhang, and Bo Li. Joint routing and scheduling in multi-radio multi-channel multi-hop wireless networks. pages 631–640 Vol. 1, Oct. 2005.

[139] Shengli Zhang. Hot topic: physical-layer network coding. In *in Proc. of ACM Mobicom*, pages 358–365, 2006.

[140] Qing Zhao and L. Tong. A multiqueue service room mac protocol for wireless networks with multipacket reception. *Networking, IEEE/ACM Transactions on*, 11(1):125 – 137, feb 2003.

[141] Dong Zheng, Weiyan Ge, and Junshan Zhang. Distributed opportunistic scheduling for ad-hoc communications: an optimal stopping approach. In *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, MobiHoc '07, pages 1–10, New York, NY, USA, 2007. ACM.

[142] Peng Xuan Zheng, Ying Jun Zhang, and Soung Chang Liew. Multipacket reception in wireless local area networks. In *Communications, 2006. ICC '06. IEEE International Conference on*, volume 8, pages 3670 –3675, june 2006.