謝詞

本篇論文完成，代表研究所求學時代的結束，也是人生另一個階段的開始。首先要感謝我的父母：高森林先生與高周美女女士。感謝父母給我良好的環境與全力的支持，使我能順利走完研究的漫長旅程。

在大學及研究所的求學過程中，感謝許多老師鼓勵與支持，使這篇論文能順利完成。感謝恩師林永松博士，從大學的課程起，老師總是不厭其煩的給我指導與方向。在研究所論文寫作期間，老師更帶領我深入各個領域，並提供良好環境，使本研究能更深入、完善。這五年來在老師身邊，我不僅僅學到研究精神，更學到做人做事的道理。而口試期間，承蒙台灣科技大學陳郁堂博士、本系莊裕澤博士、世新大學顏宏旭博士對學生論文的指正及提出建議，使本論文能更為嚴謹，成果更為豐碩。

感謝研究所同學在論文寫作期間的協助：佳宏、晉華、至鈞、麒安、克行，與你們共同切磋，使我的論文更加完善。在台大資管待了六年，對這裡有著複雜的情感。台大資管，我永遠不會忘記在這裡的點點滴滴。

最後要感謝我的女友：宛儒，謝謝妳這些日子以來對我的包容。人生的旅途有妳相伴，一切都美好。

高茂原　謹識

于台大資訊管理研究所

民國九十二年七月

論文摘要

近年來，光通訊網路技術中的多波長分工網路逐漸成為被廣泛期許的傳輸標準之一。由目前發展顯示，多波長分工網路未來在橫跨大區域之網際網路骨幹網路中將扮演實體傳輸層要角。基於多波長分工網路技術，目前許多學者及業界人士正致力於將現有的網路流量以訊量匯整的方式負載於其上，讓多波長分工網路技術得以充分利用其所帶來之大量頻寬。

目前相關研究大多著重在現有 SONET 的環狀網路架構下，如何最大化由於多波長分工網路上支援多網路流量共同使用同一波長光路徑而提高的網路流量；以及如何降低其網路架構成本。為能有效的利用多波長分工網路所帶來的大量頻寬，本篇論文將研究在一般網狀網路架構下之多波長分工網路初始規劃建置問題。在既有的動態網路流量需求及其他限制條件下，希望能建立流量總阻塞機率最小化的網路。

我們將整個問題仔細地分析並轉換成一個最佳化數學模型，這個數學問題在本質上是一個非線性混和整數規劃問題，問題的本身具有高度的複雜性和困難度。我們採用以拉格蘭日鬆弛法為基礎的方法來處理此一複雜問題。

關鍵詞：多波長分工網路、網路規劃、路由、波長分配、阻塞機率、最佳化、拉格蘭日鬆弛法、數學規劃。

THESIS ABSTRACT

GRADUATE INSTITUTE OF INFORMATION MANAGEMENT

NATIONAL TAIWAN UNIVERSITY

NAME：MAO-YUAN KAO **MONTH/YEAR：JULY, 2003**

ADVISER：YEONG-SUNG LIN

MINIMIZATION OF OVERALL BLOCKING PROBABILITY FOR TRAFFIC GROOMING IN OPTICAL WDM MESH NETWORKS

Recently, wavelength division multiplexing (WDM) has been considered as a promising transmission technology in optical communication networks. Researchers and optical networking industries are now trying to find a way to support multiple low-speed traffic streams onto a single high-speed wavelength in order to minimize the network-wide cost in terms of line terminating equipment and electronic switching.

In the proposed research, the problem of traffic grooming in optical WDM networks is studied. The problem is formulated as a combinatorial optimization problem, of which the objective function is to minimize the total traffic blocking probability upon the underlying WDM mesh network subject to physical link capacity constraints, wavelength continuity constraints, and wavelength add/drop port constraints The decision variables in the formulations include traffic requirement routing design, WDM layer lightpath routing assignment and WDM layer wavelength assignment.

The basic approach to the algorithm development for this thesis is Lagrangean relaxation in conjunction with a number of optimization techniques. We evaluate the performance for different network topologies.

Keywords: WDM, Traffic Grooming, Network Planning, Routing, Wavelength Assignment, Blocking Probability, Optimization, Lagrangean Relaxation, Mathematical Programming.

Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

Over the last few years, the continual growth of bandwidth demand in IP backbones pushes ahead the introduction of the wavelength division multiplexing (WDM) technology in today's optical communication networks. WDM technology is considered to be a promising transport mechanism for improving fiber bandwidth utilization. Recent optical technology shows that a single fiber can contain up to 160 channels (wavelengths), each with capacity of OC-192 (10 Gbps) [29]. Furthermore, channels operating at OC-768 (40 Gbps) will soon be commercially available in the near future [21].

WDM is an approach that can exploit the potential ample bandwidth of fibers by multiplexing multiple WDM channels, which is used by only one traffic requirement, on the same fiber, as depicted in Figure 1-1. Current research and development on optical technology indicates that WDM technology will be deployed mainly in a backbone network for large regions [21]. WDM can also enhance an optical network's capacity without expensive re-cabling and can tremendously reduce the cost of network upgrades.

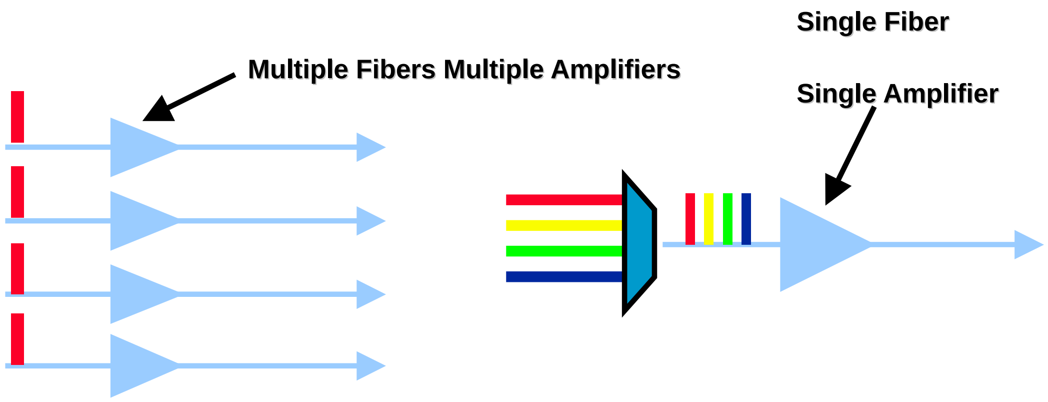


Figure 1-1 Illustration of Wavelength Multiplexing

In the beginning, WDM technology is deployed for a fixed point-to-point communication. This deployment was driven by the increasing demands on bandwidth. WDM is a more cost-effective alternative compared to laying more fibers. Afterwards, significant advances in optical component technologies (e.g., amplifiers, fibers, filter and laser sources) have brought about more advanced WDM sub-systems providing wavelength routing functions, such as static/reconfigurable optical add/drop multiplexers (OADMs), waveguide grating routers (WGRs), and reconfigurable optical cross-connects switches (OXCs). The routing function of Static OADMs and WGRs are nonreconfigurable, which means the routes of different wavelengths are fixed in these components. For reconfigurable OADMs and OXCs, their routing functions are reconfigurable and can be controlled electronically. In a WDM network employing static OADMs and WGRs in network nodes, a lightpath can be established using a wavelength, which means that the WDM network provides wavelength routing capability. This allows for overlaying "virtual" higher-layer protocol topologies on the top of the physical layer topology (WDM network), where the lightpaths represent virtual links. But static OADMs and WGRs are nonreconfigurable so the virtual topologies are not changeable. With reconfigurable OADMs and OXCs, the routing capacity of the WDM network is more flexible, because the components provide additional control in setting up connections by electronic signals.

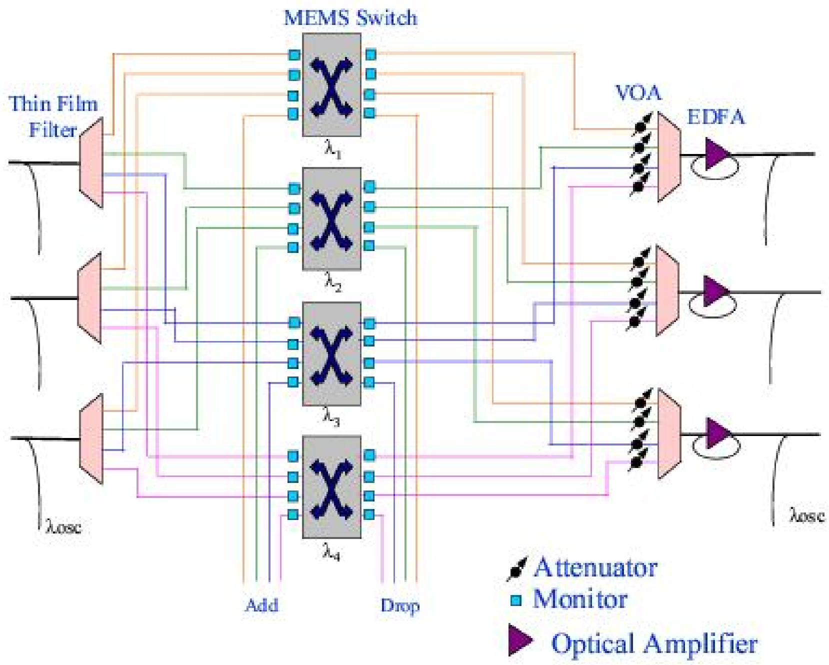


Figure 1-2Example Architecture of OXC

In Figure 1-2, we illustrate an architecture example of OXC. It is now a common architecture of OXCs in the market. In this architecture, a specific wavelength MEMS switch, which performs the lightpath switching function, can add and drop limited number of lightpaths simultaneously. For example, in the λ_1 MEMS switch we can add and drop only one of λ_1 lightpath simultaneously. In our formulation, we consider this as a constraint. But if the architecture changed and add /drop functions of MEMS switches are not a constraint anymore, we can just adjust add/drop related parameters to adapt the new condition. We also assume there is no wavelength conversion for each established lightpath, because the O-E-O wavelength conversion slows down the network performance. However, we allow traffic demands being sent through different lightpaths on a single path and use the hop-count constraint to limit the downgrade of performance from the O-E-O conversions. We don't take the all-optical wavelength conversion into consideration in our model since the cost of the conversion is still too high so far.

Nowadays, most of the traffic connections in a communication system have smaller bandwidth requirement comparing to the large granularity of the wavelength capacity. With such traffic patterns and characteristics, assigning one wavelength for each connection request could lead to poor bandwidth utilization, and hence link capacity could be greatly under-utilized. Therefore, if there is a way to aggregate the traffic of multiple connections into a single wavelength channel, the wavelength capacity would be highly utilized. Traffic grooming is such a technique that divides a lightpath into multiple time slots and traffic demands can be multiplexed on the same lightpath. It can considerably improve the utilization of bandwidth of wavelength channels. The resulting multi-wavelength time-division multiplexed networks are referred to as WDM-TDM networks or WDM grooming networks [26].

Recent progress in optical switching technology that results in faster switches up to sub-nanosecond switching times, and the introduction of fiber delay lines as time-slot interchangers make all-optical TDM networks potentially feasible in the future. An architectural prototype of a TDM routed node, called *time wavelength space router*, is presented in [10][9]. It is capable of routing data in the dimension of time, wavelength, and space in a TDM wavelength-routed network.

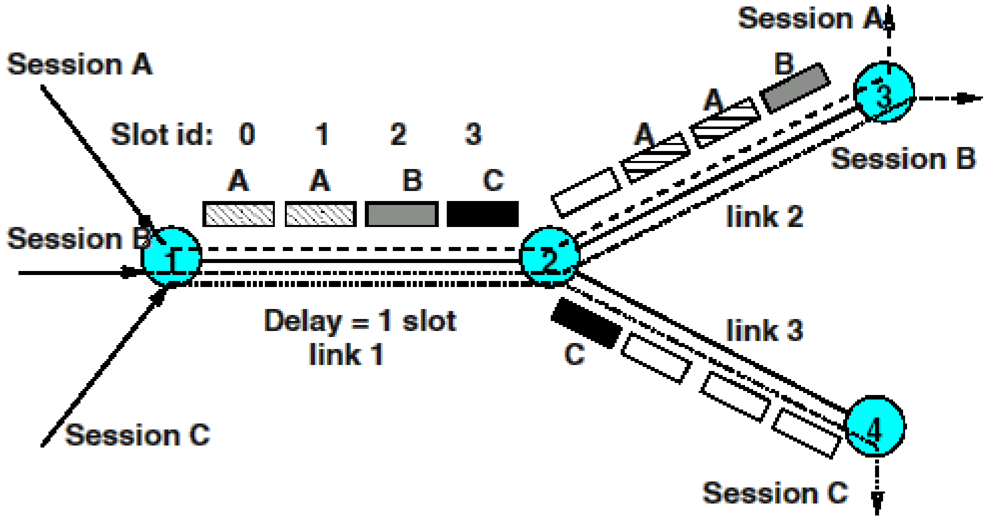


Figure 1-3Time-slot Allocation for a TDM Wavelength-routed Network

An example showing how the TDM works on WDM networks is illustrated in Figure 1-3 (from [29]). In the figure, each wavelength is divided into frames composed of four time-slots. Session A's traffic is routed from node 1 to 3 through node 2. It is assigned slots 0 and 1 on link 1; and slots 1 and 2 on link2. The link propagation and node processing delay result in a shifted time-slot allocation, with propagation delay dominating in WANs. The combined link delay is considered to be an integral number of time-slots, with optical synchronizers used to align data to slots. In the example a delay of one slot is assumed for link 1, shifting the allocated slots on link 2 by one slot. Similarly, the time-slot allocations for session B and C can be seen in the figure.

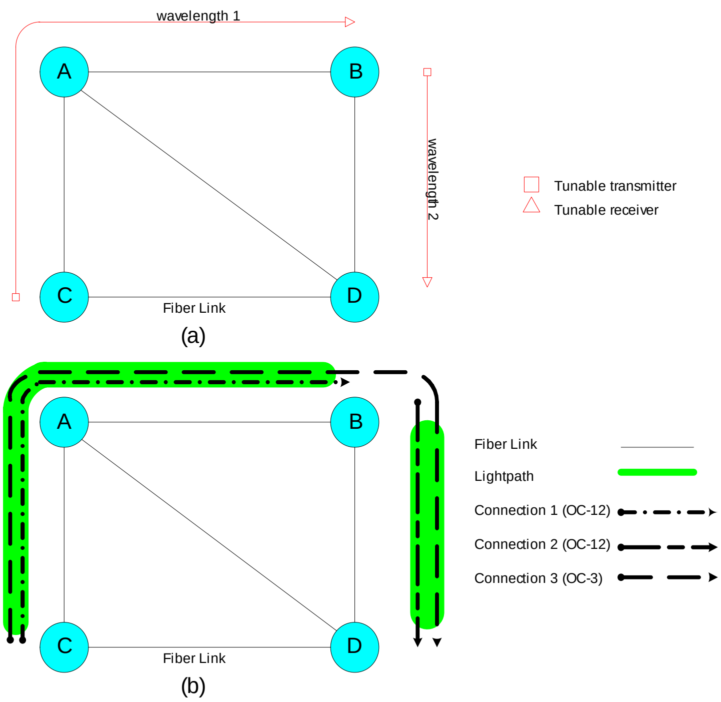


Figure 1-4 Example of Traffic Grooming

Figure 1-4 is an illustrative example of traffic grooming in a WDM mesh network. Figure 1-4(b) shows a simple network with four nodes. We assume that each fiber has two wavelength channels and the capacity per wavelength channel is OC-48 (2.5 Gbps). Each node is assumed to have only one tunable transmitter and receiver which can be tuned to any wavelength. The traffic demand matrix is shown as the following table:

	Source	**Destination**	**Requirement**
Connection 1	C	B	OC-12
Connection 2	B	D	OC-12
Connection 3	C	D	OC-3

Table 1-1 Traffic Requirement of the Grooming Example

In order to satisfy the traffic requirement, two lightpaths are established under the original WDM network without traffic grooming support; however, due to the resource limitation, a direct lightpath between node C and node D cannot be created. And thus, the connection request 3 cannot be fulfilled unless more transceivers are added at each node and available wavelength channels on fibers are increased. On the contrary, with traffic grooming support, the connection 3 can be carried by sharing the free capacity of these two existing lightpaths. The grooming result is illustrated in Figure 1-4(b).

1.2 Motivation

In most previous studies on optical networks, traffic demands are usually assumed to be wavelength demands. In practice, optical networks are typically required to carry a large number of lower rate traffic demands. Therefore, traffic grooming is an important mechanism by which low-rate traffic requests could be assigned to wavelengths in order to improve the wavelength channel utilization and to minimize the amount of electronic multiplexing equipment.

Recently, a significant amount of research has been carried out in this area. However most of the research on traffic grooming has been focused on SONET rings [27][28][30][31]. However, due to the growth in Internet traffic, an increasing number of networks are being arranged in a general mesh topology. Therefore, there is a need to extend the grooming work to general mesh networks. For improving utilization of wavelength channels over WDM networks in the near future, we will propose a mathematical formulation and an optimization-based algorithm considering the traffic grooming, routing, and wavelength assignment problem over WDM mesh networks.

1.3 Literature Survey

In this section, we firstly summarize some similar concepts about the grooming problem from relative works in ATM networks and point out the differences. Afterward, survey about traffic grooming problems on different WDM network topologies is discussed. At last, Lagrangean relaxation method is introduced.

1.3.1 Similar Concepts in ATM Networks

The virtual path assignment and virtual circuit routing problem in ATM networks has been a popular issue in the research area [1][16][17]. The use of virtual paths in ATM networks reduces the call set-up delays, simplifies the hardware in the transit nodes and provides simple virtual circuit admission control. A virtual path (VP) is a logical connection for a node pair by means of a label in the header of an ATM cell named Virtual Path Identifier (VPI). Each VP is considered as a logical link for a certain service, and therefore VP subnetworks for different services can be built within an ATM network. Each VP is assigned a number of physical links and an effective capacity (in terms of the maximum number of virtual circuits allowed) to assure quality of service (QoS) requirements. Several VPs may be multiplexed on the same physical link.

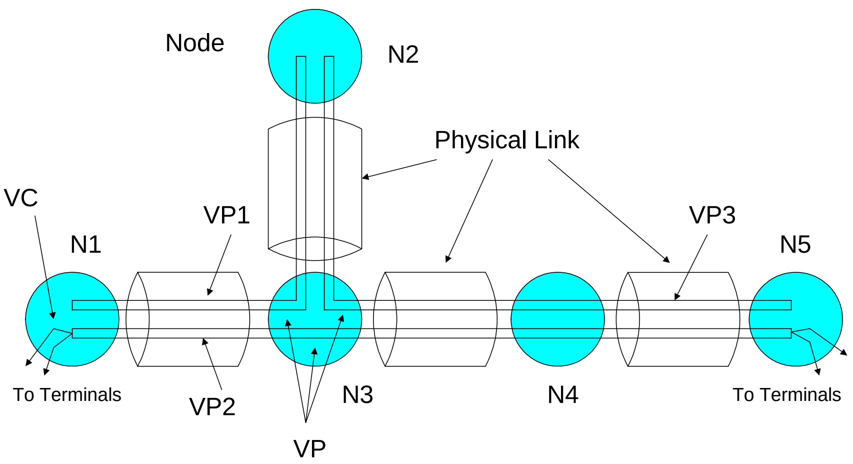


Figure 1-5 Schematic Illustration of Virtual Paths

Figure 1-5 (from [23]) shows the concept of virtual paths. There are three virtual paths VP_1, VP_2 and VP_3, shown in the figure. VP_2 is a virtual path between end nodes N_1 and N_5. Nodes N_3 and N_4 are transit nodes of virtual path VP_2.

The concept of virtual path assignment and virtual circuit routing problem is similar to the problem we want to solve here. We separate both problems into two layers and discuss the differences between them:

1. **Physical layer:** In WDM networks, the physical layer means the optical fiber connections between each OXCs. In ATM networks, the physical layer is the set of optical trunks on ATM networks.
2. **Logical layer:** In WDM networks, the logical layer is constructed by lightpaths. Each lightpath is a logical link passing through multiple underlying physical links. Each lightpath does not change its wavelength during transmission. In other words, lightpaths obey wavelength continuity constraint. When establishing lightpaths on each OXC, we have to take the OXC property into consideration, since there are limitations on the available transmitters and receivers at each OXC. In ATM networks, the logical layer is established by VPs. Each virtual path is considered as a logical link. Several VPs may be multiplexed on the same physical link. This is similar to the multiple wavelength channels on a single fiber. The bandwidth assigned to each virtual paths could be different. However, the capacities of lightpaths are all the same because each lightpath uses one wavelength. Later in our model, we will extend this constraint to allow each logical link contains more than one wavelength in order to make the model more flexible.

In WDM networks, multiple traffic demands can be multiplexed on a lightpath in order to improve the wavelength utilization. Each traffic demand is fulfilled by a logical path, which may pass through several lightpaths. Each traffic demand can only be sent as a whole by one selected path. In ATM networks, on the contrary, on the arrival of a virtual circuit, paths satisfying the QoS constraints are selected. Moreover, the traffic flow of a virtual circuit can be sent through more than one candidate paths.

In [16], the virtual path assignment problem and the virtual circuit routing assignment problem are jointly considered for the first time. A general problem formulation is presented in [16] and two special cases of the general problem are considered. One is to assume that the network is with abundant capacity and hence each virtual circuit use exactly one virtual path so that the call set-up delay is kept minimum. This implementation is favorable when the call set-up time is considered as a critical performance measure and the network has abundant capacity so that the achievable total call blocking rate is within an acceptable range. Another special case of the general problem is assumed that each physical link is considered a virtual path. This scheme is recommended for networks with very limited capacities and capacity sharing is highly desirable. Algorithms for solving the two special cases are proposed and implemented.

In [17], a solution procedure to the joint problem of assigning virtual paths and determining virtual circuit routing assignments for ATM networks is proposed, which is not solved in [16]. The solution consists of two phases. In the first phase, the capacity of each virtual path is fixed and the routing assignments are adjusted to reduce the overall call blocking rate. In the second phase, the virtual path capacity assignments are adjusted to reduce the overall call blocking rate, using the fixed virtual circuit routing assignments derived in the first phase.

In [1], the problem is further extended to include the Quality of Service constraints. The author focused on how to minimize the total call blocking rate on ATM networks with traffic demands of different QoS types, while at the same time, satisfying different call set-up time constraints for various traffic demands. There are many ways to manage how virtual circuits are created on virtual paths. Two methods are proposed in [1]. The author first introduced the separate traffic management, which fulfills traffic requirements with different service types by different VP. In this way, the management of VPs can be simplified. However, it indirectly increases the expected call blocking rate. The other way the author offered is to allow different traffic demands to be fulfilled on same paths.

1.3.2 Introduction to Traffic Grooming on WDM Networks

In order to utilize bandwidth more effectively, new models of optical networks are continuously proposed and improved by the academic communities and industries. One is to allow several independent traffic streams to share the bandwidth of a lightpath. If the multiplexing and demultiplexing of lower-rate traffic components is performed at the boundaries of the network only, and the aggregate traffic transparently traverses the network, this problem is equivalent to RWA problem. However, it is generally impossible to set up lightpaths between every origin and destination of traffic demands due to wavelength continuity constraint and number limitations of optical transceivers. Therefore, it is natural to consider optical networks at which traffic on terminating lightpaths is electronically switched (groomed) onto new lightpaths toward the destination node. Introducing some amount of electronic switching within the network has two advantages: it can considerably increase the degree of virtual connectivity, while at the same time it may drastically reduce the wavelength requirements within the optical network for a given traffic demand.

Most studies about traffic grooming on WDM networks are trying to balance the trade-off between the degree of virtual connectivity and the increase in network cost due to the introduction of expensive active components, for example optical transceivers or electronic switches. And the objectives can be summarized to minimize the following:

- Total number of lightpaths
- Total number of optical transceivers
- Total amount of electronic switching
- Total number of wavelengths used
- Maximum number of lightpaths at a node

In our model, we will solve the traffic grooming problem from another aspect, attempting to minimize the total blocking probability.

1.3.3 Traffic Grooming in Ring Networks

Most of current optical network infrastructures are built around Synchronous Optical Network (SONET) rings. Given the widespread use of SONET networks, it is not surprising that recent work has focused on ring topologies. In general, a SONET ring is established with fibers to connect SONET add/drop multiplexers (ADMs). SONET ADM can aggregate low-rate SONET signals onto a single higher-rate SONET stream.

Yoon [30] considered a traffic grooming and lightpath routing problem in WDM ring networks. The author formulates the problem as a mixed integer programming problem, trying to minimize the number of wavelengths required to the ring with hop-count constraint. Since information forwarding from lightpath to lightpath is performed in the electronic domain, frequent optical-electronic transformation would be required. Therefore it may downgrade the performance of transmission networks such as delay and throughput. Due to the computational complexity for finding an optimal solution in large-scale network within a reasonable time, a heuristic method is proposed. The author deals with the traffic grooming in the first stage, and designs the routing for each lightpath at the second stage. The computation time of the proposed heuristic was agreeable both in speed and quality. However, the number of wavelengths on a single fiber is assumed to be unlimited under this model. If the whole design process can take this into consideration, it would be more practical.

Jian Wang, et al. [28], first provided the formal mathematical specifications of the traffic grooming problem in several ring networks, i.e., single-hop and multi-hop cases of unidirectional and bidirectional rings. Then they proposed a simulated-annealing-based traffic grooming algorithm for the single-hop case and a greedy heuristic approach for the multi-hop case. It was shown that the greedy algorithm is usually good enough for non-uniform traffic when compared with the simulated-annealing method. Though the multi-hop approach could reduce the number of ADMs when the grooming ratio is neither too small nor too large, it usually results in more wavelength usage due to the prolonged connection length.

1.3.4 Traffic Grooming in Mesh Networks

In [32], traffic grooming problem is formulated as an Integer Linear Program (ILP) optimization problem. The objective is to maximize network throughput for given static traffic demand, which is defined by the amount of traffic demands that are actually accepted, subject to network resource constraints. In other words, the model does not necessarily serve all the given traffic requirements. As a consequence, some problem instances, which may be not feasible for models that are constrained to serve all traffic requests, may be feasible here. It is somewhat like those models associated with dynamic traffic scenarios.

The authors assume that the common bifurcation is not allowed, and require the data traffic on a connection request always to be routed by the same route. In addition, unlimited mux/demux capability on each node is assumed. This might underestimate the total cost or overestimate the network resources.

Since RWA optimization problem was proved as NP-hard, it is easy to see that the traffic grooming problem in a mesh network is also a NP-hard problem. The authors use a small network topology as an example for obtaining ILP result. As expected, the multi-hop case leads to higher throughput than the single-hop case. However, fast heuristic approaches are needed for a large network or for a dynamic traffic pattern.

The problem of traffic grooming to reduce the number of transceivers for general topologies is considered in [14]. All costs except the cost of transceivers are neglected. The authors concentrated on the topology subproblem integrated with the routing subproblem. By assuming that the number of physical links available between neighboring nodes is large enough, the number of wavelengths in the system is not important, and hence, the RWA problem is neglected.

An interesting intuitive observation is pointed out that minimizing endpoints for a given traffic is closely related to maximizing traffic for a given amount of endpoint equipment. This notion is formalized into the key concept of a precise duality, and thus it is shown that the transceiver minimization problem is equivalent to a commodity flow problem. However, the duality is precise only under the assumption that the RWA problem can always be solved and the ability to carry traffic is limited only by available network capacity. In other words, the grooming granularity has no effect. Based on this, a heuristic algorithm is presented that involves starting with an initial topology and successively deleting lightpaths after rerouting the traffic they carry. The computational experiments show that the algorithm does not give the optimal solutions.

In [15], the routes for the traffic demands are assumed to be given already. The author tries to modify a grooming factor g, to find an optimal wavelength assignment and grooming such that the number of wavelengths required in the network is minimized. In addition, the lower-rate traffic demands are assumed to be at the same rate. Though this is not true in reality but it simplifies the procedure of modifying the grooming factor. A heuristic approximation algorithm based on binary search and the LP relaxation of integer linear programming is proposed. According to the simulation results, if the grooming factor g is small, the decrease of wavelengths is faster than the case where g is large.

J. Q. Hu [8] described the grooming program as a GRWA problem. GRWA means traffic grooming, routing, and wavelength assignment. The author first formulates the GRWA problem as a integer linear programming problem and then, proposes a decomposition method that divides the GRWA problem into two smaller problems: the traffic grooming and routing problem and the wavelength assignment problem, which can then be solved much more efficiently. However, the solution is only a near optimal solution for the original GRWA problem. The author provides that under some sufficient conditions, the decomposition method did give an optimal solution.

The GRWA problem is divided into the traffic grooming and routing (GR) problem, and the wavelength assignment (WA) problem. The GR problem is a smaller ILP problem, and by relaxing some of its integer constraints, the computational efficiency is significantly improved.

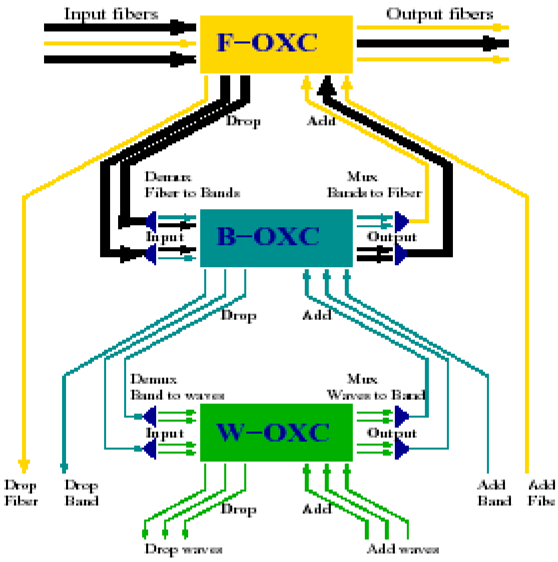


Figure 1-6 A 3-layer Switch

G. Huiban, et al, presented a framework to study the traffic grooming problem with multi-layer switches [12]. A node is considered as an 3-layer switch. Figure 1-6 is the corresponding switch used. Each level in the hierarchy is called a *layer*. Typical layers are wavelengths, bands and fibers. The idea is wavelengths (W) are included in bands (B) that are included in fibers (F). The cost of a given node depends on the number of input and output ports of each layer. The authors try to minimize the overall cost of the network for a given traffic matrix.

In this model, the wavelength continuity is not concerned. Though the authors claimed this is a multi-layer model, they only define the grooming problem over two layers in order to generate "solvable" problems. Since there are several problem simplifications (greedy-pre-processing, the use of splits, pipe filtering), as for problems for k layers, the approximated solution will be even loose by solving the problem recursively.

R. Srinivasan and Arun K. Somani illustrated the significance of dynamic routing of fractional wavelength traffic based on request characteristics in [26]. Dynamic routing in WDM networks has received very little attention so far in the literature. It has been extensively studied in the context of Quality-of-Service (QoS) routing for single-channel networks. A single-channel network is equivalent of an optical network with one wavelength. The performance of three destination-specific routing algorithms, namely shortest-widest path, widest-shortest path, and fixed alternate path routing, are studied and a request-specific routing scheme called Available Shortest Path Routing (ASPR) is developed. ASPR considers the capacity requirement of a request and selects the shortest path among those that can accommodate the request. It is shown through the simulations that ASPR enhances the performance of the network with respect to utilization and fairness metrics as compared to destination-specific routing schemes.

Hong Huang, et al., proposed a hybrid wavelength and sub-wavelength routing scheme (HWSR) which aims at striking a right balance between the benefits of optical bypass in wavelength routing and multiplexing gain in sub-wavelength routing. The key idea of the scheme is to partition wavelength channels in the network into two sets: dedicated and shared channels. Wavelength routing takes place in the set of dedicated channels, where a wavelength channel on a certain link can only belong to one connection. Sub-wavelength routing takes place in the set of shared channels. In such case, a wavelength channel can be shared by multiple connections. HWSR scheme provides a cost-effective and flexible way for network service providers to accommodate a customer base with diverse bandwidth demand granularities.

1.4 Proposed Approach

We model the traffic grooming problem as an optimization problem. The mathematical programming problem is a nonlinear nonconvex mixed integer-programming problem. As we expected, the problem is by nature highly complicated and difficult.

To the best of our knowledge, the proposed approach is the first attempt to consider the traffic grooming mechanism for statistical traffic demands with QoS traffic types on WDM networks and we formulate it rigorously. We then apply the Lagrange relaxation method and the subgradient method to solve the problem.

Chapter 2 Problem Formulation

2.1 Problem Description

In this chapter, we want to provide a well formulated model to describe the grooming problem. This model can used for minimizing the overall blocking probability under WDM grooming networks without violating the QoS guarantees for each traffic demands.

Problem assumptions:

1. The transmission channels are separated into three levels: paths for traffic demands, virtual links and fiber links (physical links). A path for traffic demand connecting a specific O-D pair is composed of more than one virtual links. A virtual link or a logical link is a lightpath or an aggregation of lightpaths with different wavelengths connecting the same O-D pair. Both terms are interchangeably used in this paper. A lightpath, which utilizes one specific wavelength, employs several fiber links. A fiber link is the physical transmission link.
2. The traffic flow and QoS requirements of traffic demands between each O-D pair are given. QoS requirements are satisfied by the hop distance in terms of

number of virtual links passed for each O-D pair in the network.

3. The OXCs used in the optical network are lack of the capability of wavelength conversion.
4. All nodes have grooming capability.
5. Wavelength continuity

Table 2-1 Problem Assumptions

Given:

1. The optical layer topology.
2. The available number of wavelengths per fiber link.
3. The capacity per wavelength channel
4. The traffic demands for different types of each O-D pair
5. The space requirement of each class.
6. The number of available add/drop ports for each OXCs

Objective:

To minimize the overall blocking probability.

Subject to:

1. Capacity constraints of components in the WDM network.
2. Single-path routing constraints for traffic demands.
3. QoS constraints guaranteed by limiting the maximum hop distance for each class in the network.
4. Wavelength continuity constraint.
5. Add/drop port number constraints at each OXC.

To determine:

1. The selected path for each traffic demand in the network.

2. The routing and capacity of logical links that support those selected paths. 3. The wavelength assignment.

Table 2-2 Problem Description

2.2 Notation

Given Parameters	
Notation	**Descriptions**
N	The set of nodes
L	The set of physical links
J	The set of available wavelengths on each physical link
V	The set of candidate logical links
T_n	The set of candidate outgoing paths from node n
R_n	The set of candidate input paths to node n
W	The set of all O-D pairs
q^+	The first logical link in the path q
q^-	The last logical link in the path q
C	Per wavelength capacity
c_v	Capacity for logical link v
γ_w^s	Traffic arrival rate of type s for origin-destination pair w (session/sec)
δ_{qv}	1 if path q uses virtual link v; otherwise 0
σ_{vjl}	1 if lightpath v with lambda j uses physical link l; otherwise 0
H_q	All virtual links that path q uses
$H_{q,v}$	The set of all virtual links prior to v that path q uses
θ^s	The hop count limit for traffic demand of type s
a_n	Number of add ports at node n

d_n	Number of drop ports at node n
Q_w	The set of candidate paths for O-D pair w

Table 2-3 Notation of Given Parameters

Decision Variables	
Notation	**Descriptions**
B_v^s	The joint blocking probability for logical link v with type s.
x_q^s	1 if path q is used by traffic request of type s; otherwise 0.
z_{vj}	1 if lightpath v uses wavelength j; otherwise 0.
g_v^s	The estimated aggregation flow on the logical link v with type s.

Table 2-4 Notation of Decision Variables

2.3 Problem Formulation

Optimization Problem:

Objective function:

$$Z_{IP1} = \min \sum_{q \in Q_w} \sum_{w \in W} \sum_{s \in S} x_q^s \gamma_w^s (1 - \prod_{v \in H_q} (1 - B_v^s)) \quad \text{(IP1)}$$

subject to:

$$g_v^s = \sum_{q \in Q_w} \sum_{w \in W} x_q^s \delta_{qv} \gamma_w^s \prod_{v' \in H_{q,v}} (1 - B_{v'}^s) \quad \forall v \in V, s \in S \quad \text{(IP1.1)}$$

$$\sum_{q \in Q_w} x_q^s = 1 \quad \forall w \in W, s \in S \quad \text{(IP1.2)}$$

$$x_q^s = 1 \text{ or } 0 \quad \forall w \in W, q \in Q_w, s \in S \quad \text{(IP1.3)}$$

$$\sum_{v \in V} z_{vj} \sigma_{vjl} \le 1 \quad \forall l \in L, j \in J \quad \text{(IP1.4)}$$

$$z_{vj} = 1 \text{ or } 0 \quad \forall v \in V, j \in J \quad \text{(IP1.5)}$$

$$\sum_{q \in Q_w} \sum_{v \in V} x_q^s \delta_{qv} \le \theta^s \quad \forall w \in W, s \in S \quad \text{(IP1.6)}$$

$$\sum_{q \in T_n} \sum_{j \in J} z_{q^+ j} \le a_n \quad \forall n \in N \quad \text{(IP1.7)}$$

$$\sum_{q \in R_n} \sum_{j \in J} z_{q^- j} \le d_n \quad \forall n \in N \quad \text{(IP1.8)}$$

$$c \sum_{j \in J} z_{vj} = c_v \quad \forall v \in V \quad \text{(IP1.9)}$$

$$c_v \in \{0, 1C, 2C, \ldots, |J|C\} \quad \forall v \in V \quad \text{(IP1.10)}$$

$$B_v^s = \sum_{s=0}^{b^s - 1} q_v(c_v - s) \quad \forall v \in V, \forall s \in S \quad \text{(IP1.12)}$$

$$\sum_{s=1}^{|S|} g_v^s b^s q_v(i - b^s) = j q_v(i) \quad \forall i \in \{0, 1, 2, \ldots, c_v\} \quad \text{(IP1.13)}$$

$$q_v(i) = 0 \quad \text{for } i < 0 \quad \forall v \in V \quad \text{(IP1.14)}$$

$$\sum_{i=0}^{c_v} q_v(i) = 1 \quad \forall v \in V. \quad \text{(IP1.15)}$$

The objective function is to minimize the total blocking rate (to maximize the total throughput) in the network.

Constraint (1): The aggregate flow on the logical link v taking into account the traffic loss along the path.

Constraints (2) and (3): Require that the traffic data with different types of service between each O-D pair must be transmitted as a whole over exactly one candidate path.

Constraint (4) and (5): To ensure every wavelength in each physical link will be used at most by only one lightpath.

Constraint (6): The hop-count number of paths for each traffic demand with different types of service is within the hop-count limitations as defined.

Constraint (7): To ensure the number of each wavelength used by outgoing paths of each traffic demand must not exceed the number of add ports of the corresponding OXC.

Constraint (8): To ensure the number of each wavelength used by the incoming paths of each traffic demand must not exceed the number of drop ports of the corresponding OXC.

Constraints (9) and (10): The total capacity of a logical link.

Constraints (11)-(15):

B_v^s is the joint blocking probability for logical link *v* with type *s*. We assume that different traffic demands can be managed jointly with the resource complete sharing policy. We can calculate the blocking probability $B_{joint}^s(g_v, c_v)$ by the algorithm developed by Kaufman [13]. Let g_v^s be the offered load of traffic type *s*, and we have $g_v = \sum_{s \in S} g_v^s$. With spatial requirement being b^s, the blocking probability for traffic type *s* on virtual link *v* is:

$$B_v^s(g_v, c_v) = \sum_{s=0}^{b^s-1} q(c_v - s), \quad s = 1,...,|S|$$

and the distribution *q*(.) should satisfy the following constraints:

$$\sum_{s=1}^{|S|} g_v^s b^s q(i - b^s) = jq(i), \;\; i = 0,1,...,c_v$$

where $q(x) = 0$ for $x < 0$ and $\sum_{i=0}^{c_v} g(i) = 1$.

2.4 Simplified Formulation

Two approximations are made to reduce the degree of nonconvexity of Problem (IP1). Although the approximations basically will lead to overestimation of the call blocking rate for each O-D pair, it can be easily seen that the surrogate formulation is an accurate approximation when the blocking probability on each logical link is small. The first approximation is to replace the end-to-end call blocking probability $1 - \prod_{v \in H_q} (1 - B_v^s)$ for each path *q* by $\sum_{v \in H_q} B_v^s$. It can be easily shown that $\sum_{v \in H_q} B_v^s$ is an upper bound on $1 - \prod_{v \in H_q} (1 - B_v^s)$. In addition, if each B_v^s is small, the bound will be tight.

The second approximation is to ignore the traffic losses when aggregate link flows are calculated. There is an effect of overestimating the aggregate logical link flows and thus the call blocking probability for each logical link. It is also clear that when the value each B_v^s is small, this approximation leads to tight upper bounds on the logical link flows and call blocking probabilities.

Combining the above two approximations, the total blocking rate can be expressed as $\sum_{v \in V} \sum_{s \in S} g_v^s B_v^s$, which, as discussed above, is an upper bound on the exact total blocking rate as given in (IP1). A surrogate formulation of (IP1) with the aforementioned approximations is given below.

$$Z_{IP2} = \sum_{v \in V} \sum_{s \in S} g_v^s B_v^s \qquad \text{(IP2)}$$

subject to:

g_v^s	$=$	$\sum_{q \in Q_w} \sum_{w \in W} x_q^s \delta_{qv} \gamma_w^s$	$\forall v \in V, s \in S$	(IP2.1)
$\sum_{q \in Q_w} x_q^s$	$=$	1	$\forall w \in W, s \in S$	(IP2.2)
x_q^s	$=$	1 or 0	$\forall w \in W, q \in Q_w,$ $s \in S$	(IP2.3)
$\sum_{v \in V} z_{vj} \sigma_{vjl}$	$\le$	1	$\forall l \in L, j \in J$	(IP2.4)
z_{vj}	$=$	1 or 0	$\forall v \in V, j \in J$	(IP2.5)
$\sum_{q \in Q_w} \sum_{v \in V} x_q^s \delta_{qv}$	$\le$	θ^s	$\forall w \in W, s \in S$	(IP2.6)
$\sum_{q \in T_n} \sum_{j \in J} z_{q^+ j}$	$\le$	a_n	$\forall n \in N$	(IP2.7)
$\sum_{q \in R_n} \sum_{j \in J} z_{q^- j}$	$\le$	d_n	$\forall n \in N$	(IP2.8)
$c \sum_{j \in J} z_{vj}$	$=$	c_v	$\forall v \in V$.	(IP2.9)
c_v	$\in$	$\{0, 1C, 2C, \ldots, \|J\|C\}$	$\forall v \in V$	(IP2.10)
B_v^s	$=$	$\sum_{s=0}^{b^s - 1} q_v(c_v - s)$	$\forall v \in V, \forall s \in S$	(IP2.12)
$\sum_{s=1}^{\|S\|} g_v^s b^s q_v(i - b^s)$	$=$	$j q_v(i)$	$\forall i \in \{0, 1, 2, \ldots, c_v\}$	(IP2.13)
$q_v(i)$	$=$	0 for $i < 0$	$\forall v \in V$	(IP2.14)
$\sum_{i=0}^{c_v} q_v(i)$	$=$	1	$\forall v \in V$.	(IP2.15)

In Figure 2-1 and Figure 2-2, we illustrated the relations between blocking rate and its corresponding traffic load. The link capacity is 50 and the spatial requirement for class 1 and 2 are 1, and 11. As we can see, when B_v^s is a function of g_v^s, blocking rate B_v^s is not a convex function. It increases the complexity and difficulty to solve this problem. Thus, we establish a feasible region: A_v on each $v \in V$.

Figure 2-1 Blocking Rate for Class 1

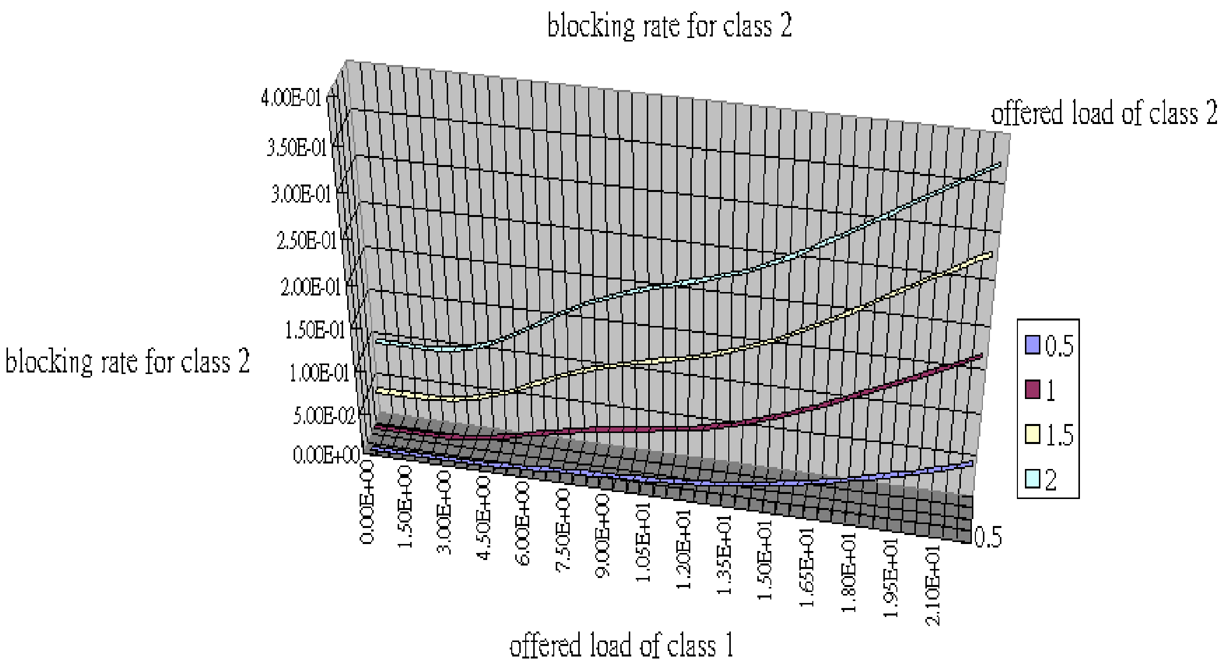


Figure 2-2 Blocking Rate for Class 2

In order to simplify the complex of the problem and make it easier to use Lagrangean relaxation method, we remodel our problem into:

$$Z_{IP3} = \sum_{v \in V} \sum_{s \in S} g_v^s B_v^s \qquad \text{(IP3)}$$

subject to:

$$\sum_{q\in Q_w}\sum_{w\in W} x_q^s \delta_{qv} \gamma_w^s \le g_v^s \qquad \forall v\in V, s\in S \qquad \text{(IP3.1)}$$

$$(g_v^s, s\in S)\in A_v \qquad \forall v\in V \qquad \text{(IP3.2)}$$

$$g_v^s \in G_v^s \qquad \forall v\in V, s\in S \qquad \text{(IP3.3)}$$

$$\sum_{q\in Q_w} x_q^s = 1 \qquad \forall w\in W, s\in S \qquad \text{(IP3.4)}$$

$$x_q^s = 1 \text{ or } 0 \qquad \forall w\in W, q\in Q_w, s\in S \qquad \text{(IP3.5)}$$

$$\sum_{v\in V} z_{vj}\sigma_{vjl} \le 1 \qquad \forall l\in L, j\in J \qquad \text{(IP3.6)}$$

$$z_{vj} = 1 \text{ or } 0 \qquad \forall v\in V, j\in J \qquad \text{(IP3.7)}$$

$$\sum_{q\in Q_w}\sum_{v\in V} x_q^s \delta_{qv} \le \theta^s \qquad \forall w\in W, s\in S \qquad \text{(IP3.8)}$$

$$\sum_{q\in T_n}\sum_{j\in J} z_{q^+ j} \le a_n \qquad \forall n\in N \qquad \text{(IP3.9)}$$

$$\sum_{q\in R_n}\sum_{j\in J} z_{q^- j} \le d_n \qquad \forall n\in N \qquad \text{(IP3.10)}$$

$$c\sum_{j\in J} z_{vj} = c_v \qquad \forall v\in V. \qquad \text{(IP3.11)}$$

$$c_v \in \{0, 1C, 2C, \ldots, |J|C\} \qquad \forall v\in V \qquad \text{(IP3.12)}$$

$$B_v^s = \sum_{s=0}^{b^s-1} q_v(c_v - s) \qquad \forall v\in V, \forall s\in S \qquad \text{(IP3.14)}$$

$$\sum_{s=1}^{|S|} g_v^s b^s q_v(i-b^s) = j q_v(i) \qquad \forall i\in\{0,1,2,\ldots,c_v\} \qquad \text{(IP3.15)}$$

$$q_v(i) = 0 \quad \text{for } i<0 \qquad \forall v\in V \qquad \text{(IP3.16)}$$

$$\sum_{i=0}^{c_v} q_v(i) = 1 \qquad \forall v\in V. \qquad \text{(IP3.17)}$$

In Constraint (IP3.3), G_v^s is a set that contains all the possible value of g_v^s. To solve this problem, we assume the elements in G_v^s to be discrete values. This makes Problem (IP3) to be the approximation of the original problem but it can improve the performance of the algorithm. By modifying the density of the elements in G_v^s, we can improve the preciseness of Problem (IP3).

Chapter 3 Solution Approach

3.1 Introduction to Lagrangean Relaxation Method

Lagrangean method is originally used in scheduling and the general integer programming problems in 1970's [4], because it can provide proper solutions for those problems. Gradually, it became one of the best tools for optimization problems, such as integer programming, linear programming combinatorial optimization, and non-linear programming.

Lagrangean relaxation has several advantages. For example, we can utilize Lagrangean relaxation to decompose mathematical models in many different ways. By doing so, the solution approach becomes very flexible. And thus, the complexity of the original problem can be considerably reduced and those decomposed subproblems can be solved separately as stand-alone problems. As for solving subproblems, we can optimally solve them by using any proper algorithm [4][7].

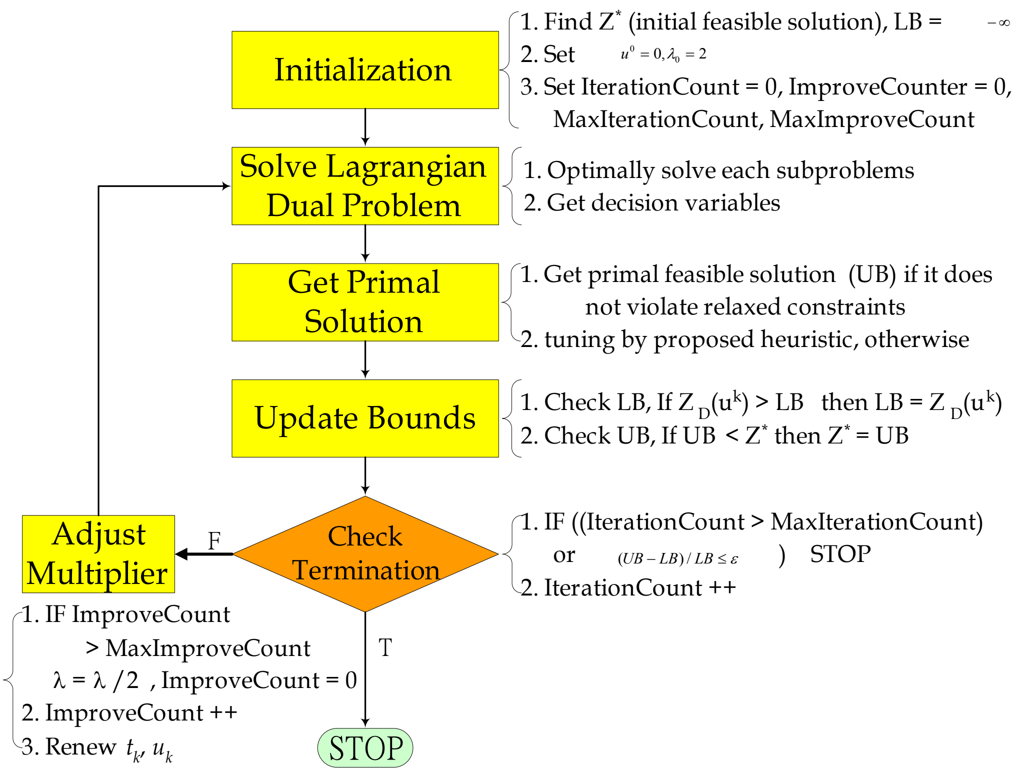


Figure 3-1 Lagrangean Relaxation Algorithm

Lagrangean relaxation helps us to find the boundary of the objective function. Also, it can be exploited to develop heuristic algorithms so as to get feasible solutions. It is a flexible solution strategy that permits modelers to make use of the underlying structure in any optimization problems by relaxing complicated constraints.

What's more, it permits us to "pull apart" models by removing constraints and instead, places them in the objective function with associated Lagrangean multipliers. The optimal value of the relaxed problem is always a lower bound (for minimization problems) on the objective function value of the problem. To obtain the best lower bound, we need to choose the minimization multiplier so that the optimal value of the Lagrangean subproblem is as large as possible.

We can solve the Lagrangean multiplier problem in a variety of ways. The subgradient optimization technique is possibly the most popular technique for solving the Lagrangean multipliers problem [4][7]. The procedure of Lagrangean relaxation is illustrated in Figure 3-1.

3.2 Lagrangean Relaxation

By using the Lagrangean relaxation method, we can transform the primal problem (IP2) into the following Lagrangean relaxation problem (LR) where constraints (IP3.1), (IP3.6), (IP3.9), (IP3.10) and (IP3.11) are relaxed. With a vector of non-negative Lagrangean multipliers, a Lagrangean relaxation problem of IP3 is given below.

Optimization problem (LR):

$$Z_D(u_1,u_2,u_3,u_4,u_5) = \min \sum_{v\in V}\sum_{s\in S} g_v^s B_v^s + \sum_{v\in V}\sum_{s\in S} u_{1vs}\left(\sum_{q\in Q_w}\sum_{w\in W} x_q^s \delta_{qv}\gamma_w^s - g_v^s\right) + \sum_{v\in V} u_{2v}\left[c\sum_{j\in J} z_{vj} - c_v\right]$$
$$+\sum_{n\in N} u_{3n}\left[\sum_{q\in T_n}\sum_{j\in J} z_{q^+ j} - a_n\right] + \sum_{n\in N} u_{4n}\left[\sum_{q\in R_n}\sum_{j\in J} z_{q^- j} - d_n\right] + \sum_{l\in L}\sum_{j\in J} u_{5lj}\left[\sum_{v\in V} z_{vj}\sigma_{vjl} - 1\right]$$

subject to:

		$(g_v^s, s\in S)\in A_v$	$\forall v\in V$	(LR.1)
		$g_v^s \in G_v^s$	$\forall v\in V, s\in S$	(LR.2)
$\sum_{q\in Q_w} x_q^s$	=	1	$\forall w\in W, s\in S$	(LR.3)
x_q^s	=	1 or 0	$\forall w\in W, q\in Q_w,$ $s\in S$	(LR.4)
z_{vj}	=	1 or 0	$\forall v\in V, j\in J$	(LR.5)

$$\sum_{q\in Q_w}\sum_{v\in V} x_q^s \delta_{qv} \quad \le \quad \theta^s \qquad \forall w\in W, s\in S \qquad \text{(LR.6)}$$

$$c_v \quad \in \quad \{0, 1C, 2C, \ldots, |J|C\} \qquad \forall v\in V \qquad \text{(LR.7)}$$

$$B_v^s \quad = \quad \sum_{s=0}^{b^s-1} q_v(c_v - s) \qquad \forall v\in V, \forall s\in S \qquad \text{(LR.8)}$$

$$\sum_{s=1}^{|S|} g_v^s b^s q_v(i-b^s) \quad = \quad jq_v(i) \qquad \forall i\in\{0,1,2,...,c_v\} \qquad \text{(LR.9)}$$

$$q_v(i) \quad = \quad 0 \quad \text{for } i<0 \qquad \forall v\in V \qquad \text{(LR.10)}$$

$$\sum_{i=0}^{c_v} q_v(i) \quad = \quad 1 \qquad \forall v\in V \qquad \text{(LR.11)}$$

where u_1, u_2, u_3, u_4, u_5 are the vectors of $\{u_{1vs}\}$, $\{u_{2v}\}$, $\{u_{3n}\}$, $\{u_{4n}\}$, $\{u_{5lj}\}$ respectively, and u_1, u_2, u_3, u_4, u_5 are Lagrange multipliers and $u_1, u_3, u_4, u_5 \ge 0$. To solve (LR), we continue to decompose (LR) problem into the following three independent and easily solvable optimization sub-problems.

3.2.1 **Subproblem 1** (related to decision variable B_v^s, g_v^s, c_v)

$$Z_{sub1}(u_1,u_2) = \min \sum_{v\in V}\sum_{s\in S} g_v^s B_v^s - \sum_{v\in V}\sum_{s\in S} u_{1vs} g_v^s - \sum_{v\in V} u_{2v} c_v$$

subject to:

$$(g_v^s, s\in S) \in A_v \qquad \forall v\in V \qquad \text{(sub1.1)}$$

$$g_v^s \in G_v^s \qquad \forall v\in V, s\in S \qquad \text{(sub1.2)}$$

$$c_v \quad \in \quad \{0, 1C, 2C, \ldots, |J|C\} \qquad \forall v\in V \qquad \text{(sub1.3)}$$

$$B_v^s \quad = \quad \sum_{s=0}^{b^s-1} q_v(c_v - s) \qquad \forall v\in V, \forall s\in S \qquad \text{(sub1.4)}$$

$$\sum_{s=1}^{|S|} g_v^s b^s q_v(i-b^s) \quad = \quad jq_v(i) \qquad \forall i\in\{0,1,2,...,c_v\} \qquad \text{(sub1.5)}$$

$$q_v(i) \quad = \quad 0 \quad \text{for } i<0 \qquad \forall v\in V \qquad \text{(sub1.6)}$$

$$\sum_{i=0}^{c_v} q_v(i) \quad = \quad 1 \quad \forall v \in V. \qquad \text{(sub1.7)}$$

This problem can be decomposed into |V| subproblems. For $g_v \in A_v$, find the g_{v^*}, which minimize $Z_{sub1}(u_1, u_2)$.

3.2.2 Subproblem 2 (related to decision variable x_q^s)

$$Z_{sub2}(u_1) = \min \sum_{v\in V}\sum_{s\in S} u_{1vs} \sum_{q\in Q_w}\sum_{w\in W} x_q^s \delta_{qv} \gamma_w^s$$

subject to:

$$\sum_{q\in Q_w} x_q^s \quad = \quad 1 \quad \forall w \in W, s \in S \qquad \text{(sub2.1)}$$

$$x_q^s \quad = \quad 1 \text{ or } 0 \quad \forall w \in W, q \in Q_w, s \in S \qquad \text{(sub2.2)}$$

$$\sum_{q\in Q_w}\sum_{v\in V} x_q^s \delta_{qv} \quad \le \quad \theta^s \quad \forall w \in W, s \in S \qquad \text{(sub2.3)}$$

Subproblem 2 is decomposed into $|W||S|$ shortest path problems. For each O-D pair w with service type *s*, u_{1vs} is used as the arc weights of path *q*. Because of the hop count constraint, we use Bellman-Ford shortest path algorithm [2][5] to solve the hop-constrained shortest path problem. Since the time complexity of Bellman-Ford algorithm is $O(|V| \times |E|)$, where |*E*| is the number of edges, the time complexity of this subproblem is $O(|W| \times |S| \times |V| \times |E|)$.

3.2.3 Subproblem 3 (related to decision variable z_{vj})

$$Z_{sub3}(u_2,u_3,u_4,u_5)=\min\sum_{v\in V}u_{2v}\left[c\sum_{j\in J}z_{vj}\right]+\sum_{n\in N}u_{3n}\left[\sum_{q\in T_n}\sum_{j\in J}z_{q^+j}\right]+\sum_{n\in N}u_{4n}\left[\sum_{q\in R_n}\sum_{j\in J}z_{q^-j}\right]$$
$$+\sum_{l\in L}\sum_{j\in J}u_{5lj}\left[\sum_{v\in V}z_{vj}\sigma_{vjl}\right]$$

subject to:

z_{vj}	$=$	1 or 0	$\forall v\in V, j\in J$	(sub3.1)
$\sum_{j\in J}z_{vj}$	$\le$	a_v	$\forall v\in V$	(sub3.2)
$\sum_{j\in J}z_{vj}$	$\le$	d_v	$\forall v\in V$	(sub3.3)

Let

a_{q^+}	$=$	a_n	$\forall q\in T_n, n\in N$
d_{q^-}	$=$	d_n	$\forall q\in R_n, n\in N$

We add two redundant constraints for subproblem 3 to improve the solution quality. Constraint (sub3.2) and (sub3.3) state that the number of wavelengths used by a virtual link v must not exceed the add/drop port number of the corresponding OXC. Constraint (sub3.2) and (sub3.3) are relaxed versions of Constraint (IP3.9) and (IP3.10).

We can reformulate the Subprobem 3 into the following formulation:

$$Z_{sub3}(u_2,u_3,u_4,u_5)=\min\sum_{v\in V}\sum_{j\in J}\left[u_{2v}c+u_{3v}+u_{4v}+\sum_{l\in L}u_{5lj}\sigma_{vjl}\right]z_{vj}$$

This problem can be further decomposed into |V| subproblems. Since the multiplier u_{2v} can be positive or negative, the following algorithm can optimally solve this subproblem:

Step 1. For each virtual link *v*, check if u_{2v} is positive. If so, set all its corresponding z_{vj} to 0, else go to step 2.

Step 2. For each virtual link with different wavelength, assign weight of u_{5lj} to physical links and apply Dijkstra's shortest path algorithm to get the value of σ_{vjl} and calculate the weight for each z_{vj}. If the weight is larger than 0, it's set to 0, otherwise set to1.

3.3 The Dual Problem and the Subgradient Method

According to the weak Lagrangean duality theorem [7], for any set of multipliers $(u_1, u_3, u_4, u_5) \geq 0$, $Z_D(u_1, u_2, u_3, u_4, u_5)$ is a lower bound on Z_{IP3}. The following dual problem is then constructed to calculate the tightest lower bound.

Dual Problem (D)

Z_D = **max** $Z_D(u_1, u_2, u_3, u_4, u_5)$

s.t. $(u_1, u_3, u_4, u_5) \geq 0$

There are several methods for solving the dual problem (D), of which the subgradient method [11] is the most popular and is adopted as our solution approach to the dual problem. Let a vector *k* be a subgradient of $Z_D(u_1, u_2, u_3, u_4, u_5)$. Then, in iteration *p* of the subgradient procedure, the multiplier vector $\lambda = (u_1, u_2, u_3, u_4, u_5)$ is updated by

$$\lambda^{p+1} = \lambda^p + t^p k^p$$

the step size t^p is determined by

$$t^p = \delta \frac{Z_{IP3}^{h} - Z_D(\lambda_p)}{\| k^p \|^2}$$

Z_{IP3}^{h} is the primal objective function value for a heuristic algorithm and δ is a constant, $0 \leq \delta \leq 2$. δ is initiated with a value of 2 and halved whenever the largest objective function value doesn't improve within *i* iterations, where *i* is the limit of unimprovement counter. The multipliers are initiated with a zero value and reset to λ^{p+1} if $\lambda^{p+1} \geq 0$, else set to 0.

Chapter 4 Getting Primal Feasible Solutions

4.1 Lagrangean Relaxation Results

By using Lagrangean relaxation and subgradient method to solve large scale optimization problems, we can get important information. It not only helps to get a theoretical lower bound of the primal problem, but also points out how to obtain feasible solution in each iteration of the dual problem.

If the calculated decision variables happen to satisfy the relaxed constraints in the primal problem, then a primal feasible solution is found. Otherwise, the infeasible primal solution will be used as a good starting point to get a primal feasible solution.

4.2 Getting Primal Heuristics

We will separate the original problem into three subproblms. At first, the construction of virtual layer topology is considered. And then, we deal with WDM network lightpath routing and wavelength assignment problem. At last, network traffic rerouting subproblem is considered.

4.2.1 Constructing Virtual Layer Topology Subproblem

Step 1. Sort all possible virtual links descendingly according to their aggregate flow obtained in Subproblem 2.

Step 2. For each virtual link, sort its corresponding z_{vj} ascendingly according to the weight, $u_{2v}c + u_{3v} + u_{4v} + \sum_{l \in L} u_{5lj}\sigma_{vl}$, which is calculated in Subproblem 3.

Step 3. Pick out the first z_{vj} in the list, and put it into a set denoted by *O* if it does not violate the add port constraint at the front node of the link and does not violate the drop port constraint at the tail node of the link. If it does not satisfy the constraints, we put it into another set denoted by *M*.

Step 4. If the list is not empty, go to Step 3, otherwise go to 4.2.2.

4.2.2 Lightpath Routing and Wavelength Assignment Subproblem

Step 1. Select the first z_{vj} in set *O*. Build the underlying lightpath by using weight u_{5lj} and running shortest path algorithm. If the corresponding lightpath can not be established for the specific z_{vj}, we drop it out of set *O*.

Step 2. If set *O* is not empty, go to Step 1, otherwise go to Step 3.

Step 3. Select the first z_{vj} from set *M* and try to build the lightpath too if it does not violate the add port constraint at the front node of the link and does not violate the drop port constraint at the tail node of the link.

Step 4. If set *M* is not empty, go to Step 3, otherwise go to Step 5.

Step 5. Calculate the capacity of each virtual link according to the underlying established lightpaths.

4.2.3 Rerouting Network Traffic Subproblem

Step 1. Check if the capacity of each virtual link equals to the c_v obtained in Subproblem 1. If not, recalculate the best traffic demand distribution for the capacity.

Step 2. For each virtual link *v,* check if the aggregate flow calculated in Subproblem 2 is larger than the demand distribution in Step 1. If so, reroute OD pairs that use virtual link *v* using RE_ROUTE algorithm.

Step 3. Use ADV_ROUTE to adjust the distribution of the traffic flow.

4.2.4 RE_ROUTE Algorithm

Step 1. Define all the O-D pairs that pass through this virtual link as a set w^v.

Step 2. Sort O-D pairs in set w^v descendingly according to their traffic demand.

Step 3. Pop out the first O-D pair in set w^v and reroute its traffic request until the aggregate flow on *v* is no long larger than $g_v^s, \forall s \in S$.

4.2.5 ADV_ROUTE Algorithm

Step 1. Calculate the blocking rate for each virtual link, and sort them descendingly. Give each virtual link a flag as feasible.

Step 2. Pick out the one with largest blocking rate and a feasible flag.

Reroute all the O-D pairs that pass through this virtual link. If feasible, go to step 3, else we set the flag of the link to infeasible.

Step 3. Check if all the flags of virtual links are infeasible. If so, it means that we don’t have to reroute anymore. If not, continue to step 2.

Chapter 5 Computational Experiments

In order to show the difference between the results from our Lagrangean relaxation method and other primal heuristics, we would like to do some experiments for the purpose of showing the effectiveness of Lagrangean results. With the comparison of the results, we can not only examine the quality of the primal heuristics, but also get some implications from the Lagrangean multipliers to find a feasible solution.

5.1 Simple Algorithm (SA)

Step 1. Identify the set of direct virtual links between O-D pairs with traffic, denoted by O.

Step 2. Sort $v \in O$ descendingly based on total traffic on it.

Step 3. Set P to empty.

Step 4. Denote $v \in O$ with highest traffic as v_{high}. If there are several links with the same traffic, randomly select one.

Step 5. Build virtual link v_{high} by constructing the corresponding wdm path for each lambda. The number of wavelengths used cannot violate the add/drop port constraints. Add v_{high} to set P

Step 6. Delete v_{high} from O

Step 7. If set O is not empty, go to Step 4, else go to Step 8

Step 8. Route traffic demands on the established virtual topology.

Step 9. Apply ADV_ROUTE algorithm.

After applying the algorithm above, we can find the total blocking probability, which is the objective of our problem.

5.2 Lagrangean Relaxation Based Algorithm (LR)

We provide an algorithm that utilizes Lagrangean relaxation based heuristics to deal with this network design problem. The algorithm is as follows:

Step 1. Read configuration file to construct IP network topology and WDM wavelength network *G*.

Step 2. Initialization stage: Calculate constant parameters; assign Lagrangean Relaxation unimprovement counter to 50.

Step 3. Initialize multipliers.

Step 4. By using the given multipliers, optimally solve subproblem 1, subproblem 2 and subproblem 3 of Chapter 3 to get the value of Z_D.

Step 5. Apply the heuristics of Chapter 4, calculate the value of Z_{IP3}.

Step 6. If Z_{IP3} is smaller than Z_{IP3}*, we assign Z_{IP3} to Z_{IP3}*. Otherwise, we decrement 1 from the unimprovement counter.

Step 7. Recalculate step size and update Lagrangean relaxation multipliers.

Step 8. Iteration counter is incremented by 1. If iteration counter is beyond the threshold of the system, terminate the program. And, Z_{IP3}* is our best solution. Otherwise, repeat Step 4.

5.3 Parameters and Cases of the Experiment

Number of Iterations	1000	
Maximum Unimprovement Counter	50	
Begin to Get Primal Solution	100	
Initial Scalar of Step Size	2	
Class Number	2	
Space Requirement	Class 1	1
	Class 2	2
Hop Count Limit	Diameter of the network model	

Table 5-1 Initial Testing Parameters

Case	Lambda	Demand	Add/Drop Port	Network Model
1	4	0~3	6	NSF
2	8	0~7	15	NSF
3	4	0~3	6	Mesh
4	8	0~7	15	Mesh
5	4	0~3	6	GTE
6	8	0~7	15	GTE

Table 5-2 Test Cases With 10 Channels per Lambda

Case	Lambda	Demand	Add/Drop Port	Network Model
7	4	0~6	6	NSF
8	4	0~6	6	Mesh
9	4	0~6	6	GTE

Table 5-3 Test Cases With 20 Channels per Lambda

The following network topologies are used in the experiment: NSF network (Figure 5.1), mesh network (Figure 5.2), and GTE network (Figure 5.3).

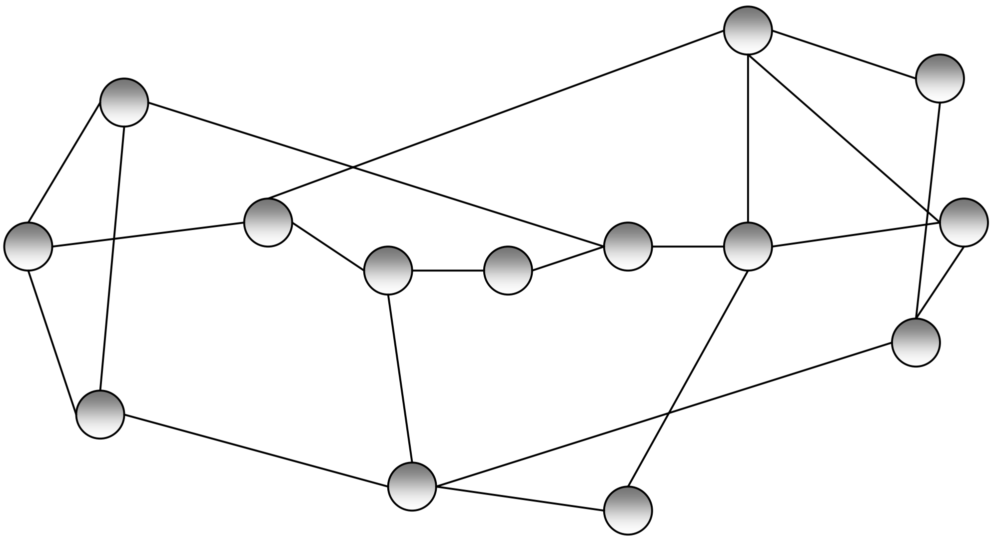


Figure 5-1 14-node 42-link NSF Network

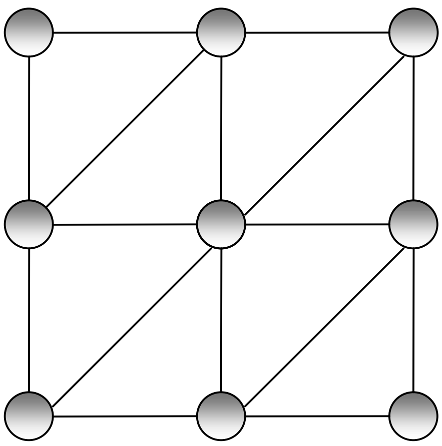


Figure 5-2 9-node 32-link Mesh Network

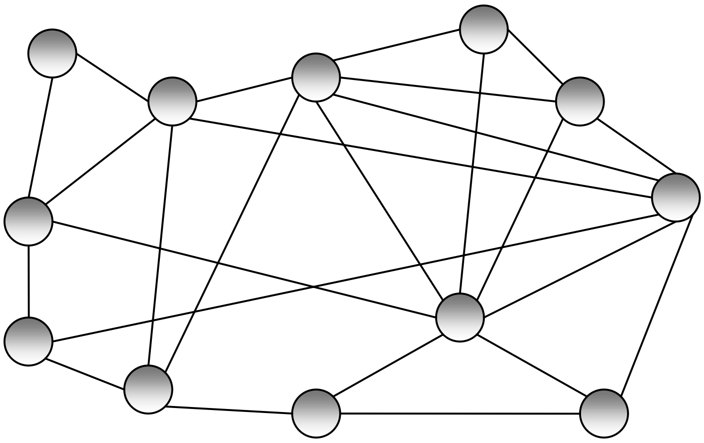


Figure 5-3　12-node 50-link GTE Network

5.3 Experiment Results

Case 1

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
3.157956	3.068887	2.903339	5.70	2.82046362
3.294065	3.135097	2.941949	6.56	4.82589141
3.393895	3.185464	3.027293	5.22	6.14135087

Table 5-4　The Result of Case 1

Case 2

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
11.149524	9.645071	8.271819	16.60	13.4934281
10.071403	9.113952	7.537180	20.91	9.50662981
12.134440	9.605261	8.620891	11.41	20.8429808

Table 5-5　The Result of Case 2

Case 3

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
5.164334	4.558854	3.930731	15.97	11.7242611
5.354078	4.744762	3.689849	28.58	11.3804095
5.387052	4.040760	3.919723	3.08	24.9912568

Table 5-6   The Result of Case 3

Case 4

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
7.906869	7.368961	5.149524	43.09	6.80304682
6.320506	5.268376	4.506935	16.89	16.6462938
8.045898	7.329010	6.693856	9.48	8.90998121

Table 5-7   The Result of Case 4

Case 5

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
8.903074	8.578479	6.825312	25.68	3.64587557
5.783097	5.364670	3.937740	36.23	7.23534466
8.830267	7.852991	5.796889	35.46	11.0673437

Table 5-8   The Result of Case 5

Case 6

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
9.793609	9.187989	8.030272	14.41	6.18382866
6.006056	5.068526	4.836856	4.78	15.6097446
7.904976	7.376418	6.367018	15.85	6.68639601

Table 5-9 The Result of Case 6

Case 7

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
4.423319	4.267376	3.847239	10.92	3.525475
4.380748	4.203608	3.662008	14.79	4.043602
4.392976	4.246701	3.500948	21.30	3.329747

Table 5-10 The Result of Case 7

Case 8

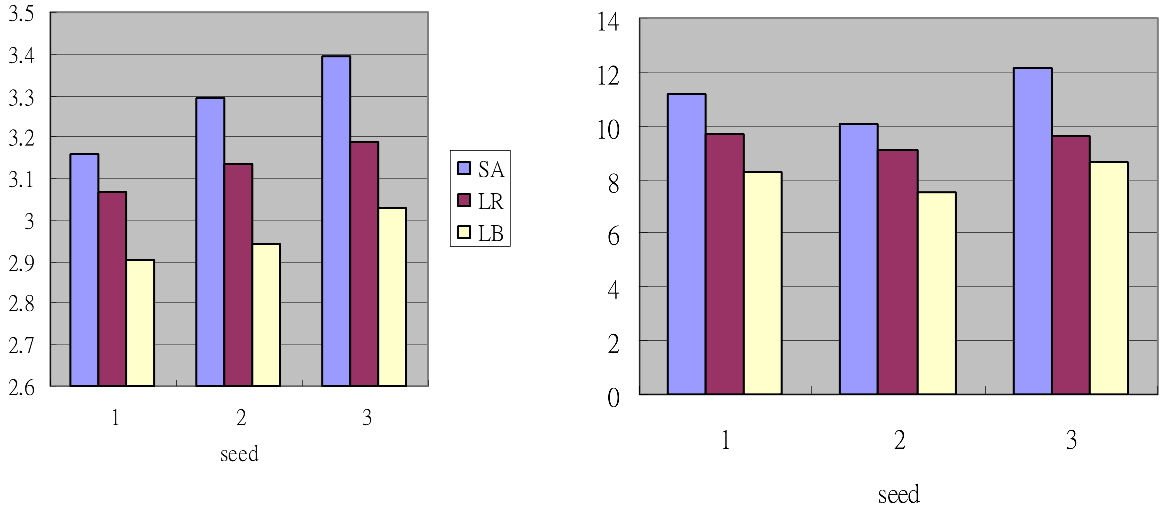
SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
2.562406	2.486273	1.916997	29.70	2.971153
2.431580	2.120436	1.643874	28.99	12.79596
2.512072	2.414360	1.819700	32.68	3.889697

Table 5-11 The Result of Case 8

Case 9

SA	LR	Lower Bound	Gap (%)	(SA-LR)/SA (%)
3.660636	3.454487	2.601587	32.78	5.631508
3.690342	3.500410	2.526022	38.57	5.146732
3.630515	3.560664	2.700352	31.86	1.923997

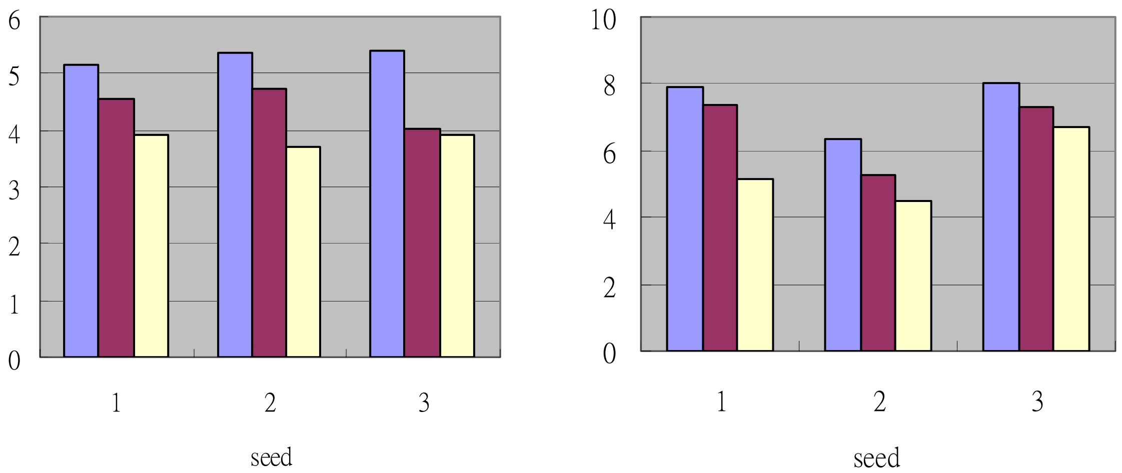
Table 5-12 The Result of Case 9



[a] **[b]**

Figure 5-4 Comparisons of Different Number of Lambdas in NSF Network

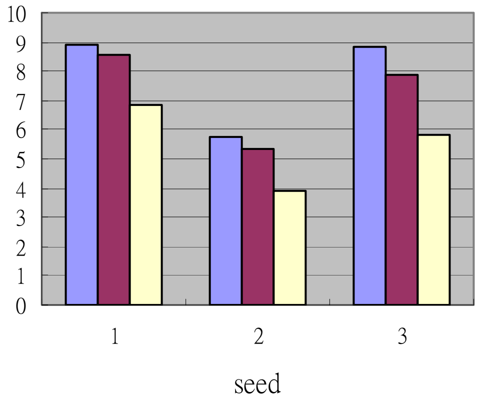
[a] Case 1; [b] Case 2



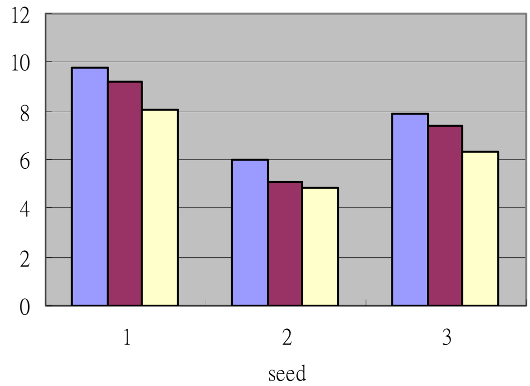
[a] **[b]**

Figure 5-5 Comparisons of Different Number of Lambdas in Mesh Network:

[a] Case 3; [b] Case 4



[a]



[b]

Figure 5-6 Comparisons of Different Number of Lambdas in GTE Network:

[a] Case 5; [b] Case 6

5.5 Computation Time

Model	Lambda	LR	Avg. Time	SA
NSF	4	507.716667	0.507717	0.346667
	4	504.583333	0.504583	0.254099
	4	504.466667	0.504467	0.392374
	8	6190.683333	6.190683	0.725538
	8	6197.816667	6.197817	0.682332
	8	6194.983333	6.194983	0.764137
Mesh	4	194.650000	0.194650	0.158969
	4	194.383333	0.194383	0.110509
	4	194.266667	0.194267	0.116012
	8	2422.566667	2.422567	0.497544
	8	2389.416667	2.389417	0.513333
	8	2417.783333	2.417783	0.566910
GTE	4	367.083333	0.367083	0.177213
	4	361.916667	0.361917	0.142499
	4	361.133333	0.361133	0.167018
	8	4458.716667	4.458717	0.397533
	8	4444.833333	4.444833	0.481210
	8	4460.833333	4.460833	0.345089

Table 5-13 Computation Time When Channels per Lambda is 10

Model	Lambda	LR	Avg. Time	SA
NSF	4	2736.700000	2.736700	0.673811
	4	2746.500000	2.746500	0.654899
	4	2716.583333	2.716583	0.706174
Mesh	4	1483.566667	1.483567	0.561173
	4	1475.800000	1.475800	0.635419
	4	1493.333333	1.493333	0.572727
GTE	4	3739.733333	3.739733	0.689124
	4	3720.733333	3.720733	0.736652
	4	3719.466667	3.719467	0.678490

Table 5-14 Computation Time When Channels per Lambda is 20

The experiment platform is a pc with Pentium4 1.5GHz CPU and 256 MB DRAM, running Mandrake Linux 9.0 Edition. And the experiment program is written in programming language C.

The computation time of SA and LR is slightly different when the number of lambdas is small. However, as the number of lambdas increases or as the channels per wavelength increases, the average computation time of LR becomes larger than that of SA. The cause of the increasing difference in computation time is mainly because we use brute force algorithm to solve Subproblem 1 in the dual problem. We tried to get the best traffic allocation for each possible capacity in Subproblem 1. However, the SA algorithm does not take this into consideration.

5.6 Result Discussion

According to the result tables, results of LR algorithm are all better than SA algorithm. There are two main reasons that LR algorithm works better than SA algorithm. First, SA algorithm makes routing decision only based on the hop count and traffic demands, whereas LR makes use of multipliers, which includes the influence of O-D pairs, link capacity, wavelengths used, and path selected.

Second, LR is iteration-based and is guaranteed to improve the result iteration by iteration. Moreover, the result of each iteration can also be used as a good hint to improve the lower bound of the problem, which leads to good feasible solutions.

Chapter 6 Summary and Future Work

6.1 Summary

While the fiber-optic backbone networks gradually migrates from interconnected SONET rings to arbitrary mesh topology, traffic grooming capabilities on WDM mesh networks becomes an extremely important research problem.

To address this problem, we propose an approach to design traffic routing problem, lightpath routing and wavelength assignment on grooming WDM mesh networks, while taking into account various constraints.

The achievement of the thesis can be expressed in terms of experiment performance and mathematical formulation. In terms of experiment performance, Lagrangean relaxation based algorithm has more significant solutions to the problem as we can see from the experiment results. In terms of formulation, we model a precise mathematical expression to describe the overall design problem of traffic routing problem, lightpath routing and wavelength assignment on grooming WDM networks.

Different network topologies, including NSF network model, mesh network model and GTE network model are tested in experiment. Different number of lambdas, and different types of traffic demands are also used to make this thesis more generic. Due to the complexity of this problem, Lagrangean relaxation and subgradient method are used as our main methodology. When utilizing these mathematical tools, they provide us some hints to improve our heuristics.

6.2 Future Work

In this paper, we only solve the problem under the unicast environment. However, as WDM technology matures, multicast applications become increasingly popular. Supporting multicast at the WDM layer becomes an important and yet challenging topic. Based on the algorithms developed in this thesis, we can try to extend our work into the multicast environment. In addition, only joint traffic management is studied here because it can reduce the expected call blocking rate. However, separate traffic management can simplify the task of traffic management. So it's a good extension from this work too.

Hop count limitation is the only QoS measurement used here. In the future, we can add more QoS related constraints, for example, delay constraints, to make the model more generic. Moreover, the sharing policy of virtual links used in this research is complete sharing policy. The resource sharing policy is another important issue. Due to the variety of network demands, partial shared policy can be considered to improve the flexibility of the planning problem in the future.

Besides, if the underlying WDM network technology is getting more advanced, for example, the wavelength conversion is no longer a cost issue, and becomes widely available, we can modify our model to adapt this new characteristic of the WDM network.

References

[1] 魏子順，“承載多訊務型態之 ATM 網路中虛擬路徑及虛擬通道之配置”，國立台灣工業技術學院電子工程技術研究所碩士論文，民國八十五年。

[2] R. E. Bellman, “Dynamic programming,” NJ: Princetion University Press, 1957.

[3] Rudra Dutta and George N. Rouskas, “Traffic Grooming in WDM Networks: Past and Future,” *IEEE Network*, pp. 46-56, November/December 2002.

[4] M. L. Fisher, “The Lagrangean relaxation method for solving integer programming problems”, *Management Science*, vol. 27, pp. 1-18, 1981.

[5] L. R. Ford and D. R. Fulkerson, “Flows in Networks,” NJ: Princeton University Press, 1962.

[6] Aysequl Gencata and Biswanath Mukherjee, “Virtual-Topology Adaptation for WDM Mesh Networks Under Dynamic Traffic,” INFOCOM 2002, *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceeding*, vol. 1, pp. 48-56, 2002.

[7] A. M. Geoffrion, “Lagrangean relaxation and its use in integer programming”, *Math. Programming Study*, vol. 2, pp. 82-114, 1974.

[8] J.Q. Hu, B. Leida, “Traffic Grooming, Routing, and Wavelength Assignment in Optical WDM Mesh Networks,” manuscript, April 2002.

[9] Hong Huang, John A. Copeland, “Hybrid Wavelength and Sub-Wavelength Routed Optical Networks,” *Proc. IEEE Globecom*, 2001.

[10] Nen-Fu Huang, Guan-Hsiung Liaw, and Chuan-Pwu Wang, “A Novel All-optical Transport Network with Time-shared Wavelength Channels,” *IEEE Journal on Selected Areas in Communications*, vol.18, Issue 10, pp 1863-1875, October 2000.

[11] M. Held, *et al.*, “Validation of subgradient optimization,” *Math. Programming*, vol. 6,

pp.62-88, 1974.

[12] G. Huiban, S. Perennes and M. Syska, "Traffic Grooming in WDM Networks with Multi-Layer Switches," *IEEE International Conference on Communications*, vol. 5, pp. 2896-2901, 2002.

[13] Joseph S. Kaufman, "Blocking in a Shared Resource Environment," *IEEE Transactions on Communications*, pp. 1474-1481, October 1981.

[14] V. R. Konda, T. Y. Chow, "Algorithm for Traffic Grooming in Optical Networks to Minimize the number of Transceivers,", *IEEE Workshop on High Performance Switching and Routing*, pp. 218-221, 2001.

[15] Deying Li, Zhenqiang Sun, Xiaohua Jia and Sam Makki, "Traffic Grooming for Minimizing Wavelength Usage in WDM Networks," *Eleventh International Conference on Computer Communications and Networks*, pp. 460-465, 2002.

[16] F. Y.S. Lin and K.T. Cheng, "Virtual Path Assignment and Virtual Circuit Routing in ATM Networks," *Proc. IEEE INFOCOM*, San Francisco, California, April 1993.

[17] F. Y.S. Lin and K.T. Cheng, "On the Joint Virtual Path Assignment and Virtual Circuit Routing in ATM Networks," *Proc. IEEE GLOBEOCOM*, San Francisco, California, November 1994.

[18] M. Mellia, E.Leonardi, M. Feletig, R. Gaudino, F. Neri, "Exploiting OTDM technology in WDM networks," *Optical Fiber Communication Conference and Exhibit*, pp 605 -606, March 2002.

[19] Eytan Modiano, Philip J. Lin, "Traffic Grooming in WDM Networks," *IEEE Communications Magazine*, pp. 124-129, July 2001.

[20] Biswanath Mukherjee, Dhritiman Banerjee, S. Ramamurthy, and Amarnath Mukherjee, "Some Principles for Designing a Wide-Area WDM Optical Network," *IEEE/ACM Transactions on Networking.* vol. 4 no. 5, October 1996.

[21] Biswanath Mukherjee, "Optical Communication Networks," McGraw-Hill Companies,

Inc., 1997.

[22] Biswanath Mukherjee, "WDM Optical Communication Networks: Progress and Challenges," *IEEE Journal on Selected Areas in Communications,* vol. 18, no.10, October 2000.

[23] S. Ohta and K. I. Sato, "Dynamic Bandwidth Control of the Virtual Path in an Asynchronous Transfer Mode Network," *IEEE Transactions on Communications* 40-7, pp. 1239-1247, July 1992.

[24] D. A. Schupke, D. Sellier, "Lightpath Configuration of Transparent and Static WDM Networks for IP Traffic," *IEEE International Conference on Communications*, vol.2, pp. 494-498, 2001.

[25] R. Srinivasan, "Dynamic Routing in WDM Grooming Networks," *DCNL Technical Report*, 2001.

[26] R. Srinivasan and Arun K. Somani, "Request-Specific Routing in WDM Grooming Networks," *IEEE International Conference on Communications*, vol. 5, pp. 2876-2880, 2002.

[27] Jian Wang, V. Rao Vemuri, Wonhong Cho, Biswanath Mukherjee, "Improved Approaches for Cost-effective Traffic Grooming in WDM Ring Networks: Non-uniform Traffic and Bidirectional Ring," *Journal of Lightwave Technology*, vol. 19, Issue: 11, pp. 1295-1299, 2000.

[28] Jian Wang, V. Rao Vemuri, Wonhong Cho, Biswanath Mukherjee, "Improved Approaches for Cost-effective Traffic Grooming in WDM Ring Networks: ILP Formulations and Single-Hop and Multihop Connections," *Journal of Lightwave Technology*, vol. 19, no. 11, pp. 1645-1653, November 2001.

[29] Bo Wen and Krishna M. Sivalingam, "Routing, Wavelength and Time-Slot Assignment in Time Division Multiplexed Wavelength-Routed Optical WDM Networks," INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and

Communications Societies. Proceedings. IEEE, vol. 3, pp. 1442-1450, 2002.

[30] Moon-Gil Yoon, "Traffic Grooming and Light-path routing in WDM Ring Networks with Hop-count Constraint," *IEEE International Conference on Communications*, vol. 3, pp. 731-737, 2001.

[31] John Yu, Eileen Y. Zhou, "An Integrated DWDM Optical Network Design with Traffic Grooming," *IEEE International Conference on Communications*, vol. 4, pp. 2131-2136, 2002.

[32] Keyao Zhu and Biswanath Mukherjee "Traffic Grooming in an Optical WDM Mesh Network,", *IEEE International Conference on Communications*, vol. 3, pp.721-725, 2001.