JOINT LINK/PACKET SCHEDULING, RATE ALLOCATION AND ROUTING OPTIMIZATION IN STDMA BASED WIRELESS MESH NETWORKS

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

AND THE INSTITUTE OF ENGINEERING AND SCIENCES OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

By

Ahmet Emrah SEZGİN

August 2009

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Ezhan KARAŞAN(Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Nail AKAR

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. İbrahim KÖRPEOĞLU

Approved for the Institute of Engineering and Sciences:

Prof. Dr. Mehmet Baray
Director of Institute of Engineering and Sciences

ABSTRACT

JOINT LINK/PACKET SCHEDULING, RATE ALLOCATION AND ROUTING OPTIMIZATION IN STDMA BASED WIRELESS MESH NETWORKS

Ahmet Emrah SEZGİN
M.S. in Electrical and Electronics Engineering
Supervisor: Assoc. Prof. Dr. Ezhan KARAŞAN

August 2009

Wireless Mesh Networks (WMN) are a promising solution for next generation wireless access networks since they cherish benefits of having a reliable backbone in an ad hoc nature. This thesis investigates the joint scheduling and routing problem in minimizing the maximum delay required for delivering a given packet traffic to the intended destinations in spatial reuse Time Division Multiple Access (STDMA) based single channel multi-rate WMNs. Firstly, an Integer Linear Programming (ILP) model is developed by considering the constraints inherent to the problem, such as the Signal-to-Noise and Inteference Ratio (SINR) and capacity. The model is then improved by using additional constraints, such that these constraints do not affect the final solution but narrow down the search space. Because of the computational complexity, this ILP based optimization model is limited to small sized networks. We then consider different approaches to obtain sub-optimum solutions and find good lower and upper bounds on the optimum solution. LP and Lagrangian relaxations are used for obtaining lower bounds. Improving LP relaxation with cutting planes and Lagrangian relaxation, we obtain lower bounds that are up to 100% tighter than the simple LP relaxation. Next, a greedy heuristic approach is employed as an upper bounding technique. Tabu Search technique is implemented to improve the upper bound provided by the greedy approach, and around $10 - 20\%$ tighter bounds are obtained. The sub-optimum solutions obtained by using the heuristic Tabu Search algorithm are shown to provide maximum delays that are within $10 - 50\%$ of the bounds obtained by using the cutting planes and Lagrangian relaxation for the networks considered in this thesis.

Keywords:Wireless mesh networks, joint routing and scheduling, Integer Linear Programming, Lagrangian Relaxation, Tabu Search

ÖZET

STDMA TABANLI TEK-KANALLI KABLOSUZ ÖRGÜ AĞLARDA BİRLEŞİK LİNK/PAKET PLANLAMASI, HIZ ATAMASI VE YÖNLENDİRME EN İYILEMESI

Ahmet Emrah SEZGİN
Elektrik ve Elektronik Mühendisliği Bölümü Yüksek Lisans
Tez Yöneticisi: Doç. Dr. Ezhan KARAŞAN
Ağustos 2009

Güvenilir bir alt yapının yanısıra tasarsız bir tabiata sahip olmaktan faydalanan kablosuz örgü ağlar, yeni nesil kablosuz erişim için umut vadeden bir yaklaşımdır. Bu tez, STDMA tabanlı tek-kanallı/çoklu-hızlı kablosuz örgü ağlarda, belirlenen bir paket miktarının, istenen hedeflere ulaştığı zamanın enküçültmesinde birleşik planlama ve yönlendirme problemini inceler. İlk olarak, enterferans ve kapasite gibi probleme içsel sınırlamalar göz önünde bulundurularak, bir tam sayı doğrusal program çerçevesi modellendi. Bu model, daha sonra, en son sonucu etkilemeyen ancak arama uzayını daraltıp hesaplama hızını arttıran ek kısıtlamalarla iyileştirildi.Ancak, hesaplamaya dayalı zorluk dolayısıyla, belirtilen eniyileme, göreli olarak küçük boyutlu ağlarla sınırlıdır. İkinci olarak, eniyiye alt ve üst sınırlar elde etmek için buluşsal yöntemler uygulandı. İlk olarak LP ve Langranj gevşetmeleri denendi. LP gevşetmesini kesen düzlemlerle, Lagranj gevşetmesini de probleme uygun değişikliklerle iyileştirdiğimizde, birçok durumda 100%'e yakın daha sıkı alt sınırlar elde edildi. Bundan sonra, açgözlü tekrarlamalı bir çözüm yöntemi, üst sınır aramada önerildi. Bu açgözlü tekrarlamalı yöntemle edinilen üst sınır, Tabu arama yöntemi uygulanarak iyileştirildi,

ve $10-20\%$ etrafındaki oranlara varan daha sıkı üst sınırlar elde edildi. Bulunan üst sınırların, bu tezde incelenen ağ örneklerinde alt sınırlara $10-50\%$ arasında değişen yakınlıklarda olduğu gözlemlendi.

Anahtar Kelimeler: Kablosuz Örgü Ağlar, birleşik yönlendirme ve planlama, tam sayı doğrusal programlama, Lagranj gevşetmesi, Tabu araması

ACKNOWLEDGMENTS

I am sincerely grateful to Assoc. Prof. Dr. Ezhan Karaşan for his supervision, guidance, insights and support from selection of the topic, to the determination of the methodology, throughout this work. His broad vision and profound experiences in engineering has been an invaluable source of inspiration for me.

I would like to thank to the members of my thesis jury, Assoc. Prof. Dr. Nail Akar and Asst. Prof. Dr. İbrahim Körpeoğlu for reviewing this dissertation and providing helpful feedback.

I would like to especially thank to my family for their love and supports my whole life.

I would like to acknowledge Fazlı Kaybal for useful discussions, friendship and contributions to my work, Mehmet Can Kerse, Kemal Gürel, and Şeref Burak Selvi for their friendship, support and motivation, Halil İbrahim Şahin for his friendship, support and trust in me.

Many thanks to Altan Toksöz, Ömür Arslan, Vahdettin Taş, Murat Celal Kılınç, Mehmet Akif Koca, Ahmet Güngör, Mehmet Köseoğlu, Osman Gürlevik, Fırat Karataş and Volkan Açıkel for their friendship.

Financial support of The Scientific and Technological Research Council of Turkey (TUBITAK) for the Graduate Study Scholarship Program is gratefully acknowledged.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

The huge success of Wireless Local Area Networks led to the great interest in wireless networking technologies. Significant research has been done on Wireless Ad Hoc Networks and more recently on their successors, Wireless Mesh Networks. Wireless Mesh Networks (WMNs) have emerged as one of the enabling network technologies, and become a promising approach for next generation wireless access networks. WMNs gained particular interest as they are built over a multi-hop ad hoc-like backbone and cherish benefits of having a reliable backbone and an ad hoc nature.

Communications in a WMN are of multi-hop nature where radio nodes operate as routers, relaying packets on behalf of other nodes that may not be within direct wireless transmission range of their destinations. These routers interconnect and dynamically self-organize themselves so as to form a wireless backbone in a mesh structure. This backbone is a self-standing network offering inter-user connectivity, and if connection is available through the routers acting as gateways, it might be considered as a local wireless extension to the Internet. The routers, having gateway/bridge functionalities, will help WMNs integrate with the existing technologies like cellular, wireless-fidelity (Wi-Fi) and world-wide inter-operability for microwave access (WiMAX) networks. Thus, WMNs will greatly help clients to have connectivity anywhere and anytime.

One of the major problems in wireless mesh networks is the capacity reduction due to interference caused by multiple parallel transmissions [1]. Research on routing, channel access schemes and rate adaptation is targeted to improve the network performance severed by this problem. In addition to this, cross-interaction between different layers of the protocol stack (cross-layer design) is viewed as a promising technique in order to increase the network throughput by jointly allocating the network resources.

The physical layer in contemporary wireless communication devices provides multi-rate communication by implementing adaptive modulation and coding (AMC) techniques. According to AMC, in order to achieve a rate, SINR value at a receiver node must be above the defined SINR threshold corresponding to that rate. Certainly, high transmission rates require higher SINR thresholds than low transmission rates, therefore, require a careful planning.

In order to achieve high rates, and improve network throughput, medium access layer must be carefully planned. A widely used medium access scheme is the Time Division Multiple Access (TDMA) technique. TDMA is digital transmission technology that allows a number of users to access channel without interference by allocating unique time slots to each user within each channel. In wireless mesh networks, an extended version of the TDMA scheme, spatial reuse Time Division Multiple Access (STDMA), is used. The difference from the classical TDMA is that, STDMA enables more than one link in a single TDMA slot if the receivers at all links satisfy the requirements of the interference model adapted in the physical layer. Thus, link scheduling is a key aspect in STDMA based WMNs.

The route planning in the network layer is another crucial aspect in network performance improvement. This involves determining the best paths along which flow is routed from the source to the destination. Unlike ad hoc networks, in wireless mesh networks routers are stationary, forming a steady backbone. So, the routing is performed in the backbone, and this brings some additional advantages as the stability of mesh routers on the backbone level allows gathering some interesting statistics about the links qualities. These statistics lead to cross-layer development, which is a viable alternative for the design of architecture in wireless mesh networks.

The classical approach in network design is planning each layer separately. However, as wireless communications and networking occupy center stage in research and development in the area of communication networks, the suitability of this approach, the layered protocol architecture, is being questioned by the research community. Several arguments state that, the layered architecture is not as suitable for wireless networks as cross-layered designs.

There are several interpretations of cross-layer design in the literature. This thesis provides a cross-layer analysis involving physical, link, and network layers. Specifically, the problem of determining the shortest maximum delay within which a specified traffic can be jointly routed and scheduled among WMN nodes is analyzed. We study the joint optimization of link scheduling, rate allocation and routing in STDMA based single-channel multi-rate wireless mesh networks. The joint routing and scheduling design has the advantage of gaining inter-layer information and preventing sub-optimal solutions acquired by optimizing the two problems separately. Rate adaptation is the third aspect of the problem, and it has been used for some time now thanks to the application of adaptive modulation and coding.

The problem studied in this thesis, the determination of the shortest maximum delay, can be seen in scenarios such as VoIP applications which have maximum packet delay requirements. Therefore, in this thesis, the traffic is modeled as packets, rather than bit-rates, in parallel with such scenarios. The consideration of STDMA in the problem is realistic in WMNs that are built over the IEEE 802.16 technology. The obtained results can be saved in large look up tables to provide optimal link activation/routing patterns for known traffic scenarios.

Firstly, the problem is formulated through an Integer Linear Programming (ILP) framework, considering the characteristics of the Wireless Mesh Networks and the radio channel. In our framework, interference is considered as a major constraint. We use the physical interference model, introduced in [1], which requires that the signal to interference and noise levels (SINR) at the receivers of every ongoing transmission exceed certain thresholds that depend on link transmission rates.

Secondly, we implement heuristic approaches to solve for the problem for relatively larger networks. These heuristic approaches fall under two categories depending on their result with respect to the optimum solution: lower bounding and upper bounding techniques.

For lower bounding, we use Linear Programming (LP) and Lagrangian relaxations [2]. LP relaxation is strengtened using "additional cutting planes" that are satisfied by all integer feasible solutions. Significant improvement in the lower bound over LP relaxation is obtained using these cuts, and the resulting lower bound is as good as the one obtained by Lagrangian relaxation. Next, we suggest a modification in the Lagrangian relaxation to suit better to our problem, and the results seem promising as we observe lower bounds that are significantly tighter in many cases.

The next approach involves techniques to find an upper bound on the optimum (a suboptimum solution), and these techniques involve the greedy and "Tabu Search" [4], [5] solutions. The greedy solution is an iterative methodology where at each iteration an objective is maximized, but does not consider the overall optimization of the problem. The tabu search is a metaheuristic technique which iteratively explores the local search space in the neighborhood of a solution to move to a different solution until a stopping condition is reached. We use the outcome of the greedy approach as the initial solution to the tabu search. The upper bound obtained using this method is observed to converge to the lower bound, and they are noted within close proximity of each other.

The remainder of the thesis is organized as follows. In Chapter 2, a deeper insight and related work on routing, medium access, rate allocation and cross-layer design on wireless mesh networks will be given. In the following two chapters the solution methodology will be explained in detail. First the ILP model will be presented in Chapter 3, the (meta)heuristic approaches aimed at finding bounds will follow in Chapter 4. Finally, the conclusions will be drawn in Chapter 5.

Chapter 2

LITERATURE REVIEW

This chapter presents a brief introduction to Wireless Mesh Networks, and provides the motivation and background information with some of the related work on routing, medium access, rate allocation and cross-layer design. The chapter ends with the discussion of our contributions.

2.1 Wireless Mesh Networks

The last decade witnessed wireless networks evolving toward simple ways and better services for users to get connected while on the move. This evolution did not slow down and still carries on. In this perspective, Wireless Mesh Networks represents a key technology for the next generation wireless networks. Wireless Mesh Networks are an emerging two-tier architecture with two main objectives: (a) to offer connectivity to end-users; (b) to form a self-organized wireless backbone. These two characteristics make WMNs different from ad hoc and other infrastructure-based networks. Wireless mesh networks offer a dynamically self-organized and self-configured backbone where radio nodes automatically establish and maintain mesh connectivity among themselves. Communications in a wireless mesh network are of multi-hop nature where radio nodes operate as routers, relaying packets on behalf of other nodes that may not be within direct wireless transmission range of their destinations.

Nodes in a WMN consist of two types: wireless mesh routers, and mesh clients [6]. The mesh routers interconnect and dynamically self-organize themselves so as to form a wireless backbone in a mesh structure. This backbone is a self-standing network offering inter-user connectivity, and if connection is available through the routers acting as gateways, it might be considered as a local wireless extension to the Internet. These routers cover a region as access points supplying connectivity to the mesh clients in their region. The routers, having gateway/bridge functionalities, will help WMNs integrate with the existing technologies like cellular, wireless-fidelity(Wi-Fi) and worldwide inter-operability for microwave access(WiMAX) networks. Thus, wireless mesh networks will greatly help clients to have connectivity anywhere and anytime.

As their contribution to connectivity is significant, wireless mesh networks also take part in a number of applications as well. Among those applications are building automation, broadband home networking, community and neighborhood networks, video surveillance and perimeter security, mines and industrial sites, military communication, sports events, emergency and hostility formed networks, railway and highway corridors, VoIP phone applications [6].

According to the capabilities of the radio nodes, and in terms of the architectural design, wireless mesh networks can be classified into three categories: infrastructure, infrastructure-less and hybrid networks.

- Infrastructure based mesh networks are two-tier networks composed of a wireless backbone of interconnected self-organizing mesh routers. In this structure, end-users move freely in the environment and wireless mesh routers are fixed providing a fixed backbone. The backbone can be built using various types of radio technologies like IEEE 802.11 and IEEE 802.16. Mesh routers are responsible for routing as well as operating as access points and providing network access to mesh clients. The clients have no special routing capabilities.

- Infrastructure-less mesh networks are one-tier networks where client nodes are part of the routing architecture and mesh routers do not exist. They are basically wireless ad hoc networks.

- Hybrid mesh networks are the combination of infrastructure based and infrastructure-less networks. As well as having a mesh backbone consisting of mesh routers, mesh clients have routing capabilities. Mesh clients can access the network through mesh routers as well as directly connecting to other mesh clients.

Although these three types exist, the term wireless mesh networks generally applies to infrastructure based mesh networks which this thesis focuses on. This eliminates the complexity caused by dealing with high mobility hybrid and infrastructure-less states bring. Some of the problems when trying to improve performance of infrastructure based wireless mesh networks are channel access, routing and rate adaptation and this thesis tackles these problems.

2.2 Rate Adaptation In WMNs

The growing success of wireless networks is mainly due to the ratification of the standards like IEEE 802.11 that have led the way to readily available wireless devices that can give service in high data rates [7]. Rate adaptation is commonly used in contemporary wireless networks by dynamically switching data rates to match the channel conditions, with the goal of selecting the rate that will give the optimum throughput for the resources available [8], [9], [10].

Rate adaptation involves two aspects : estimation of channel quality and rate selection [7]. For estimation of the channel quality, the time-varying state of the wireless channel is measured and then future quality is predicted. The issues here are : selection of metrics to be used as indicators of channel quality (e.g., signal-to-noise-interference ratio, signal strength, symbol error rate, bit error rate), selection of predictors to be used, and decision of whether predictions should be short-term or long-term, etc. [11]. The second aspect of rate adaptation, rate selection, involves using the channel quality predictions to select an appropriate rate. A common technique is threshold selection, where the value of an indicator is compared against a list of threshold values representing boundaries between the data rates [9], [12].

In this thesis, signal-to-noise and interference ratio (SINR) at the receiver nodes is used as the channel quality indicator and it will be compared against a list of threshold values representing nominal data rates, and data rates will be modeled with constraints defining the number of information units (packets) that must be transmitted on each link per frame (TDMA slot).

2.3 Channel Access in WMNs

The wireless industry developed several access schemes to the shared wireless medium to improve network performance. The general approach categorizes the access schemes into two: contention based access, and contention free access [13]. A most widely used contention free channel access scheme which we will focus on is the Time Division Multiple Access (TDMA) technique. TDMA is digital transmission technology that allows a number of users to access a single radio-frequency (RF) channel without interference by allocating unique time slots to each user within each channel. The users transmit in rapid succession, one after the other, each using its own time slot which allows multiple radio nodes to share the same transmission medium.

This classical TDMA scheme provides rules for a group of nodes to activate links among each other in a small area. However, in a wireless mesh network, depending on the position of nodes and propagation conditions, a larger area is to be considered. Spatial reuse Time Division Multiple Access (STDMA) is one such scheme proposed for this problem. Considering the area the network takes up as well as the constraints in link activation, some transmitter-receiver pairs share the time slot in STDMA instead of a single pair. A link between two radio nodes exists if transmitting a signal from the transmitter, the Signal-to-Noise-and-Interference Ratio (SNR) at the receiver is sufficiently high to correctly decode the signal. Moreover, when parallel transmissions allowed by the STDMA scheme are inserted into the picture, the major problem is that these parallel transmissions generate interference at receiving nodes that may affect the correct decoding. Therefore, some constraints on resource reuse must be considered in order to guarantee correct network operation.

According to the well known classification proposed in [1], two possible interference models can be considered to provide guidelines for resource reuse: the Protocol Interference Model and the Physical Model. The Protocol Interference Model describes interference constraints according to a conflict graph. Not taking into account the cumulative effect of interference, it only defines basic communication and interference regions for nodes. If node i is transmitting to node j, no other node h can transmit if link $h \rightarrow j$ is in the conflict region defined by this interference model. On the other hand, the Physical Model directly considers SINR constraints at receivers. Within a protocol interference model context, scheduling problem has some similarities with the graph coloring problem and several solutions have been proposed both for point-to-point and broadcast/multicast transmissions [14], [15],[16].

The protocol interference based models do not consider that the effect of several interfering transmissions is cumulative. Thus, the carrier sense threshold must be set to a quite low value and this results scheduling policies not to be very efficient [12]. To take into consideration the cumulative effect, physical interference model is a more suitable and practical approach as it is based on SINR levels at receivers. Since SINR levels are closely related to the transmit power levels, joint scheduling and power control are quite popular [17], [18].In [19], a scheduling and power control strategy is proposed. The main objective is the maximization of network throughput. The solution approach proposed does not carry a joint framework, but rather is based on the decomposition of the problem into two sub-problems. The first sub-problem has the goal of finding the maximum subset of parallel transmissions where transmitting nodes are separated by at least a given distance, which in other words mean a protocol interference model is used. The second sub-problem sets powers according to a scheme. The work presented in [17] is similar with the addition of using a greedy approach and SINR constraints.

Addition of rate adaptation to the scheduling problem is also quite popular. Approaches similar to this thesis are presented in [20], [21], and [12]. A mixed integer linear programming (MILP) formulation for the joint scheduling, power control, and rate control problem is proposed in [20]. With goal being throughput maximization, the solution approach is based on a greedy algorithm. A-two step solution approach to joint scheduling and rate adaptation problem is presented in [12]. First, a MILP formulation is provided pursuing an optimal solution, and as the problem proves hard to solve, heuristics are applied to find bounds to the solution as a second step.

2.4 Routing in WMNs

The routing problem is determining the paths along which flow is routed from the source to the destination. Being their descendants, routing protocols designed for ad hoc networks provide a basis for the ones in wireless mesh networks. Dealing with a dynamic mobile environment, these protocols operate on hop count as a metric. Since there is no infrastructure, clients themselves are responsible for routing as they form the network. Some examples of such protocols designed for ad hoc networks include Optimized Link State Routing [22], Adhoc On-Demand Distance Vector (AODV) [23], and Dynamic Source Routing (DSR) [24] protocols. These protocols designed for ad-hoc networks are rather complex as they need to deal with the possibility of highly mobile nodes that the ad-hoc structure applies. There has, therefore, been significant work on route discovery and maintenance. Although ad-hoc networks' routing algorithms can be directly applied in WMNs, the relatively stationary topology of WMNs suggests that much more simplified routing algorithms could be developed. Further, the traffic pattern in WMNs is such that alternate protocols are likely preferred [25]. The major difference between infrastructure based mesh networks and ad hoc networks from routing point of view is that in mesh environments routers are stationary forming a steady backbone. The routing is performed in the backbone level of the two-tier architecture among mesh routers, and mesh clients do not have a participation in the process. The work of Raniwala [26] presents this kind of a protocol. Dealing with the backbone has some additional advantages as the stability of mesh routers on the backbone level allows gathering some interesting statistics about the links qualities that can be then used as metrics to improve routing performance [27], [28], [29]. This essentially leads to cross-layer development which this thesis focuses on.

In [25], routing protocols for wireless mesh networks are investigated and they also propose algorithms of their own to minimize the packet delay. Packet delay is caused by various reasons, including collision resolution during packet forwarding, packet buffering, and different scheduling algorithms. They start with the assumption that a smaller number of hops would lead to less packet delay, so to minimize the packet delay the shortest path is used. However, it is observed that collisions must be minimized, as highly-contended paths that are shortest showed not to be necessarily ideal [30]. In order to minimize collisions it is concluded that routing could benefit from jointly considering the scheduling.

2.5 Cross-Layer Optimization In WMNs

Traditional packet-based network architectures assume that communication functions are organized into nested levels of abstraction called protocol layers.This layered architecture divides the overall networking task into layers and defines a hierarchy of services that will be provided by each of the layers [31]. In fact, the network functionalities and services are commonly classified and modelled through the well-known Open System Interconnection (OSI) network model. The ISO/OSI model, which was standardized in 1984, establishes a 7-layer protocol stack where each layer defines the specifications for a particular network aspect (medium access, congestion control, etc.) and provides services to the upper layers. In this model, the services are realized by designing protocols for the different layers. The architecture forbids direct communication between non-adjacent layers, only communication between adjacent layers is possible and limited with standard interfaces such as procedure calls and responses.

In this kind of layered architectures, the designers have two choices when designing a protocol. The traditional choice is designing the protocol by strictly regarding the classical layered architecture. This means that protocols will be designed for specific layers and the designer needs to be concerned only about the layer in process. There is no need for a detailed information about what goes on in lower layers. Certainly, same applies to the upper layers only standard responses will be carried on to the upper layer and the detailed information will be kept inside the layer.

On the other hand, protocol can be designed by letting some violations on the layered architecture. Rather than following the general rules of the layered architecture, a layer can share information with other layers and work according to the feedback from those layers Figure. This approach falls under the category of cross-layer design and has been quite popular in wireless sensor networks [32], [33], ad hoc networks [34], [35], and more recently in wireless mesh networks [36], [37], [38].

Joint consideration of the physical and the network layers is a basic example of cross-layer approach for multi-rate networks [39].The role of the routing protocol is to select from the set of available links to form a path from source to destination. While in single-rate networks all links are equivalent, multi-rate networks support each available link to be able to operate at a different rate. Thus the routing protocol is presented with a much more complex problem when dealing with rate adaptation. It has to choose between sets of trade-offs. Long distance links can cover the distance to the destination in few hops, but then the links would be forced to operate at a low speed. Short links can operate at high rates, but more hops are required to reach the destination. In addition, the path selected by the routing protocol will not only affect the packets moving along that path, but will affect the link activations as well as their transmission rates within the interference range.

Cross-layer optimization of rate allocation and link scheduling is another basic approach. Some studies on this kind of cross-layer solutions include [20], [21], and [12]. A mixed integer linear programming (MILP) formulation for the joint scheduling, power control, and rate control problem is proposed in [20]. With goal being throughput maximization, the solution approach is based on a greedy algorithm. A two step solution approach to joint scheduling and rate adaptation problem is presented in [12]. First a MILP formulation is provided pursuing an optimal solution, and as the problem proves hard to solve heuristics are applied to find bounds to the solution as a second step.

A general observation regarding cross-layer optimization is that the efficiency of the scheduler depends on the efficiency of the routing protocol. A scheduling strategy that is supposed to operate well can achieve very poor performance if the routes are not optimized [30]. This brings in mind the idea of addressing link scheduling and routing jointly. This means routes will not be defined prior to scheduling and they will be derived from the link activation pattern over time. Adding rate allocation to this joint routing and link scheduling framework we obtain the problem addresed in this thesis. There are several works that also add channel assignment to this problem [40], [16], [41]. However, the objective of this thesis is to provide and solve an optimization framework, which with the addition of channel assignment becomes intractable with the state of the art MILP solvers. Thus, we stick with a single channel framework.

As this thesis focuses on joint routing, link scheduling and rate adaptation in a single channel framework, it serves in the same category as the works of [42], [43], [38], [44]. In [43], first an integer linear programming model for gateway placement is generated. Next, throughput maximization for known placements is targeted. They conclude that even for very small networks with very small traffic, the problem is infeasible to solve due to the extremely large number of variables. Carello et. al [42] and Molle et. al [43] reach to the same conclusion although the routes are almost previously defined and they suggest using column generation to provide good suboptimal results. The strategy followed in [38] is similar with the difference of using Genetic Algorithms to provide bounds for the optimum solution.

Although cross-layer approaches have received great interest recently, and their significance in providing improvement in terms of performance is widely accepted, there is also various research on the negative consequences of these approaches. These opposing works generally state that the layered architecture can be attributed a considerable measure of credit for the revolutionary success of the Internet. In [45], authors state that architecture violations introduced by cross-layer design undermines the importance of the architecture as architecture no longer represents the system. Moreover, too many violations over time will lead to architecture completely losing its meaning. Even though asserting the wireless medium is fundamentally different from the wired one, authors of [46] try to show that the conventional layered architecture provides a reasonable way to operate wireless networks. They demonstrate that good architectural design results in proliferation and longevity using some successful historical examples like von Neumann's architecture for computer systems, Shannon's architecture for communication systems and the plant controller feedback paradigm in control systems. Two of the problems associated with cross-layer design are tight-coupling and unbridled stack design.

The major element in layered architecture's success is the fact that it enables protocol designers to focus on specific layers without considering the rest of the protocol stack. In cross-layer designs, the protocols become tightly coupled, and highly dependent on each other [33]. This dependency may result in consequences as far as instability. Unbridled cross-layer design is another problematic issue [46]. Further improvements in the design will become difficult as it will be hard to address satisfactorily how a new modification will interact with the multiplicity of already existing dynamics. Second, it will be difficult from maintenance point of view. Modifying one layer or adding new layers will require changes in the whole system.

Although these negative aspects are valid in general, substantial and complicated interdependence among layers in wireless communication systems make the cross-layer approach a viable alternative for the design architecture of wireless mesh networks.

2.6 Contributions of the Thesis

In the last few years, wireless mesh networks (WMNs) have emerged as one of the enabling network technologies, and become a promising approach for next generation wireless access networks. The huge success of wireless LANs led to the great interest in wireless networking technologies. Significant research has been done on wireless ad hoc networks and more recently on their successors, wireless mesh networks. Wireless mesh networks gained particular interest as they are built over a multi-hop ad hoc-like backbone and cherish benefits of having a reliable backbone and an ad hoc nature. A series of issues related to the wireless mesh network research such as routing, medium access, rate adaptation, power and channel allocation are stated previously, and the main contribution of this thesis is targeting many of these issues at once.

This thesis provides a cross-layer approach to three major aspects present in wireless mesh networks: (1) link scheduling (medium access), (2) routing , and (3) rate adaptation (multi-rate). Specifically, the problem of determining the shortest deadline within which a specified traffic can be jointly routed and scheduled using multi-rate among WMN nodes is analyzed. Although we do not consider channel and power allocation issues, these three aspects we investigate constitute a detailed and comprehensive research problem and the optimization takes up a considerable time with state of the art solvers even for relatively small sized networks.

We propose a mixed integer linear programming model (MILP) and meta-heuristics to solve for joint routing, link scheduling and rate allocation problem in spatial reuse time division multiple access(STDMA) based wireless mesh networks. The proposition assumes a centralized solution to, as STDMA implies, a synchronized, time slotted network architecture. Another important assumption is that routes are not predefined as they are selected in conjuction with link activation/scheduling. Hence, multiple paths may exist to every destination.

The joint routing and scheduling design has the advantage of gaining inter-layer information and preventing sub-optimal solutions acquired by optimizing the two problems separately. For instance, we are informed about the traffic to a destination, the possible routes to that destination, which links to activate for those routes, what links are affected by the SINR caused by this activation, what other links can be activated concurrently without violating the SINR, etc... Rate adaptation is the third aspect of the problem, and it has been used for some time now thanks to the application of adaptive modultion and coding. Investigating the problem for a variety of power levels and considering these three aspects altogether, we make a series of decisions of whether to use concurrent active links with low rates, or to use nonconcurrent but fast links, whether to use multiple hops causing little interference or to open long range direct links, when and which links to activate - at what rate as well, which routes to deliver a certain traffic and so on.

The main objective of the thesis is to determine the shortest maximum packet delay (in terms of STDMA slots) within which a specified traffic can be delivered to intended destinations, by using joint routing and scheduling among WMN nodes. We make observations for several traffic demands ranging from sparse traffic destined to a small subset of the network (a few gateways), to uniform traffic destined to every member of the network (all gateways).

The thesis mainly focuses on the determination of a joint link scheduling and routing pattern that provides the shortest maximum packet delay, and this problem can be seen in scenarios such as VoIP applications which have maximum packet delay requirements. Therefore, in this thesis, the traffic is modeled as packets, and the rates are considered as packets/slot (STDMA slot), rather than actual bit-rates, to have parallelism with such scenarios. STDMA is used as the medium access scheme, and it is realistic for WMNs that are built over the IEEE 802.16 technology. The results obtained by the approach followed in this thesis can be used to build large look up tables that provide optimal link activation/routing patterns for the delivery of the defined packet traffics.

Our solution approach is two-fold: (1) MILP solution (2) (meta)heuristics. The first part of our study considers an MILP solution approach to the proposed problem. We optimize network throughput by implementing rate adaptation, routing, and link scheduling altogether considering several constraints. These constraints stem from several factors such as the properties of the wireless medium, the resources involved (single channel, radio), and traffic demand. The problem targeted in this thesis do not extend to directional antennas, so transmission in wireless medium is of broadcast nature. This, in turn, creates a severe interference on other intended transmissions. Of the two popular interference models introduced in [1], we use the physical interference model, which requires that the signal to interference and noise levels (SINR) at the receivers of every ongoing transmission exceed certain thresholds that depend on link transmission rates. The MILP solution has shortcomings in terms of computation time and can only be applied to networks consisting of five to six routers with small traffic demands even with performance increasing modifications. This fact necessitated designing suboptimal heuristic solutions to the problem.

The second solution methodology involves using heuristics to find a range where the optimum solution lies. In order to do so, a value which we know is smaller than the optimum (lower bound) and one that is known to be greater (upper bound) must be found. There are several similar works that use heuristics, like column generation and genetic algorithms that achieve upper bounds [42], [38]. Other than that, greedy solutions, where we make an optimization at every slot separately, thus not an optimization of the whole problem, provide trivial upper bounds. These upper bounds are actually useful as they are feasible solutions to the problem as well. In this thesis, we implemented a tabu search mechanism to improve the greedy solution and generate an upper bound. Typically $10-20\%$ improvement over the initial greedy solution is achieved using the tabu search.

Using relaxation techniques is another common method to achieve bounds on a hard optimization problem [3], [2]. Relaxation corresponds to allowing the solution to violate some constraints of the problem thus get an easier-to-find, infeasible solution which corresponds to a lower bound on the optimum solution. Since our aim is throughput optimization, which is achieved through minimization of time to deliver a traffic demand, our problem involves slot minimization. We investigated methods such as linear programming(LP), and Lagrangian relaxations for obtaining lower bounds. We also developed a relaxation similar to Lagrangian but more suitable to the proposed problem and achieved significant improvements over LP and classical Lagrangian relaxations.

In the following chapter, we introduce the MILP formulation to solve the joint scheduling/routing problem in STDMA based single channel/multi-rate wireless mesh networks.

Chapter 3

JOINT SCHEDULING and ROUTING OPTIMIZATION in WMNs

In this chapter,we formulate the optimization framework for joint scheduling/routing problem as a Mixed Integer Linear Program (MILP) in STDMA based single channel/multi-rate wireless mesh networks.

3.1 Problem Overview

The broadcast nature of the wireless medium is the core limitation in wireless networks. Even when we ignore power efficiency issues, one cannot use high power to open high-speed links, because while improving signal-to-noise-ratio in that particular communication, using high power leads other communications to deteriorate. So, when and at what speed to open a link between two devices becomes a key issue in improving throughput in a wireless network. In a wireless mesh network, given enough transmit power, nodes can talk to each other directly, or use relay nodes in a multi-hop fashion. When the nodes to be used as relays are not predefined earlier, the routing becomes another problem to solve. This, together with the previous issue, constitutes the joint routing and scheduling problem.

The aim of this study is designing a joint routing and scheduling scheme using multi-rate communication such that traffic in a wireless mesh network is delivered in the smallest amount of time. Suppose we have a network consisting of a number of nodes. By nodes we mean the wireless mesh routers. These routers provide backbone access to clients in the neighborhood and are connected to each other in a multihop fashion. Therefore, these nodes are not mobile, but have static placements. We assume that the distances between every node pair are known. That way we can use the distances in finding the propagation loss, thus, signal to interference and noise ratios (SINR) for different communication scenarios. One important parameter in our problem is having multirate communication. In wireless communication a link between two nodes can only be active if the SINR threshold for communication is satisfied. This threshold is determined by the sensitivity of the radios, and is standardized for different communication technologies. In our case, a link can be active with a certain rate if the corresponding SINR threshold for that rate is satisfied. A node cannot transmit and receive at the same time since it has a single radio. Also having a single radio and a single channel, a node can only listen to a single node at any one time.

Given a network and a traffic demand, our aim is to find an efficient joint routing and link scheduling solution where each link is assigned a rate. Since each transmission causes interference on all the others, SINR is a major constraint in the problem. A second constraint is that a node can only receive or transmit in a time slot but not concurrently. The objective is to finish the traffic demand in the minimum number of slots.

3.2 MILP Formulation

After this brief introduction, the MILP formulation to solve this aforementioned problem is explained next.

3.2.1 Assumptions

The following assumptions are used in the mathematical models for the joint scheduling, rate allocation and routing problem presented in Section the formulation.

- The system operates in a time-slotted mode (TDMA).
- We consider a WMN with a fixed number of nodes, N.
- Nodes are static and their positions are known.
- We only consider large-scale fading when modeling the path loss.
- There is a single channel.
- Each node has a single radio.

3.2.2 Sets and Parameters

The following three sets are used in the formulation

The first set corresponds to the nodes denoted $\mathcal{N}$. The nodes selected from this set will be assigned transmitter-receiver pairs and destinations.

$$\mathcal{N} := \{1,...,|N|\}$$

The second set is used to define rates used in communication, denoted as $\mathcal{R}$. Essentially, we use the term rate as the number of packets that can be sent in a single slot.

$$\mathcal{R} := \{1,...,|R|\}$$

A rate is assigned to each connection in the network. Each assigned rate is selected from $\mathcal{R}$ and has to satisfy corresponding SINR threshold β_k given in the following table:

minimum SINR threshold (β_k)	R_k
63.0957	9
6.3096	4
1.9953	2

Table 3.1: *minimum SINR threshold vs. Rate for* 802.16*e with BER* $= 1e-4$

The final set used in our problem is the time slots $\mathcal{S}$.

$$\mathcal{S} := \{1,...,|S|\}$$

Next come the parameters used in the formulation.

$T \equiv [\ T_{ij}\ ]$ is the initial traffic matrix among nodes, where T_{ij} is the number of packets destined from node i to node j, $\forall i, j \in \mathcal{N}$, and $T_{ii} = 0\ \forall i \in \mathcal{N}$.

$D \equiv [D_{ij}]$ is the distance matrix among nodes, where D_{ij} is the distance between node i and node j, $\forall i, j \in \mathcal{N}$, and $D_{ii} = 0\ \forall i \in \mathcal{N}$.

$P \equiv$ Power is radiated from node, constant throughout the network.

$L \equiv [L_{ij}]$ is the path loss matrix among nodes, where L_{ij} is the path loss between node i and node j, $\forall i, j \in \mathcal{N}$, and $L_{ii} = 1\ \forall i \in \mathcal{N}$. L_{ij} is given as:

$L_{ij} \approx D_{ij}^{-\alpha}$, where α is the path loss exponent.

$N_0 \equiv$ Thermal noise at receivers

3.2.3 Variable Declarations

In the formulation, we need to use a set for defining which nodes are transmitting. Another set will be used to denote which nodes are receiving. If the nodes that receive the packets are not final destinations of the packet they will be used as relays. Therefore, we also need another set to denote the final destinations of the transmission. These transmissions will create SINR values for every other transmission taking place in the same time, also the movement of the traffic needs to be updated after each transmission, so we need another set for time. Finally, high SINR values will allow high rates in the transmissions, while low ones will decrease the rate. So a final set is used on rates. We have binary variable x that uses the information above:

$$x_{ijks} = \begin{cases} 1 & \text{if } \textit{node i transmits to node j with rate k at slot s} \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

The variable x is not enough to describe routing, so we need variable y to do that job:

$y_{ijms} \equiv$ number of packets sent from i to j destined to m at slot s.

The variable y just represents the delivered packets in a time slot, but we need variable u to keep track of the remaining number of packets:

$u_{ijms} \equiv$ number of packets destined to m in node i at the beginning of slot s.

Finally binary variable w is used to construct the objective which is minimizing the number of slots used.

$$w_s = \begin{cases} 1 & \text{if } \textit{slot is used, i.e., there exists activated links at slot s} \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

3.2.4 Constraint Inequalities

$$\sum_{j,k}(x_{ijks} + x_{jiks}) \leq 1 \tag{3.3}$$

$$x_{ijks} \leq \begin{cases} 1 & \text{if } \dfrac{P \cdot L_{ij}}{N_0 + \sum_{m,n,r} {}_{(m,n)\neq(i,j)} x_{mnrs} P_{mn} L_{mj}} \geqslant \beta_k \\ 0 & \text{otherwise} \end{cases} \tag{3.4}$$

$$\sum_{m} y_{ijms} \leq \sum_{k} x_{ijks} R_k \tag{3.5}$$

$$u_{ims} = u_{ims-1} - \sum_{j} y_{ijms-1} + \sum_{j:i\neq m} y_{jims-1} \tag{3.6}$$

$$y_{ijms} \leq u_{ims} \tag{3.7}$$

$$u_{im0} = T_{im} \tag{3.8}$$

$$w_s \geq \frac{\sum_i \sum_m u_{ims}}{M} \tag{3.9}$$

The single-channel, single-radio nature of the problem forces us to constraints (3.3) which means that a node can only transmit/receive to/from a single node with a certain rate in a certain slot.

The physical characteristics of the problem forces us to (3.4). That is, when a node starts transmission, it radiates electromagnetic waves out in the open, and unwanted nodes may hear it. How much of it they hear is modeled by the so called "physical interference model" that we use in our problem. Suppose node i wants to transmit to node j at rate k. For that to happen the signal to interference ratio at receiving node j must be over the threshold β_k corresponding to rate k. So, if a second node pair, m and n, have an ongoing transmission, m's interference on j must not deteriorate the SINR value in node j too much. Node m's interference on node j is simply the transmit power of node m weighted by the path loss from m to j. The same applies to the received signal from i. If the received signal's ratio to interference and noise summed up is greater than the threshold, than that communication scheme is achievable for that transmitter i, receiver j, rate k, slot and co-existing transmission m to n. To linearly express this above statement we multiply the denominator with the threshold and adding some new terms, we obtain :

$$x_{ijks}P \cdot L_{ij} \geq N_0\beta_k + \beta_k * P \sum_{m,n,r:(m,n)\neq(i,j)} x_{mnrs}P_{mn}L_{mj} - M(1 - x_{ijks}) \quad (3.10)$$

where M is a large number which guarantees that the SINR constraint is only valid when the link is active. The capacity constraint of (3.5) states that the number of packets transmitted over a link in a slot is limited by the rate of communication assigned to that link in that particular slot.

Inequalities (3.6), (3.8), (3.7) describe the routing of packets. As we stated before the system is a multi-hop network. Packets may need to visit intermediate nodes in order to reach their final destination. (3.6) state that if packets sent from a node to another are actually intended for that receiving node, those packets are done. However, if the receiving node is not the final destination for that packet then it becomes a relay in that transmission. It is the receiving node's duty to send those packets to their final destinations. (3.8) simply sets the number of packets in the initial slot, while (3.7) ensures that the number of packets transmitted by a node in a time slot does not exceed the number of packets in that node at the beginning of the time slot.

Our objective is to finish the traffic in the least time possible. We defined variable w just for that purpose. We need to get w to zero as soon as possible, in other words all traffic is delivered. (3.11) comes at this point, it suggests that as long as there are packets to be delivered to their respective destinations in the network w cannot be zero. Our objective function is:

$$N_{mip} = \min \sum_{s} w_s \quad (3.11)$$

All nodes in the network use the same transmit power. The transmit power is another parameter. However, the problem is investigated for three different transmit power levels. The lowest level is the minimum power such that connectivity in the topology is ensured. Highest level ensures every node can talk to others in fast rates.

This ends the discussion on the model used. Using this model formulation, we solve the joint routing and scheduling problem in multirate networks for up to five nodes. The computation time increases enormously for larger networks. To be able to solve for these bigger topologies, improvements in the original formulation are necessary.

3.3 Additional Constraints

To speed up the process, one alternative is to use additional constraints that would help the solver do its job faster. These additional constraints should be such that they do not affect the final solution but narrow down the search space. These constraints are commonly referred to as cutting planes as they cut off a portion of the search space that only includes infeasibile points.

The first cutting plane that we use is defined as the link/rate elimination. Making an offline calculation over all possible node pairs, we check if each node pair can communicate at a certain rate. If they cannot, we set the x variable corresponding to that situation to zero. Interference is ignored in this calculation where only the noise is taken into account. The resulting constraint can be written as:

$$\frac{P \cdot L_{ij}}{N_0} \leq \beta_k \rightarrow x_{ijks} = 0 \tag{3.12}$$

At low power levels, this constraint eliminates lots of node pairs.

The second type of cutting planes also involves eliminating certain links/rates, but it is a bit different from the previous cut. The difference is that, this time interference is taken into account, and we do not set no variables to zero. Suppose we are checking node i's transmission to node j.

$$\frac{P \cdot L_{ij}}{N_0 + P \cdot L_{m_1 j}} \leq \beta_k \Rightarrow x_{ijks} + \sum_n \sum_r x_{m_1 nrs} \leq 1 \tag{3.13}$$

This means if the SINR value at j in transmission $i \rightarrow j$ cannot be accomplished while $m1$ is transmitting, either transmission $i \rightarrow j$ or transmission $m1 \rightarrow \textit{somenode}$ has to cease. Eliminations are not limited to the above case. The above constraint can be generalized as follows:

$$\begin{aligned} \frac{P \cdot L_{ij}}{N_0 + P \cdot (L_{m_1 j} + L_{m_2 j})} &\leq \beta_k \Rightarrow \\ x_{ijks} + \sum_n \sum_r x_{m_1 nrs} + \sum_n \sum_r x_{m_2 nrs} &\leq 2 \\ &\cdots \\ \frac{P \cdot L_{ij}}{N_0 + P \cdot \sum_c L_{m_c j}} \leq \beta_k \Rightarrow x_{ijks} + \sum_n \sum_r \sum_c x_{m_c nrs} &\leq \frac{N}{2} \end{aligned} \tag{3.14}$$

where N is the total number of nodes. Adding these cuts, at some point we get thousands of extra equations to the original model, slowing down the process. However, they help the solver to find the solution to six-node topologies whereas without them we could only solve for five.

3.4 Numerical Results

This section includes some sample simulations for the MILP model presented in Sections 3.2 and 3.3. The modeling is done using GAMS with CPLEX as the solver. Simulation parameters are stated in Table 3.2. Four different traffic matrices are investigated for the two 6-node topologies given in Figures 3.1, and 3.2.

Table 3.2: *Simulation Parameters*

Parameters	***Values***
Traffic Matrix	1,2,3,4
Number of Nodes	6
Transmitting Power	Topology 1 :0.0106, 0.107, 2.7 mW
	Topology 2 :0.010, 0.05, 0.9 mW
path-loss exponent	2.5
N_0	10^{-6} mW
Rate vs. SINR Table	Table 3.1

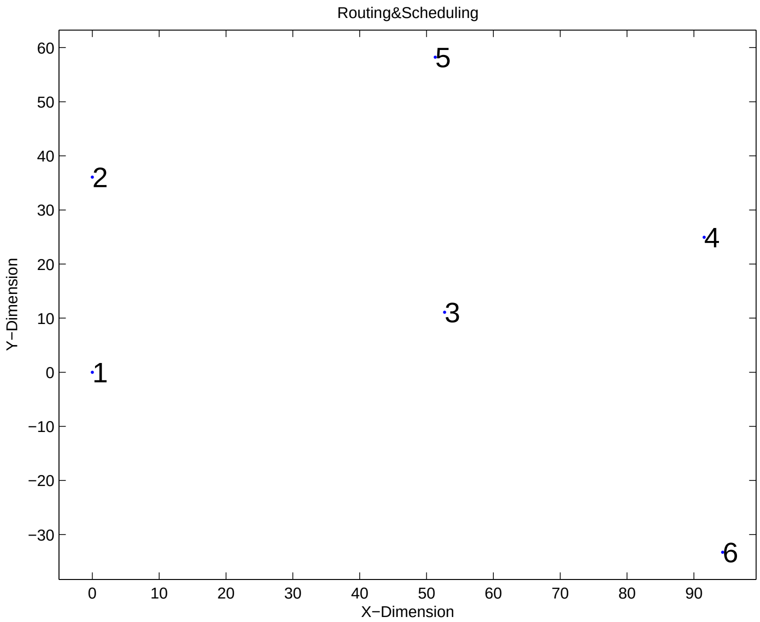

Figure 3.1: 6-node Topology 1

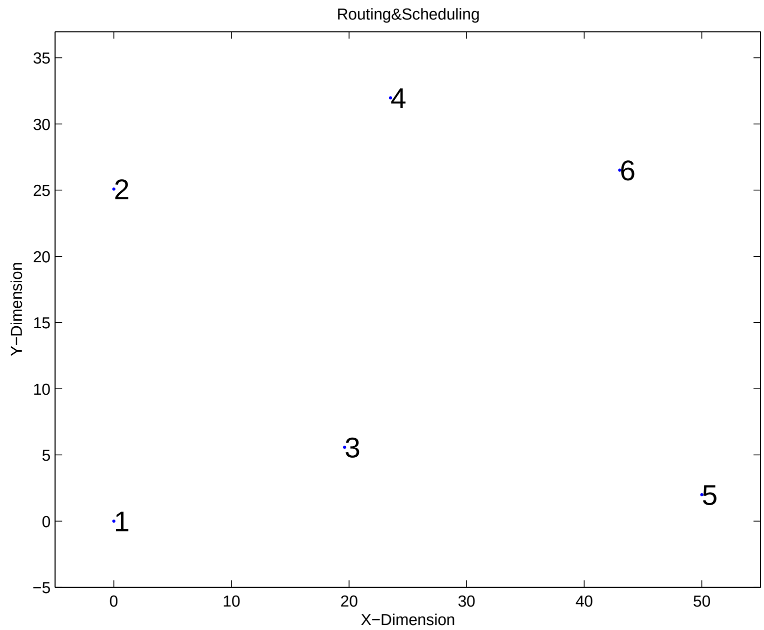

Figure 3.2: 6-node Topology 2

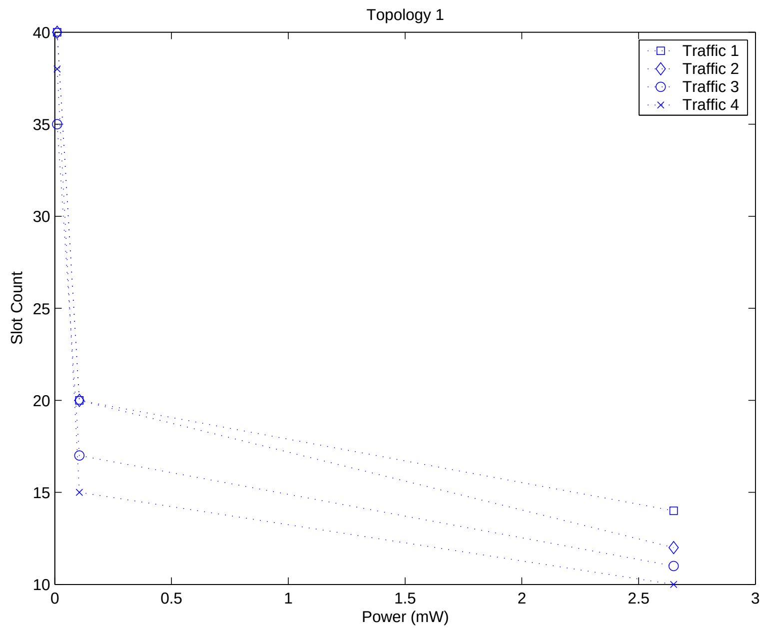

Figure 3.3: Slot counts vs power for the four traffic matrices

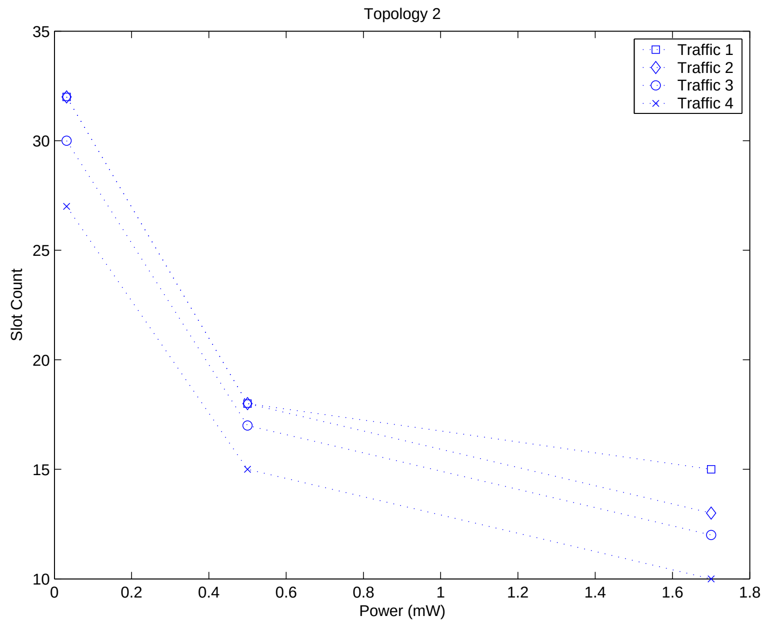

Figure 3.4: Results for the second topology

Three power levels are used in simulations, lowest power level provides just enough power to ensure connectivity while the highest one ensures high bit-rates for all possible links. Four traffic patterns are used in the simulations, they range from uniform traffic where all nodes have two packets to send to every other node, to traffic of 10 packets from all nodes destined to a single node (single gateway).

The number of necessary time slots for the two topologies as a function of the transmit power are given in Figures 3.3 and 3.4. We observed in the simulatons at both topologies and all traffic cases that as we increase transmitted power, the number of required slots to send all packets to their destinations decreases since increase in power enables of establishment high rate links.This makes sense as we deal with a small topology of 60x60m with a relatively small number of nodes to observe full impact of the concurrency. It is also observed that traffic-4 where all the traffic is destined to a single node (gateway) has the best performance in terms of slot count.

In the next chapter, the heuristic attempts at finding lower and upper bounds will be presented.

Chapter 4

Heuristic Solutions

The ILP formulation introduced in the previous chapter is a hard one to solve. Having thousands of integer variables the problem gets stuck as the size increases slightly. Estimating good upper and lower bounds on the number of necessary time slots is the issue covered throughout this chapter. First, the methodology to find lower bounds will be stated. Explanations for the methodology to get upper bounds will follow afterwards. Finally, numerical results will be presented.

4.1 Finding a Lower Bound

Using relaxation is a common method to achieve lower bounds on a hard optimization problem [2], [3], [48]. Relaxation corresponds to allowing the solution to violate some constraints of the problem thus get an easier-to-find, infeasible solution. Two relaxation methods linear programming (LP), and Lagrangian relaxations are investigated. As a third method, a relaxation scheme similar to Lagrangian but more suitable for the proposed problem is also developed.

4.1.1 Linear Programming Relaxation

Linear programming relaxation provides us with a lower bound by relaxing the integer variables to be floating points. This way the problem can be solved with very high speed. The lower bound obtained by LP relaxation alone is not a very tight one nonetheless. It seems as if the constraints we define in the problem formulation limit the integer solutions dramatically while not so much affecting the non-integer domain.

During relaxation it would help to diminish the range of values that the variables take. Rather than a whole range of decimals between zero and one for example, if decimals closer to zero are chosen for a variable meant to be zero in the integer case, and decimals close to one for variables meant to be one, the bound would become stricter than a normal relaxation. One way to achieve this is using additional equations in the model formulation. However, differently than the previous equations, these equations are not intended for speeding up the process. Using these equations, we intend to use the intrinsic relations between the decision variables to prevent them from taking a wide range of values. Following the definition of Geoffrion [3], a cutting plane is a linear equation that must be satisfied by all the feasible integer solutions to the problem, but violated by the optimum solution obtained by the usual LP relaxation.

Our aim, then, is to find cutting planes that would improve the lower bound.

Cut 1. $x - w$ relations :We defined w physically as the variable indicating the activity of the network at a certain slot. This activity was described in terms of the remaining packets in the system. If there were packets to send then the system had to be active, as system cannot be idle when there are remaining packets otherwise minimum time cannot be achieved. So, one can see the intrinsic relationship between variable w and the link activity variable x. This leads us to the cutting plane below:

$$w_s \geq x_{ijks} \forall i, j, k \tag{4.1}$$

This statement is a valid inequality to our integer problem. To show it is so, consider the following. If any one link is active in a slot, then this means packets are being sent over that link in that slot. If packets are being sent than by definition w has to be set to one in that slot. This inequality also states that when there are no packets in the system, no link can be active, which is obvious as it would serve no purpose. This equation is useful because with this equation we eliminate a range of decimal values w_s can take, at least in late slots. In the beginning of the problem w is close to 1 by its very definition, later on as remaining packets decrease the lower bound on w decreased too. However now we set the lower bound on supremum of x.

To further relate x and w a second inequality is defined:

$$w_s \leq \sum_{ijk} x_{ijks} \quad \forall i, j, k \tag{4.2}$$

Variable w is set to one if there are remaining in the system. If there are packets in the system some links may be active in order to achieve minimum time condition. If there are no packets in the system w is not set, and we do not say anything about variable x. Therefore, this inequality is valid for our integer problem.

Cut 2. $u - x$ relations : We know that when some links are active in the system, by definition packets are transmitted over that link. This transmission affects the remaining number of packets in the system, and this fact was included in the problem formulation by means of the transmitted number of packets, y.

However, the activity of the links, x, and the remaining number of packets, u, were not directly related and the following inequalities address this issue:

$$
\begin{aligned}
u_{ins} &\leq \sum_m \cdot (1 - \sum_{jk} x_{ijks}) + u_{ins-1} \\
u_{ins} &\geq \sum_m \cdot (\sum_{jk} x_{ijks} - 1) + u_{ins-1}
\end{aligned}
\tag{4.3}
$$

If link $i \to j$ is active for some j, then node i can only transmit and by single-radio constraint it cannot receive. The number of packets at node i decreases with respect to the previous slot, so the remaining number of packets at i destined to d can only decrease. If link $j \to i$ is active for some j, node i can only receive and cannot transmit. The number of packets at node i cannot decrease with respect to the previous slot, so packets destined to d also cannot decrease. If links $i \to j$ or $j \to i$ are not active for any j, then these inequalities do not bring out an extra limitation. Thus, the above inequalities are valid.

We continue relating x and u with the following:

$$
\begin{aligned}
u_{ins-1} - u_{ins} &\leq \sum_{jk} x_{ijks} \cdot R_k \\
u_{ins} - u_{ins-1} &\leq \sum_{jk} x_{jiks} \cdot R_k
\end{aligned}
\tag{4.4}
$$

If number of packets destined from node i to node d increased from the previous slot, then there exists some active link $j \to i$ for some node j. If number of packets destined from node i to node d decreased from the previous slot, then there exists some active link $i \to j$ for some node j. This concludes that above inequalities are valid for our integer problem.

Cut 3. $u - w$ relation: We defined the activity of the system, w, directly with the remaining number of packets, u, in our original problem formulation. However not considering the LP relaxation, the bound was somewhat loose between

u and w. Equations below strengthen the bound:

$$w_s \geq \frac{u_{ims}}{\sum_n T_{nm}} \quad \forall i, m, s \tag{4.5}$$

If there are any packets destined from some node i to some node m, then the system is active. This is actually close to how we defined system activity in the first place, and the inequality is valid. Notice, however, that w is now bounded from below by the supremum of u's normalized by initial traffic, rather than an average of u's as we did in the problem formulation. So, there are not as wide range of decimal values w can take, also considering (4.1), the same also applies to x.

This ends the discussion on cuts and linear programming relaxation. Next we move on with another method to obtain a good lower bound.

4.1.2 Lagrangian Relaxation

The Lagrangian Dual Problem

Using Fisher's words [2] , Lagrangian relaxation is a technique that is based on the idea that many hard integer problems can be viewed as easy problems complicated by a relatively small set of side constraints. Getting rid of these side constraints, we can get a lower bound on a minimization (or an upper bound on a maximization) problem with relatively less effort.

Consider the integer problem we defined in the previous chapter.

$$N_{MIP} = \min \sum_{s} w_s$$
$$subject\ to$$
$$A'z \leq b' \tag{4.6}$$
$$where\ z = x, y, u, w.$$

The Lagrangian relaxation depends on the observation that some constraints among $A'z \leq b'$ make the problem relatively harder than the other constraints. Without them the problem could be solved with ease. However, these constraints are not removed from the problem, we get an infeasible solution. It also provides a lower bound on the optimum solution since we relaxed a constraint.Consider (4.6) again as follows:

$$N_{MIP} = \min \sum_{s} w_s$$
$$subject\ to$$
$$\sum_{m} y_{ijms} \leq \sum_{k} x_{ijks} R_k \tag{4.7}$$
$$A''z \leq b''$$
$$where\ z = x, y, u, w$$

Here we identified the hard part of constraints and separated from others. From now on we use the notation:

$$\delta_{ijs} = \sum_{m} y_{ijms} - \sum_{k} x_{ijks} R_k \tag{4.8}$$

We can now construct the Lagrangian relaxation. The idea is simply separating the complicating constraints from others, and use those constraints in the objective:

$$\theta_\lambda = \min \sum w_s + \lambda^T \delta$$
$$A''z \leq b'' \tag{4.9}$$
$$\lambda \geq 0$$

Proposition 1. *θ_λ is a lower bound for the integer problem.*

Proof. The proof follows from the fact that λ are nonnegative. Since for minimization δ has to be negative it follows that

$$\min \sum w_s + \lambda^T \delta \leq \min \sum w_s = N_{MIP} \tag{4.10}$$

□

Since θ_λ provides a lower bound for all $\lambda \geq 0$, the best lower bound is obtained by the choice of λ that gives the optimal solution to the dual problem

$$\theta = \max_\lambda \theta_\lambda \leq N_{MIP} \tag{4.11}$$

Lagrangian relaxation, at each iteration solves for (4.9) with fixed λ, and then solve (4.11) with fixed w,δ and so on. We already constructed the model to solve for (4.9), however the dual problem defined in (4.11) is not straightforward. There are several algorithms to solve the dual problem. They vastly fall into two categories which use subgradients and cutting planes.

Methods to Solve the Lagrangian Dual

Cutting Plane Method. Previously, when trying to improve the linear programming relaxation we used cutting planes to eliminate some non-integer solutions without violating integer ones. Cutting planes can also be used to solve for (4.11). The most common method that uses this approach is Kelley's Cutting Plane Method [48]. The idea is generating cutting planes using the results of the minimization problem of (4.9) and at each iteration a new plane is added to the model. This formulation, used as it is, leads to instability and did not perform well in our problem.

Subgradient Method. Subgradient algorithm takes the current iterate λ_k in the direction of the current subgradient with a stepsize t_k. The simplest approach is to select a step size and decrease or increase it according to the performance of the method within a period of time.However, the convergence of the algorithm depends on the choice of t_k. In [49], Goffin suggests that the subgradient algorithm converges to maximum θ if

$$\begin{gathered}\sum_k t_k = \infty \\ and \\ t_k \rightarrow 0\end{gathered} \tag{4.12}$$

So any sequence satisfying these conditions would work, but the speed of convergence may not be good enough. Therefore, a practical approach in selection of t_k is presented in [49]:

$$\begin{aligned}\lambda_{k+1} &= \max(0, \theta_k + t_k * \delta) \\ where\ t_k &= \mu_k * (N_M IP - \theta_{\lambda_k}) \\ where\ \mu_k &= \frac{\mu_k}{2}\ if\ no\ increase\ in\ 3\ iterations\end{aligned} \tag{4.13}$$

The subgradient algorithm as presented above converges to a maximum value, nonetheless this value was observed to be no better than the LP-relaxation. So slight modifications were done on the model as suggested in [7]. The modifications were based on the observation that the algorithm has a Markov nature as it is. If the Markov nature is changed, by considering not only the last step but previous steps as well than the method was proposed to work better. The method did not improve by these modifications.

4.1.3 Modified Lagrangian Relaxation

Lagrangian relaxation by itself did not provide good enough results. It needs to be modified according to the nature of the problem. What we did next was taking the constraint relaxation idea of Lagrangian relaxation and use it in a way that better suits the problem at hand. The idea is simple. We cannot solve for the integer problem because of the complicating constraints. We relax the problem by moving these complicating constraints to the objective function. However, different than normal application of Lagrangian relaxation, we only add these constraints as penalties, but not as rewards. Normally, in Lagrangian relaxation if constraints are violated it exacts a penalty on the objective, and a reward when they are satisfied. By only exacting penalty on the objective we get values bigger than Lagrangian relaxation and make the problem harder as well, in fact it gets hard enough that we cannot iterate it as much as a subgradient method.

The following represents the mathematical model corresponding to the statements above:

$$\begin{aligned} \theta_\lambda = \min \sum w_s + \sum_{ij} \gamma_{ijs} \\ \gamma_{ijs} \geq \lambda_{ijs} * \delta_{ijs} \ \forall i, j, s \\ \gamma_{ijs} \geq 0 \ \forall i, j \ , \ s = 1, 2, .., s'. \\ for\ an\ experimental\ s'. \end{aligned} \tag{4.14}$$

Proposition 2. *θ_λ provides a lower bound on N_{MIP}.*

$$\theta_\lambda \leq N_{MIP} \tag{4.15}$$

Proof. In (4.14) $\gamma_{ijs} > 0$ for unsatisfied constraints and unlike ordinary Lagrange relaxation $\gamma_{ijs} = 0$ for satisfied constraints. Consider $s' \geq N_{MIP}$. Observe that minimum θ_λ in (4.14) equals N_{MIP} as all constraints must be satisfied for minimization. Next, consider $s' < N_{MIP}$. The logic stated above applies for $s \leq s'$. For $s > s'$, (4.14) is the same as the ordinary Lagrangian relaxation, providing a lower bound. □

As a matter of fact, for large enough λ and s', the relaxation equals the original formulation and

$$\max \theta_\lambda = N_{MIP} \tag{4.16}$$

However, we cannot exactly find the outcome of (4.16) as it would be to as hard as solving the original formulation. What we do to overcome this is as follows: Let $\lambda_{1:s'}$ =a big number for a suitable slot s' $\forall$ i, j and try to solve for the problem. This way found a lower bound on the optimum solution, and we search for a better bound with increasing s'. At one point, the problem will be hard to solve and get stuck. The result just before the computation gets stuck is taken as the new lower bound to the problem.

The selection of the initial value of slot parameter s' can be provided from LP or Lagrangian relaxations. This way, the method is guaranteed to improve the lower bound provided by these relaxations.

4.2 Finding an Upper Bound

The previous part investigated the methods to obtain a lower bound on the optimum solution. This section completes the picture by the search for an upper bound. Greedy solutions, where an optimization at every TDMA slot separately do not provide optimization for the problem. However, trivial upper bounds can be gathered from greedy solutions. These upper bounds are actually useful as they are feasible solutions to the problem as well. In this part, a tabu search mechanism is implemented on top of greedy solutions to generate additional upper bounds.

4.2.1 Greedy Solution

Consider the joint routing, scheduling and rate adaptation framework without time, or rather inside a single TDMA slot. Needless to say the objective for this scenario would not be the transfer of all the traffic, but maximization of some metric such as number of packets sent. Solution of this scenario would be rather fast compared to the MILP as a dimension is eliminated and the number of variables are reduced dramatically. The time will be incorporated to the system as an iteration parameter: for each and every slot until all the traffic is carried number of packets transmissible for that slot will be maximized and the parameters of the next slot will be handled. This scenario will definitely result in a feasible solution to the problem and is relatively easy to solve. However, maximizing the number of packets sent in a slot for each and every slot does not necessarily mean maximizing the throughput of the system. Hence, the outcome is an upper bound. This methodology is greedy in nature as each step takes up where the last step left and maximizes itself.

Certainly number of packets sent is not the only metric that can be applied in the maximization process during each slot. Any metric that would lead to finishing the traffic after a number of slots is an option. Some metric that come to mind easily include weighting the number of packets according to where they are sent. This way the scenario could favor fast direct links, or concurrent links, etc.

Next the explaination the MILP model used in slotwise optimization process is presented.

Sets. The same sets stated in Section 3.2.2 are used with the obvious exception of the set corresponding to the slot number.

Parameters. All parameters from Section 3.2.2 with the addition of traffic matrix as it is not a variable in this formulation:

$U_{im} \equiv$ packets destined to m residing in node i.

Variables. The same variables are used without the time dimension.

$x_{ijk} = 1$, if link i-j is active with rate k. X is binary.

$y_{ijm} \equiv$ packets sent from i to j destined to m.

The problem. Solve for

$$\max \sum_{ijm} y_{ijm} + \sum_{ijk} weight_{ijk} * x_{ijk} \tag{4.17}$$

subject to

$$\sum_{j,k} (x_{ijk} + x_{jik}) \leq 1 \tag{4.18}$$

$$x_{ijk} \leq \begin{cases} 1 & \text{if } \dfrac{P \cdot L_{ij}}{N_0 + \sum_{m,n,r} {}_{(m,n) \neq (i,j)} x_{mnr} P_{mn} L_{mj}} \geqslant \beta_k \\ 0 & \text{otherwise} \end{cases} \tag{4.19}$$

$$y_{ijm} \leq u_{im} \tag{4.20}$$

$$\sum_{m} y_{ijm} \leq \sum_{k} x_{ijk} R_k \tag{4.21}$$

This demonstrates the decision process for a slot. In order to induce it to the whole problem, the traffic matrix is updated for the next slot according to the activated link/selected route and the same MILP is solved for each slot until the traffic ends. Figure 4.2.1 illustrates this greedy methodology.

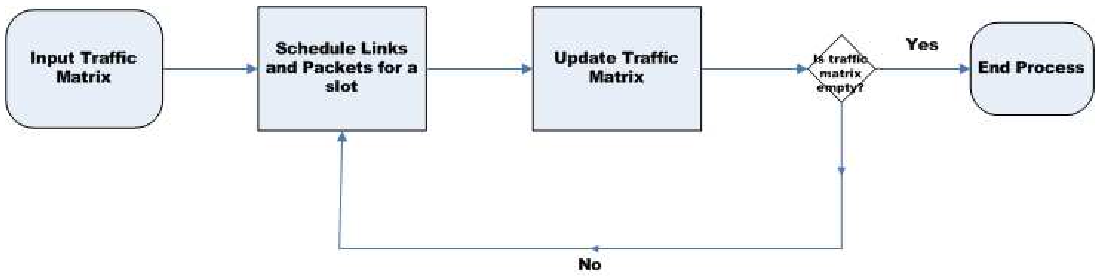

Figure 4.1: The flow-chart of the greedy solution methodolody

4.2.2 Tabu Search Solution

In this part we implement the Tabu Search technique that has been used for the difficult optimization problems. The Tabu Search technique, which with its various ingredients can be viewed as an engineer-designed approach: no clean proof of convergence is provided yet it has shown remarkable outcomes in various problems [50]. First presented by Glover [4],[5], tabu search is a metaheuristic used in solving combinatorial optimization problems, and it falls under the class of local search techniques. Tabu search iteratively explores the local search space in the neighborhood of a solution to move to a different solution until a stopping condition is reached.

A feature essential to tabu seach stems from the idea that for an efficient exploration of the search space, one needs to keep track of information to the exploration process rather than some local information like the value of the objective function. This information related to the exploration process is kept in a short-term memory defined as “tabu list”, hence the name “tabu search”. The tabu list contains previous actions that took place in the search procedure and are forbidden for some time. Thereby, the exploration process avoids from getting stuck in a local extremum and further searches the solution space.

The procedure we use in implementing tabu search in our optimization problem is as follows. As tabu search is a local search technique it must be started with an initial solution. The greedy solution explained previously is considered as such an initial solution. Other than being an initiation, greedy solution gives useful information to be used in the search procedure such as the set of links (as well as link n-tuples) that can be activated, set of routes between node pairs, etc. We use this information to narrow down the search space, and do not use constraints defined in the previous parts anymore. Although we lose several possible links and routes as all possibilities are definitely not covered in the initial solution, the computation time decreases significantly as the problem reduces to a much smaller size. This is quite similar to the column generaton idea used in [42].

Tabu search operates by moving from one solution to a neighbouring one. The neighborhood is defined as a single slot in our methodology. So, at each step of the search process a slot is selected and the links activated in that slot as well as routes corresponding to those links will be replaced by other possibilities. This replacement will be according to an objective function. The objective function used in this thesis consists of several parameters and is defined as:

$obj \equiv (w1 * par1_k + w2 * par2_k + w3 * par3_k + w4 * par4_k + w5 * par5_k) * par6_k$ for some experimental weights w1,...,w5.

- $par1_k$: number of packets that can be sent by move k.
- $par2_k$: number of packets available in the transmitters in move k.
- $par3_k$: binary parameter stating if the route selected has high speed links.
- $par4_k$: binary parameter stating if the route selected is favored by the initial solution
- $par5_k$: number of hops to the destination

- $par6_k$: binary parameter stating if the move is in the tabu list

The last parameter, $par6_k$, serves for one of the basic properties of tabu search, it ensures that the move selected is not in the tabu list. In the computations, $par1_k$ has the heighest weight, other parameters have relatively small weights that are adjusted experimentally.

The selection of which slot to change is another decision. We use two selections: random selection, and selection of a "bad" slot. A "bad" slot can be a slot where high rated links are opened but wasted due to small traffic demand or a slot with low rated link activations, etc. From these two selected slots, a search is initiated, and the slot whose search brings a better final result is accepted. The term "better final result" is a tricky one, as we change a single slot and control its overall effect to the solution (if it improved number of slots used). The effect of changing a slot to the current solution is hard to observe. As we altered the links/routes activated in that slot, a chain reaction is triggered. The routes/link activations in the forthcoming slots need to be checked. However, as this check would be arduous, we propose to generate the rest of the solution from that slot on using a greedy approach. As we only need to look at a small subset of possible links and routes obtained from the initial solution this would be rather fast. Thus, we implement an iterative tabu search scheme to improve the initial greedy solution. Algorithm 4.1 presents the steps involved in the tabu search process.

4.3 Numerical Results

This section includes some sample simulation results for the heuristics presented in this chapter. LP, Lagrangian relaxations and greedy heuristic are implemented in Gams with CPLEX as the solver. Tabu search is implemented in Matlab. Four

Algorithm 4.1 Perform Tabu Search

Step 1: Apply standard greedy solutions
  Step 1a: Obtain initial solution
  Step 1b: Obtain Set of Links/Routes that can be activated
Step 2:
  $current - solution \leftarrow initial - solution$
  $best - solution \leftarrow initial - solution$
Step 3: Initiate tabu search from $current - solution$.
  Step 3a: Select a random slot.
  Step 3b: Find the move maximizing objective in that slot.
    (objective includes tabu list check)
  Step 3c: Generate a greedy solution starting from that slot.
  Step 3d: Select a "bad slot".
  Step 3e: Repeat steps b-c for this slot.
Step 4: Compare the two new solutions.
  $current - solution \leftarrow better - of - the - two$
  Update tabu list of N latest moves with the current move.
if $slotCount_{current-solution} \leq slotCount_{best-solution}$ **then**
    $best - solution \leftarrow current - solution$
else
  if No improvement in last M iterations **then**
      $current - solution \leftarrow best - solution$
  end if
end if
if Iteration limit reached **then**
    end process
else
    Go back to Step3.
end if

different traffic matrices are investigated for the two 6-node, and two 8-node topologies given in Figures 3.1, 3.2, 4.2, and 4.3 respectively. A comparative analysis with MILP solutions of the previous chapter is presented.

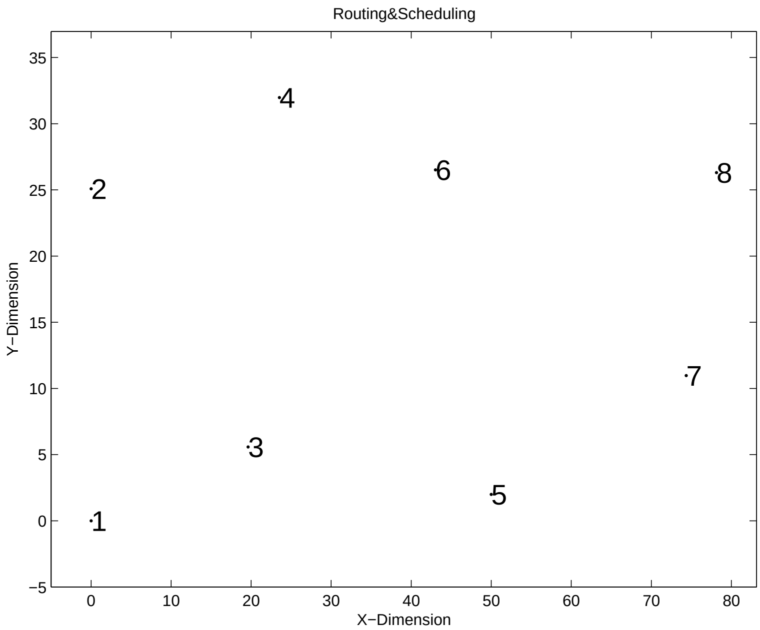

Figure 4.2: 8-node Topology 1

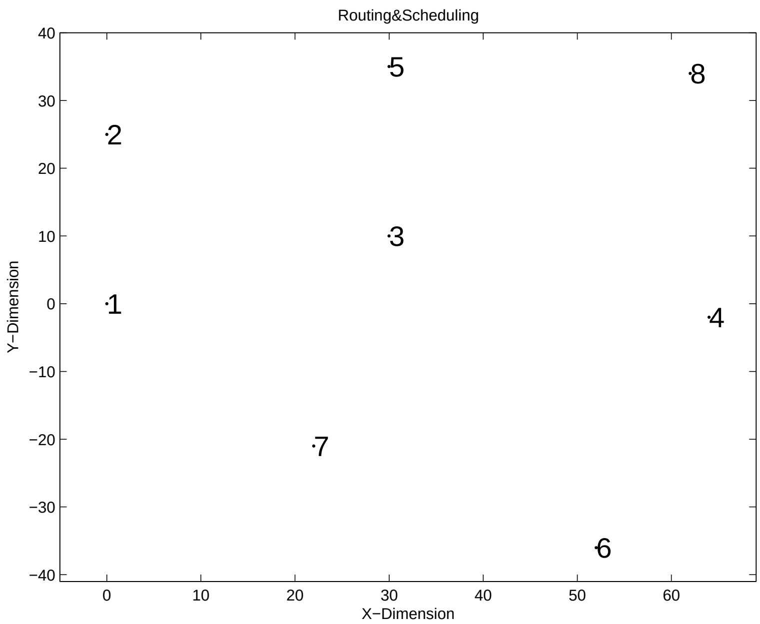

Figure 4.3: 8-node Topology 2

The parameters used in the simulations are stated in Tables 4.1 and 4.2. The traffic matrices are the same as in Section 3.4. Four traffic patterns are used in the simulations, they range from uniform traffic where all nodes have two packets to send to every other node, to traffic of 14 packets from all nodes destined to a single node (single gateway).

Table 4.1: *Simulation Parameters for 6 nodes*

Parameters	***Values***
Traffic Matrix	1,2,3,4
Number of Nodes	6
Transmitting Power	6-node Topology 1 :0.0504, 0.503, 2.6 mW
	6-node Topology 2 :0.010, 0.05, 2.7 mW
path-loss exponent	2.5
N_0	10^{-6} mW
Rate vs. SINR Table	Table 3.1

Table 4.2: *Simulation Parameters for 8 nodes*

Parameters	***Values***
Traffic Matrix	1,2,3,4
Number of Nodes	8
Transmitting Power	8-node Topology 1 :0.0503, 0.5029, 2.5143 mW
	8-node Topology 2 :0.0106, 0.053, 2.7 mW
path-loss exponent	2.5
N_0	10^{-6} mW
Rate vs. SINR Table	Table 3.1

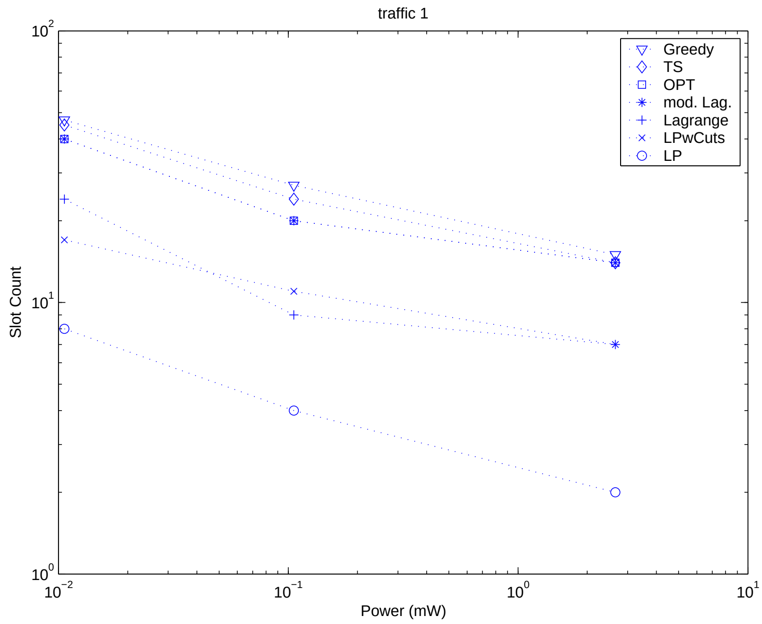

Figure 4.4: Outcomes of the methods for 6-node Topology 1, Traffic 1

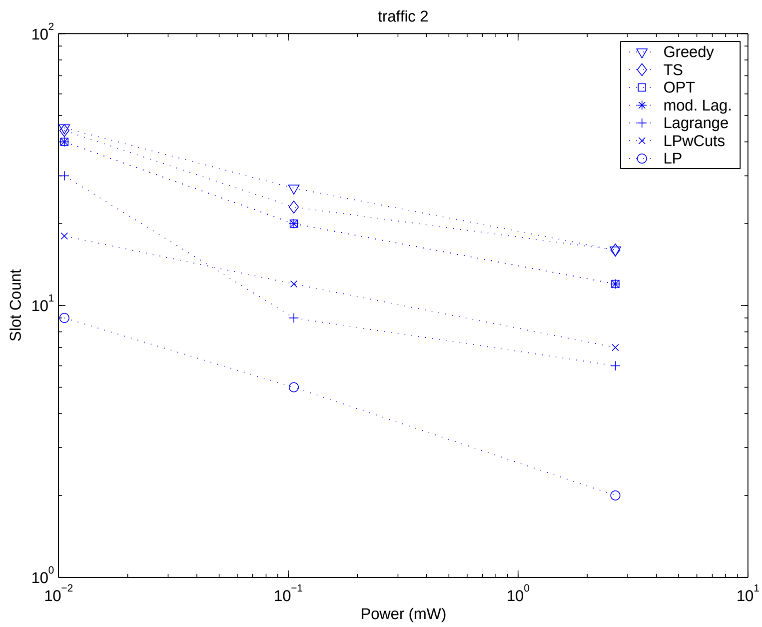

Figure 4.5: Outcomes of the methods for 6-node Topology 1, Traffic 2

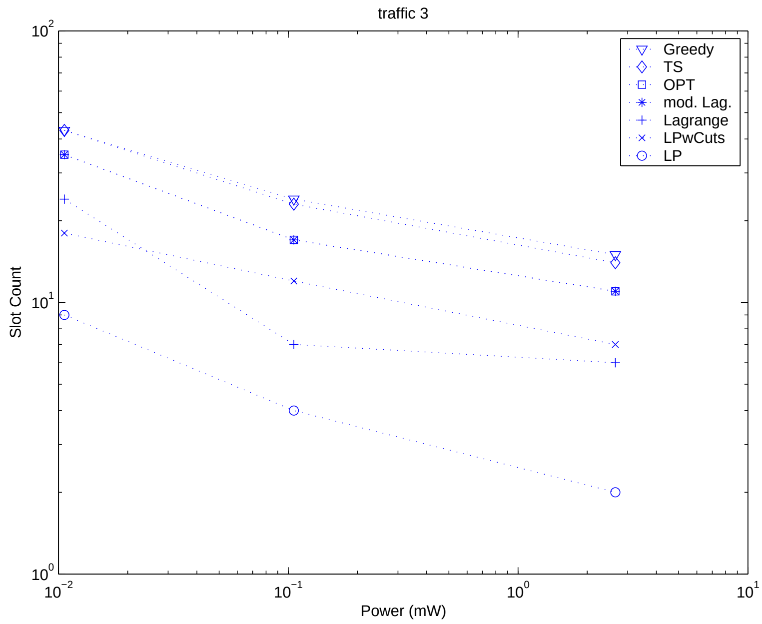

Figure 4.6: Outcomes of the methods for 6-node Topology 1, Traffic 3

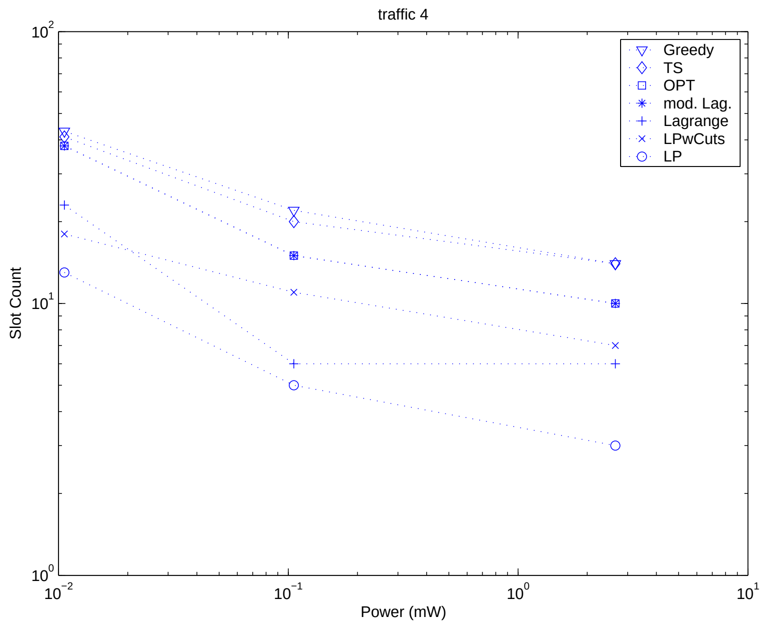

Figure 4.7: Outcomes of the methods for 6-node Topology 1, Traffic 4

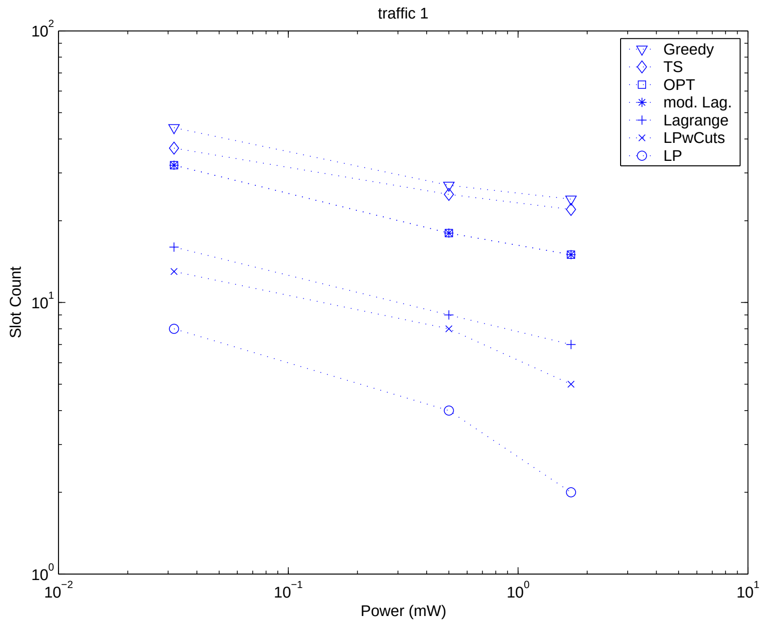

Figure 4.8: Outcomes of the methods for 6-node Topology 2, Traffic 1

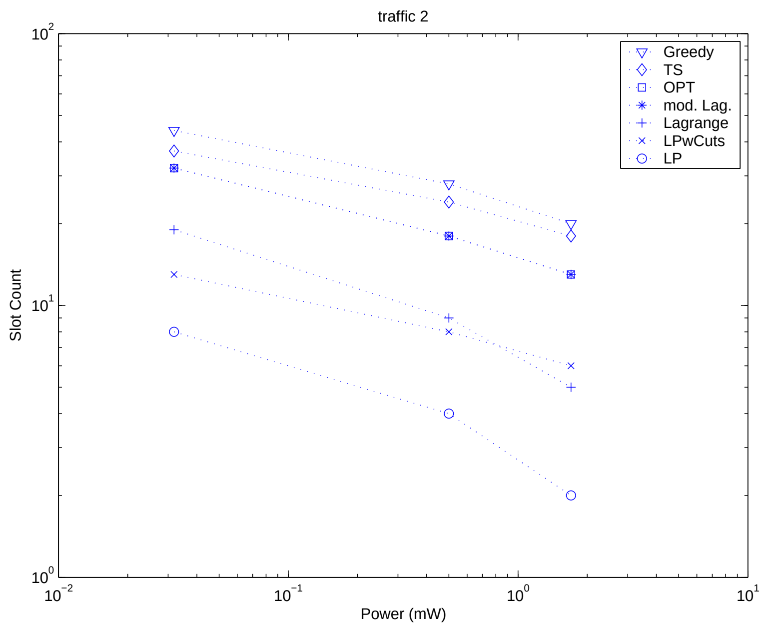

Figure 4.9: Outcomes of the methods for 6-node Topology 2, Traffic 2

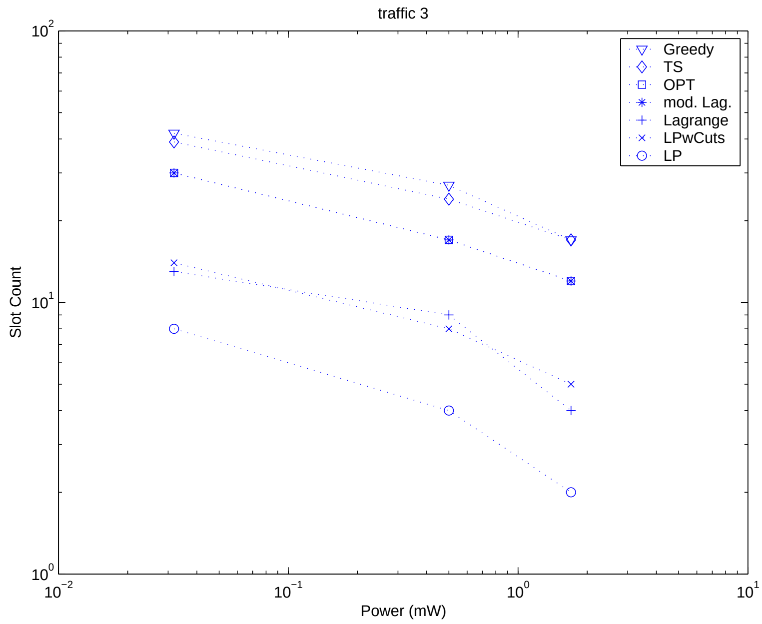

Figure 4.10: Outcomes of the methods for 6-node Topology 2, Traffic 3

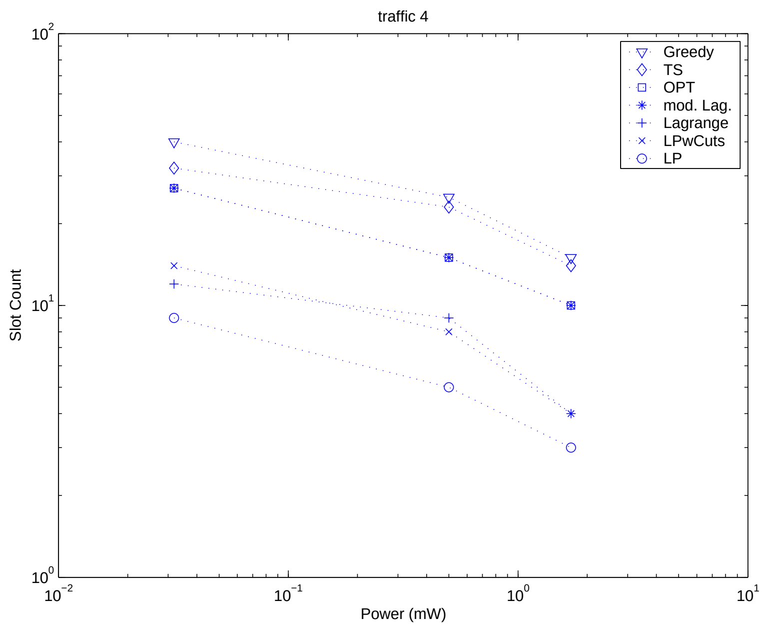

Figure 4.11: Outcomes of the methods for 6-node Topology 2, Traffic 4

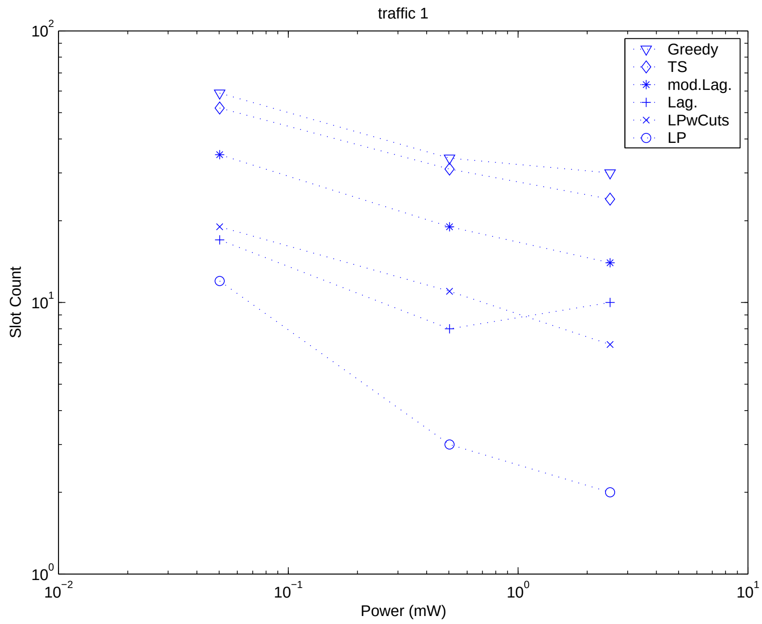

Figure 4.12: Outcomes of the methods for 8-node Topology 1, Traffic 1

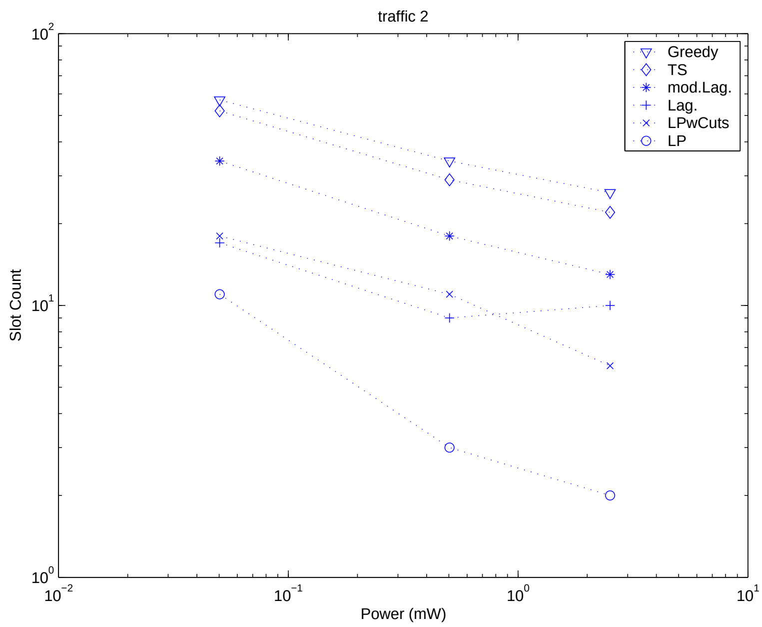

Figure 4.13: Outcomes of the methods for 8-node Topology 1, Traffic 2

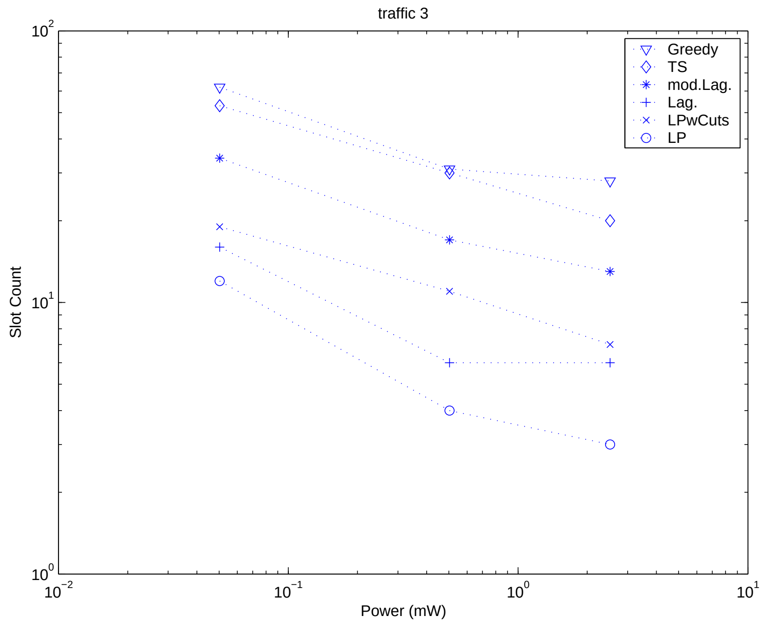

Figure 4.14: Outcomes of the methods for 8-node Topology 1, Traffic 3

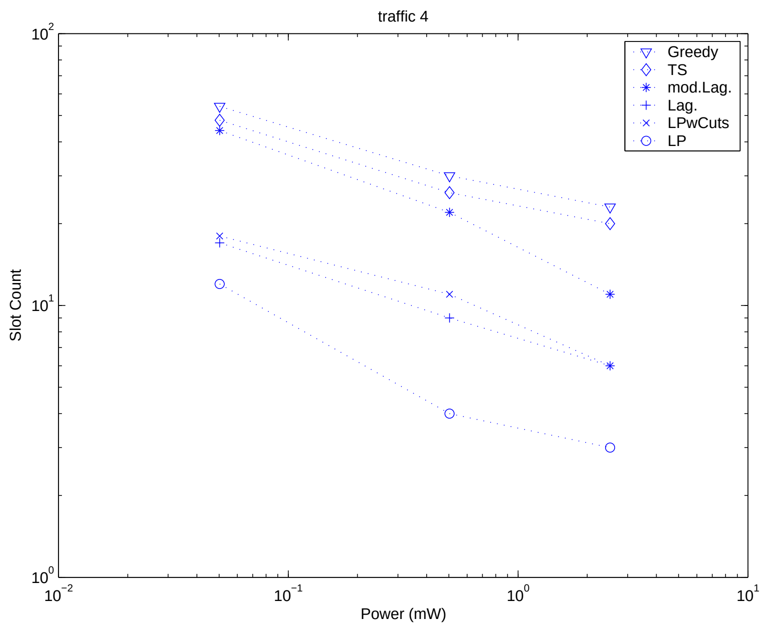

Figure 4.15: Outcomes of the methods for 8-node Topology 1, Traffic 4

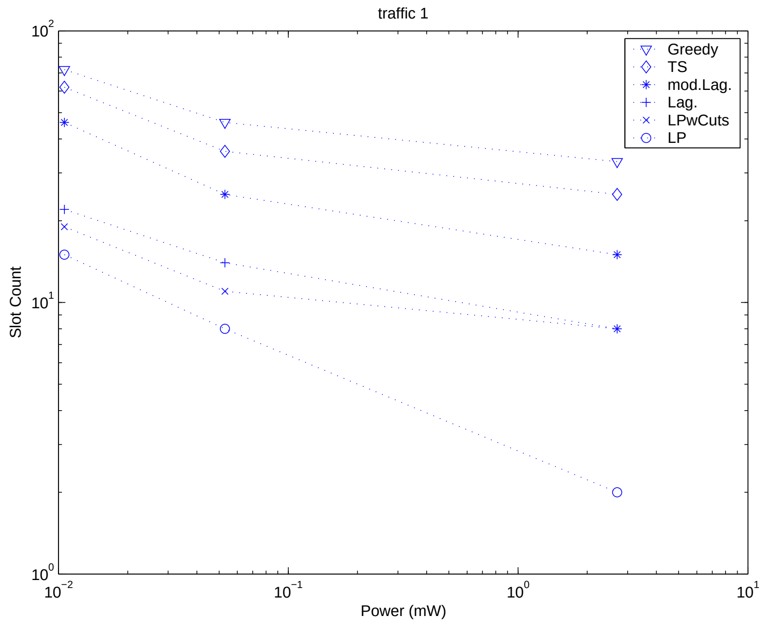

Figure 4.16: Outcomes of the methods for 8-node Topology 2, Traffic 1

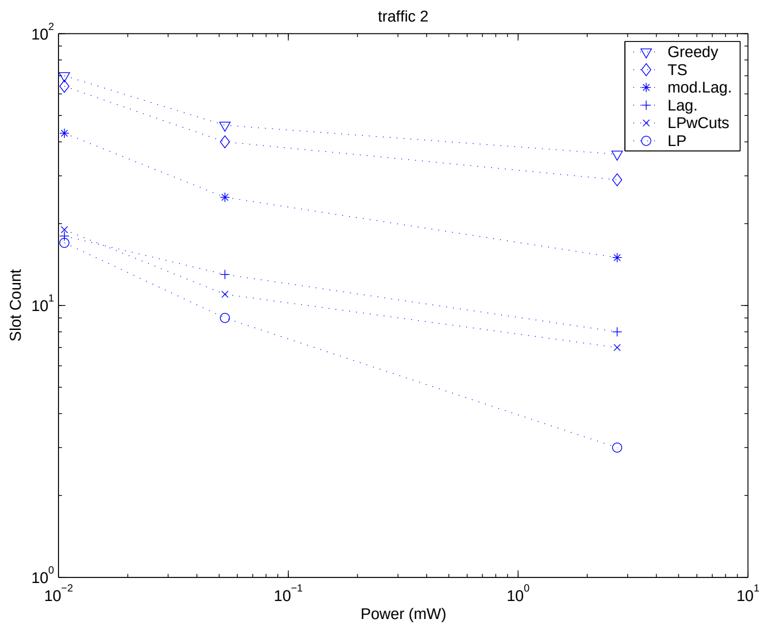

Figure 4.17: Outcomes of the methods for 8-node Topology 2, Traffic 2

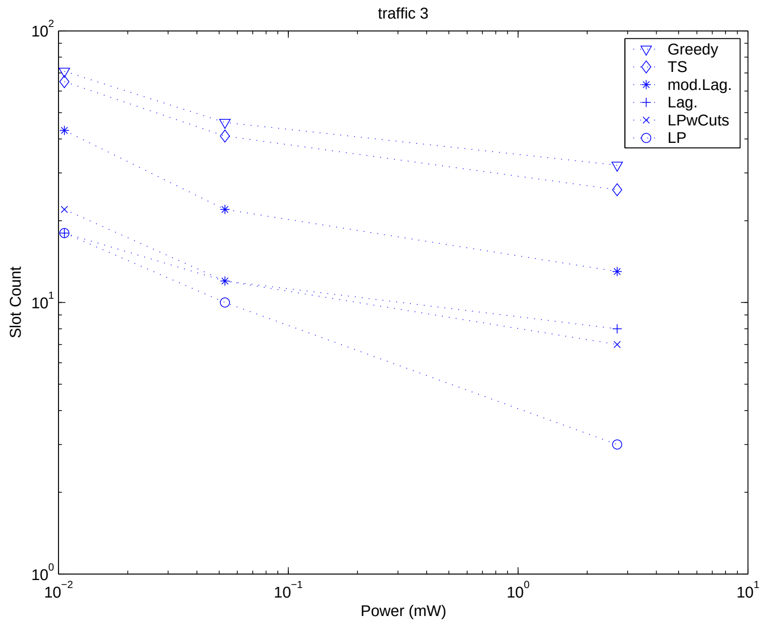

Figure 4.18: Outcomes of the methods for 8-node Topology 2, Traffic 3

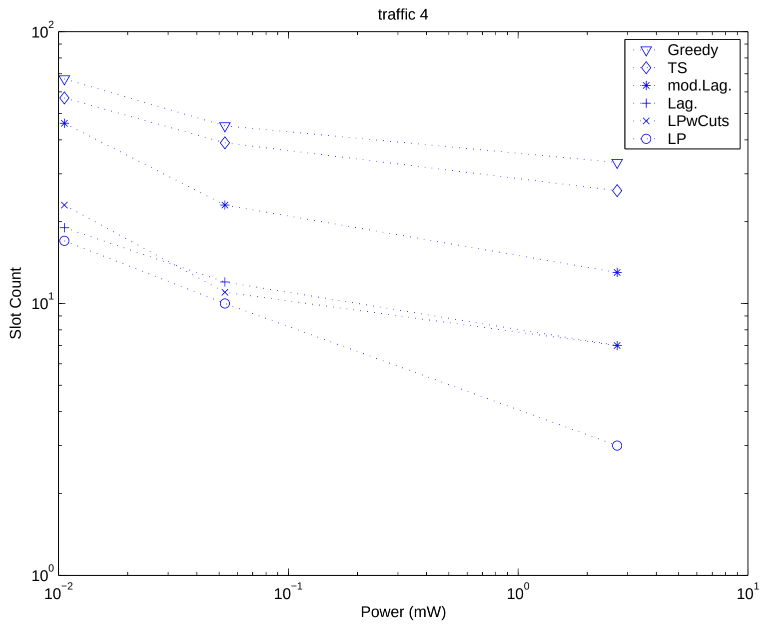

Figure 4.19: Outcomes of the methods for 8-node Topology 2, Traffic 4

		LP	***LPw/c***	***Lag***	***mod.Lag***	***OPT***	***Gre***	***TS***
P=0.011 mW	***Traf. 1***	8	17	24	40	40	47	45
	Traf. 2	9	18	30	40	40	45	44
	Traf. 3	9	18	24	35	35	43	43
	Traf. 4	13	18	23	38	38	43	41
P=0.106 mW	***Traf. 1***	4	11	9	20	20	27	24
	Traf. 2	5	12	9	20	20	27	23
	Traf. 3	4	12	7	17	17	24	23
	Traf. 4	5	11	6	15	15	22	20
P=2.650 mW	***Traf. 1***	2	7	7	14	14	15	14
	Traf. 2	2	7	6	12	12	16	16
	Traf. 3	2	7	6	11	11	15	14
	Traf. 4	3	7	6	10	10	14	14

Table 4.3: *Results (in TDMA slots) for 6-node Topology 1*

		LP	***LPw/c***	***Lag***	***mod.Lag***	***OPT***	***Gre***	***TS***
P=0.032 mW	***Traf. 1***	8	13	16	32	32	44	37
	Traf. 2	8	13	19	32	32	44	37
	Traf. 3	8	14	13	30	30	42	39
	Traf. 4	9	14	12	27	27	40	32
P=0.500 mW	***Traf. 1***	4	8	9	18	18	27	25
	Traf. 2	4	8	9	18	18	28	24
	Traf. 3	4	8	9	17	17	27	24
	Traf. 4	5	8	9	15	15	25	23
P=1.700 mW	***Traf. 1***	2	5	7	15	15	24	22
	Traf. 2	2	6	5	13	13	20	18
	Traf. 3	2	5	4	12	12	17	17
	Traf. 4	3	4	4	10	10	15	14

Table 4.4: *Results (in TDMA slots) for 6-node Topology 2*

		LP	***LPw/c***	***Lag***	***mod.Lag***	***Gre***	***TS***
P=0.050 mW	***Traf. 1***	12	19	17	35	59	52
	Traf. 2	11	18	17	34	57	52
	Traf. 3	12	19	16	34	62	53
	Traf. 4	12	18	17	44	54	48
P=0.503 mW	***Traf. 1***	3	11	8	19	34	31
	Traf. 2	3	11	9	18	34	29
	Traf. 3	4	11	6	17	31	30
	Traf. 4	4	11	9	22	30	26
P=2.514 mW	***Traf. 1***	2	7	10	14	30	24
	Traf. 2	2	6	10	13	26	22
	Traf. 3	3	7	6	13	28	20
	Traf. 4	3	6	6	11	23	20

Table 4.5: *Results (in TDMA slots) for 8-node Topology 1*

		LP	***LPw/c***	***Lag***	***mod.Lag***	***Gre***	***TS***
P=0.011 mW	***Traf. 1***	15	19	17	46	72	62
	Traf. 2	17	19	17	43	70	64
	Traf. 3	18	22	16	43	71	65
	Traf. 4	17	23	17	46	67	57
P=0.053 mW	***Traf. 1***	8	11	8	25	46	36
	Traf. 2	9	11	9	25	46	40
	Traf. 3	10	12	6	22	46	41
	Traf. 4	10	11	9	23	45	39
P=2.700 mW	***Traf. 1***	2	8	10	15	33	25
	Traf. 2	3	7	10	15	36	29
	Traf. 3	3	7	6	13	32	26
	Traf. 4	3	7	6	13	33	26

Table 4.6: *Results (in TDMA slots) for 8-node Topology 2*

In Figures 4.4 - 4.11 and Tables 4.3-4.4, which correspond to 6-node scenarios, all methods we applied in this thesis are shown. For these 6-node scenarios the optimum solution is available as stated in Section 3.4. In addition to the optimum, we observe the lower bound techniques, LP relaxation, LP relaxation with cuts, Lagrangian relaxation and modified Lagrangian relaxation, as well as the upper bound techniques, the greedy and the tabu search solutions.

In Figures 4.12 - 4.19 and Tables 4.5 - 4.6, corresponding to the 8-node scenarios we cannot find the optimum due to computational intractibility. All the bounding techniques are applied and the results are noted in the corresponding tables and shown in the figures.

In all the topologies, the computations become harder with increasing power, as this leads to an increase in the number of non-zero variables. The simulations for 6-node topologies did not have problems in terms of computation time. In the worst case, the optimum solution took around 16 hours. The bounding techniques for 6-node case were achieved at most in the order of minutes, with the exception of modified Lagrangian case which took just around the optimum solution time.

On the other hand, for the 8-node case optimum solution can only be achieved in very few of the situations, and take up time in the order of days. Most of the time, the optimum solution could not be obtained. The simulations took in the order of several hours for the lower bound case, whereas a few minutes for the upper bound. The LP relaxation is the fastest technique as expected. The modified Lagrangian relaxation performed worst in computation time, but this is certainly expected as only a very small subset of constraints are relaxed.

As a general property, all simulations point out that the throughput is maximized for the highest power in our simulations (which is a rather small power actually), and increasing the power furthermore do not increase throughput. This stems from the multirate nature of the problem, as higher power allows higher rates.

Figures 4.4 - 4.19 and Tables 4.3 - 4.6 also point out the effect of traffic. In all the cases, the total number of packets are the same but the traffic matrices differ. It is seen that the number of slots is the maximum for uniform traffic (traffic 1). Sparse traffic cases (traffic 2-3) come next and the single gateway case (traffic 4) has the minimum number of slots among the four traffic cases. That is caused by using links under their capacity,i.e. although the SINR allows it nodes do not use high-speed links to their full, as they have small amount of packets to send everywhere. These cases also differ in computation time. The uniform traffic case is the hardest to solve in terms of computation time, as it does not provide useful matrix structures like the sparse cases.

We observe that LP relaxation provides the worst lower bound in all topologies and traffic. When we integrate cuts to the LP relaxation and a significant improvement ranging from 30% to 100% occurs in the lower bound. Lagrangian relaxation performs similar to the LP relaxation with cuts, in some cases it provides a tighter lower bound, whereas in others the bound is looser. Moreover, based on the our observations, the cases that Lagrangian relaxation performs better do not share a common ground, and tend to be random in terms of traffic, topology, and power. The modified Lagrangian relaxation we suggested outperforms the other lower bound techniques in every case. That is also expected, as this technique starts from the bound obtained by LP or Lagrange relaxations (whichever is tighter), and continues until computation becomes intractable. Thus, for 6-node scenarios this technique provides a lower bound equal to the optimum, as the problem is not computationally intractable for these scenarios. As for the 8-node topologies, significant increase in the minimum number of slots with respect to LP and Lagrangian relaxations is obtained. This increase is more than 100% in most cases. So, this technique provides the tightest lower bound in our simulations.

Next, we discuss the upper bounding techniques: greedy and tabu search solutions. First, we observe the 6-node topologies where we can compare with the optimum solution. It is seen that the upper bound provided by the greedy solution is tighter compared to the lower bound provided by LP and Lagrangian relaxations. This is expected as the greedy solution provides a feasible solution by maximizing packets sent at each slot, and the number of nodes and routes possible are relatively small. Tabu search mechanism suggested in our thesis provides a tighter upper bound than the greedy solution. Typically 10% improvement over the greedy solution is achieved. Typically the upper bound obtained by this technique is $10 - 30\%$ higher than the optimum, and for some high power simulations, even the optimum could be reached.

The 8-node scenarios, where the optimum solution is not available, are discussed next. The tabu search technique achieves typically $10-20\%$ improvement over the greedy solution. This slight increase compared to 6-node scenario stems from the fact that 8-node provides more possibilities in terms of link schedules and routes, and the range of feasible solutions is wider. So, the probability of getting further away from the optimum by a greedy approach increases. In 8-node case, we cannot find the optimum, but despite this fact, upper and lower bounds are within $30 - 50$ proximity of each other.

Chapter 5

CONCLUSIONS

In this thesis, we investigated the joint link scheduling and routing strategies for STDMA based multi-rate single-channel Wireless Mesh Networks. We proposed an Integer Linear Programming (ILP) optimization framework to determine the shortest deadline within which a specified traffic can be jointly routed and scheduled among WMN nodes. The solution provides a link activation schedule for delivering a given amount of traffic to the intended destinations.

One main finding is that the throughput is maximized for the highest power in our simulations. It is also noted that further increasing the power do not improve throughput. This increase in throughput is caused by both the relatively small size of the networks we dealt with, and the multirate nature of the problem, as higher power opens up higher rates. Another finding is that the number of slots is the maximum for uniform traffic case. Sparse traffic cases and the single gateway case take less delay to be delivered, as these cases better make use of the multi-rate structure.

The integer constraint introduced in the ILP framework causes computational complexity to grow exponentially, as the number of nodes or the length of the time frame increase. The computational intractibility allows us to obtain the optimum solution for at most 6-node topologies. Thus, we proposed heuristic techniques to obtain upper and lower bounds on the optimum solution.

Firstly, we introduced lower bounding techniques which involve LP and Lagrangian relaxations. We significantly improved results of LP relaxation using cutting planes, ranging from 30% to 100% increase in the lower bound. This improvement brought the lower bound obtained by LP relaxation, around the same level with the one obtained by the Lagrangian relaxation. The final approach we implemented in lower bounding involves slight modifications on the Lagrangian relaxation, considering the nature of our problem. The method proved successful in terms of the improvement of the lower bound, more than 100% increase is observed in lowest power cases, typically over 50% increase is noted for other powers. However, this method requires relaxing the computationally intractable formulation until the complexity decreases to the maximum allowed by the state of the art solvers. So, computation time is the maximum tolerable value, which happened to be a few days in our case.

Finally, the upper bound techniques are introduced: greedy and tabu search solutions. The greedy approach provided feasible solutions to be used as an initiation to the tabu search. Tabu search resulted in tighter upper bounds on the optimum solution compared to the initial greedy solution as expected, and feasible solutions with $10 - 20\%$ smaller number of slots are achieved. Also worth noting is that, the computation complexity of our tabu search approach is not high, as we decreased the search space to a small acceptable size using information obtained from the greedy solution.

Our bounding techniques can serve as a prelude in more complex problems involving multi-radio and multi-channels. If computational complexity is overcome, addition of dynamic power allocation would also be useful in terms of power efficiency and network throughput.

Bibliography

[1] P. Gupta, and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, pp. 388 - 404, 2000.

[2] M.L. Fisher, "The Lagrangian relaxation method for solving integer programming problems," *Manage. Sci.*, 1981.

[3] A.M. Geoffrion, "Lagrangian relaxation and its uses in integer programming," *Mathematical Programming Study*, pp. 82 - 114, 1974.

[4] F. Glover, "Tabu search-Part I," *ORSA J. Comput.*, vol. 1, pp. 190 - 206, 1989.

[5] F. Glover, "Tabu search-Part II," *ORSA J. Comput.*, vol. 2, pp. 4 - 32, 1990.

[6] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445 - 487, 2005.

[7] G. Holland, N. Vaidya, P. Bahl, "A rate - adaptive MAC protocol for multi - Hop wireless networks," *Proceedings of the 7th annual international conference on Mobile computing and networking*, p.236 - 251, July 2001.

[8] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Transactions on Communications*, vol. 43, pp. 2223 - 2230, July 1995.

[9] T. Ue, S. Sampei, N. Morinaga, and K. Hamaguchi, "Symbol rate and modulation level - controlled adaptive modulation/TDMA/TDD system for high

bit - rate wireless data transmission," *IEEE Transactions on Vehicular Technology*, vol. 47, pp. 1134 - 1147, November 1998.

[10] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE Transactions on Communications*, vol. 47, pp. 884 - 895, June 1999.

[11] K. Balachandran, S. R. Kadaba, and S. Nanda, "Channel quality estimation and rate adaption for cellular mobile radio," *IEEE Journal on Selected Areas in Communications*, vol.17, pp. 1244 - 1256, July 1999.

[12] A. Capone and G. Carello, "Scheduling optimization in wireless mesh networks with power control and rate adaptation," in *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*, vol. 1, pp. 138 - 147, Sept. 2006.

[13] S. Kumar, V. S. Raghavan, and J. Deng, "Medium Access Control protocols for ad hoc wireless networks: A survey," *Ad Hoc Networks (Elsevier)*, vol. 4, pp. 326 - 358, May 2006.

[14] S. Ramanathan and E.L. Lloyd, "Scheduling algorithms for multihop radio networks," *IEEE/ACM Trans. on Networking*, vol. 1, no. 2, pp. 166 - 177, April 1993.

[15] A. Ephremides, T. Truong, "Scheduling broadcasts in multihop radio networks," *IEEE Trans. on Commun.*, vol. 38, no. 4, pp. 456 - 460, April 1990.

[16] M. Alicherry, R. Bhatia, and L. E. Li, "Joint channel assignment and routing for throughput optimization in multi - radio wireless mesh networks," in *11th Annual Internanassertional Conference on Mobile Computing and Networking (MobiCom)*, pp. 58 - 72, 2005.

[17] A. Behzad, I. Rubin, "Multiple access protocol for power - controlled wireless access nets," *IEEE Trans. on Mobile Computing*, vol. 3, no. 4, pp. 307 - 316, October - December 2004.

[18] Y. Li and A. Ephremides, "A joint scheduling, power control, and routing algorithm for ad hoc wireless networks," *Ad Hoc Networks*, vol. 5, no. 7, pp. 959 - 973, 2007.

[19] T. ElBatt, A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Trans. on Wireless Communications*, vol. 3, no. 1, pp. 74 - 85, January 2004.

[20] G. Kulkarni, V. Raghunathan, and M. Srivastava, "Joint end - to - end scheduling, power control and rate control in multi - hop wireless networks," *Proc. of IEEE GLOBECOM*, 2004.

[21] U. 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 *Twenty - third AnnualJoint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 2, pp. 1446 - 1456 vol.2, March 2004.

[22] T. Clausen, P. Jacquet, "Optimized Link State Routing Protocol (OLSR) (OLSR)," RFC 3626, October 2003.

[23] C. Perkins, E. Belding - Royer, and S. Das, "Ad hoc On - Demand Distance Vector (AODV) Routing," RFC 3561, July 2003.

[24] D. B. Johnson, D.A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, vol. 353, 1996.

[25] A. Lee, P.A.S. Ward, "A Study of Routing Algorithms in Wireless Mesh Networks," *Australian Telecommunication Networks and Applications Conference*, December, 2004

[26] A. Raniwala and T. Chiueh, "Architecture and algorithms for an IEEE 802.11 - based multi - channel wireless mesh network," in *24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 3, pp. 2223 - 2234 vol. 3, March 2005.

[27] D.S.J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "High - throughput path metric for multi - hop wireless routing," in *Proceedings of ACM MOBICOM'03*, Sept. 2003.

[28] R. Draves, J. Padhye, and B. Zill, "The architecture of the link quality source routing protocol," Technical Report MSR - TR - 2004 - 57, Microsoft Research, 2004.

[29] T. Liu, W. Liao, "On Routing in Multichannel Wireless Mesh Networks: Challenges and Solutions," IEEE Network, January/February 2008.

[30] E. Jones, "Multipath Load Balancing in Multi - hop Wireless Networks," Technical Report, University of Waterloo, 2004.

[31] V.T. Raisinghani, S. Iyer, "Crosslayer design optimizations in wireless protocol stacks," *Computer communications*, vol.27, pp. 720 - 724, October 2003.

[32] L. van Hoesel, T. Nieberg, J. Wu, and P. J. M. Havinga, "Prolonging the lifetime of wireless sensor networks by cross - layer interaction," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 78 - 86, Dec. 2004.

[33] T. Melodia, M. Vuran, and D. Pompili, "The state of the art in cross - layer design for wireless sensor networks," Proc. of EuroNGI Workshops on Wireless and Mobility. Como, Italy, July 2005.

[34] X. Yin, "Improving tcp performance over mobile ad hoc networks by exploiting cross - layer information awareness," Proc. of the 10th annual international conference on Mobile computing and networking, pp. 231 - 244, Philadelphia, USA, October 2004.

[35] G. Ferrari, S. Malvassori, M. Bragalini, and O. Tonguz, "Physical layerconstrained routing in ad - hoc wireless networks: A modified aodv protocol with power control," Proc. of IWWAN 2005, London, UK, May 2005

[36] A. Capone, I. Filippini, and F. Martignon, "Joint routing and scheduling optimization in wireless mesh networks with directional antennas," in *IEEE International Conference on Communications*, pp. 2951 - 2957, May 2008.

[37] L. Badia, A. Erta, L. Lenzini, and M. Zorzi, "A general interference - aware framework for joint routing and link scheduling in wireless mesh networks," *IEEE Network*, vol. 22, pp. 32 - 38, Jan. - Feb. 2008.

[38] L. Badia, A. Botta, and L. Lenzini, "A genetic approach to joint routing and link scheduling for wireless mesh networks," *Elsevier Ad Hoc Networks Journal*, vol. Special issue on Bio - Inspired Computing, pp. 11, April 2008.

[39] B. Awerbuch, D. Holmer, and H. Rubens, "High throughput route selection in multi - rate ad hoc wireless networks," *Lecture Notes in Computer Science*, vol. 2928, pp. 253 - 270, 2004.

[40] J. Zhang, H. Wu, Q. Zhang, and B. Li, "Joint routing and scheduling in multi - radio multi - channel multi - hop wireless networks," in *2nd International Conference on Broadband Networks*, pp. 631 - 640 Vol. 1, Oct. 2005.

[41] X. Chen, Z. Zhang, and H. Luo, "Joint optimization of power control, channel assignment and scheduling in wireless mesh network," in *Third International Conference on Communications and Networking in China*, pp. 606 - 611, Aug. 2008.

[42] G. Carello, I. Filippini, S. Gualandi, F. Malucelli, "Scheduling and routing in wireless multi - hop networks by column generation," *INOC*, April 2007.

[43] C. Gomes, C. Molle, P. Reyes, "Optimal design of wireless mesh networks," *JDIR*, January 2008.

[44] R. Cruz and A. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," in *Twenty - Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 1, pp. 702 - 711 vol.1, March - 3 April 2003.

[45] V. Srivastava, M. Motani, "Cross-layer design: A survey and the road ahead," *IEEE Commun. Mag.*, vol. 43, pp. 112 - 119, Dec. 2005.

[46] V.Kawadia , P.R.Kumar, "A cautionary perspective on cross-layer design," *IEEE Wireless Communication Magazine* , vol. 12, pp. 3 - 11, 2005.

[47] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi - hop wireless network performance," *Wirel. Netw.*, vol. 11, no. 4, pp. 471 - 487, 2005.

[48] J.E. Kelley,"The cutting - plane method for solving convex programs ," *J. Soc. Ind. Appl. Math.*, 1961

[49] J.L.Goffin, "On convergence rates of subgradient optimization methods," *Math. Program.*, 1977.

[50] E. Aarts, J.K. Lenstra, "Local search in combinatorial optimization," Wiley, Chichester, U.K., 1997.