Routing Policy on Robustness in WDM Optical Networks

by

Hungjen Wang

Submitted to the Technology and Policy Program
in partial fulfillment of the requirements for the degree of

Master of Science in Technology and Policy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2001

Author .....
Technology and Policy Program
May 11, 2001

Certified by .....
Muriel Medard
Assistant Professor
Thesis Supervisor

Certified by .....
Eytan Modiano
Assistant Professor
Thesis Supervisor

Accepted by .....
Daniel Hastings
Director, Technology and Policy Program

Routing Policy on Robustness in WDM Optical Networks

by

Hungjen Wang

Submitted to the Technology and Policy Program
on May 11, 2001, in partial fulfillment of the
requirements for the degree of
Master of Science in Technology and Policy

Abstract

Path protection and link protection schemes are the main means of protecting wavelength-division multiplexed (WDM) networks from the losses caused by a link failure such as a fiber cut. We propose a new protection scheme, which we term partial path protection (PPP), to select end-to-end backup paths using local information about network failures. PPP designates a different restoration path for every link failure of every primary path. PPP allows the re-use of operational segments of the original primary path in the protection path. Our study also consider to assign different protection paths to every segment, instead of every link, of a primary path to reduce the computational complexity. The result demonstrates that the network efficiency can be improved by using the local information of network failure. A novel approach used in this paper is that of a dynamic call-by-call model with blocking probability as the performance metric. This is in contrast with traditional approaches to restoration, which consider capacity-efficiency for batch call arrivals. Since optimizing the blocking probability is a large dynamic optimization problem, we present two heuristics for implementing PPP. We show that a simple method based on shortest path routing for which primary paths are selected first is more effective than a greedy approach that minimizes, for each call arrival, the number of wavelengths used by the primary and backup path jointly. The network reliability issues on the interconnection policy is investigated in our study. We will show that employing better protection schemes will intensify the competition in the industry.

Thesis Supervisor: Muriel Medard
Title: Assistant Professor

Thesis Supervisor: Eytan Modiano
Title: Assistant Professor

Acknowledgments

This is the acknowledgements section. You should replace this with your own acknowledgements.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A wide range of protection schemes for WDM networks have been investigated [1, 2, 3, 4, 5, 7, 9, 10, 12, 13, 14, 16, 17]. Among them, path protection and link protection have attracted the most attention [1, 10, 12, 13, 16]. Path protection requires the protection path of a request to be completely link-disjoint from the corresponding primary path, while the link protection scheme reroutes all affected requests over a set of replacement paths between the two nodes terminating the failed link. Primary capacity cannot be shared, but protection capacity can be shared as long as a single link failure does not activate more than one wavelength reserved for protection along any wavelength on any link. In general, path protection is more capacity efficient than link protection [12].

In this paper, we present a new protection scheme, the *partial path protection scheme* (PPP). In this scheme, the network identifies a specific protection path for each link along a considered primary path. Thus, similarly to the path protection scheme, the partial path protection scheme assigns "end-to-end" protection paths to primary paths. However, in PPP, one single protection path protects only one specific link failure on one primary path, instead of the whole primary path in path protection. Moreover, owing to the requirement of identifying the location of a link upon failure, the PPP uses the local information locating the link failure to activate corresponding backup paths for affected primary paths, while the path protection scheme ignores such local information since it assigns only one backup path for a primary path and

thus such a local information is unnecessary.

Our study also consider the scheme which assigns different end-to-end protection paths to every segment, instead of every link, of a primary path to decrease the computational complexity and the network management overhead. The scheme considered here is essentially the span protection [18]. Since the protection paths are established on the segmentation basis in this scheme, the needed information can be simplified as identifying the segment of a primary path affected by the link failure. Consequently, the amount (overhead) of network management information which offers local information is decreased. Furthermore, we will show that the span protection with shorter segments will do no worse than the one with longer segment in terms of network resource utilization.

In addition to establish the batch arrival model, we further consider a dynamic call-by-call system with random arrivals. The batch arrival model is reasonable when call demands are known in advance. However, static batch models do not allow for dynamic provisioning of primary and protection paths in the network. Our call-by-call dynamic model is well suited to dynamic allocation of capacity for primary and protection paths. In our call-by-call model, every new call establishes its primary and protection paths according to the traffic already present in the network when the call arrives. Given the dynamic and probabilistic nature of our model, we take the call blocking probability to be the performance metric for our schemes, rather than traditional capacity efficiency metrics.

In order to optimize call blocking probability for selecting paths using PPP over some time horizon, we would have to solve a dynamic optimization problem. The extremely large state space of a dynamic program over a reasonable network and time horizon renders such an approach impractical. The complexity of a dynamic programming approach prompts us to consider two heuristics for implementing PPP.

The first heuristic is a *greedy* approach that, for each call arrival, the system uses the fewest previously unused wavelengths to establish the primary and protection paths jointly. Wavelengths already used for protection paths can be used for new protection paths as long as a single link failure does not entail the activation of more

than one protection path on any wavelength on any link. The problem formulation is an integer linear program (ILP) [8], a common approach to network routing [4, 9, 10, 12, 14].

The second heuristic first selects the primary path, using a "shortest path" route. It then selects the protection paths using a shortest path algorithm in which wavelengths already assigned for protection can be used at no cost. The "shortest path" here refers to the minimum number of hops. Therefore, we term the whole of the second heuristic, involving the choice of primary and of protection paths, the *minimum hop* approach (MH).

We show that the MH approach is not only significantly simpler computationally than the greedy approach, but also more effective in terms of blocking probability. This result may seem surprising at first. However, since protection paths can share bandwidth, while primary paths cannot, it is reasonable to select the most economical primary first, as done by MH, rather than consider primary and protection bandwidth jointly, as done by the greedy algorithm. The MH approach, by selecting the primary path first, in effect prioritizes the efficient use of primary path resources over protection resources. The greedy approach seeks to minimize the total use of new wavelengths by primary and backup paths jointly. However, in a dynamic system, the efficient use of protection bandwidth is not as important as the efficient use of primary bandwidth, since *in the future*, protection bandwidth has a high likelihood of being shared, whereas primary bandwidth cannot be shared. The fact that MH performs better than the greedy approach highlights the significant difference between a dynamic call-by-call model and a static batch system.

In addition, we are also interested in the following two policy questions: why are the network operators willing to provide the network protection function and who incurs the additional cost of implementing the network protection schemes? By answering these questions, one can find that the provision of network protection plays an important role in the market structure.

The main contributions of our paper are the introduction of the PPP method for establishing protection paths, the introduction of the greedy and MH approaches for

implementing PPP and path protection and the use of a dynamic call-by-call model for protection. In the next chapter, we present PPP and related background. In Chapter 3, we present the formulations for batch and the call-by-call arrivals. The greedy and minimum hop approaches to implementing PPP and path protection will be presented. In Chapter 4, we present simulation results over several backbone networks to compare the performance, in terms of call blocking probability, of path protection and PPP using MH and the greedy algorithm. In Chapter 5, we investigate the incentives to design a robust WDM network from the technical, economic, and legislative aspects. We present our conclusions and directions for further research in Chapter 6.

Chapter 2

Protection schemes

In this section, we introduce PPP as well as span protection and compare them to path and link protection. We also discuss the issue of protection resource sharing.

2.1 Path protection and link protection schemes

There are two prevailing protection schemes to guard against link failure, path protection and link protection schemes. Path protection, as illustrated in Fig. 2-1, reserves network resources for a single protection path in addition to the primary path. Since it is impossible to foresee which link on the primary path will fail, the system allocates a protection path, which is completely link-disjoint from the primary path. The primary path therefore shares no common link with its associated protection path. When a link fails, the source and destination nodes of a call on the failed link are informed of the failure, and the communication is switched to the protection path.

Link protection, as shown in Fig. 2-2, reroutes all the connections on the failed link around it. When accepting a call request, the link protection scheme will reserve the network resource for the associated protection path. Note that the protection path connects the two nodes adjacent to the failed link. When a link failure occurs, the node adjacent to and upstream of the failed link immediately redirects the traffic along the predetermined protection path to the node on the other end of the failed link to restores transmission.

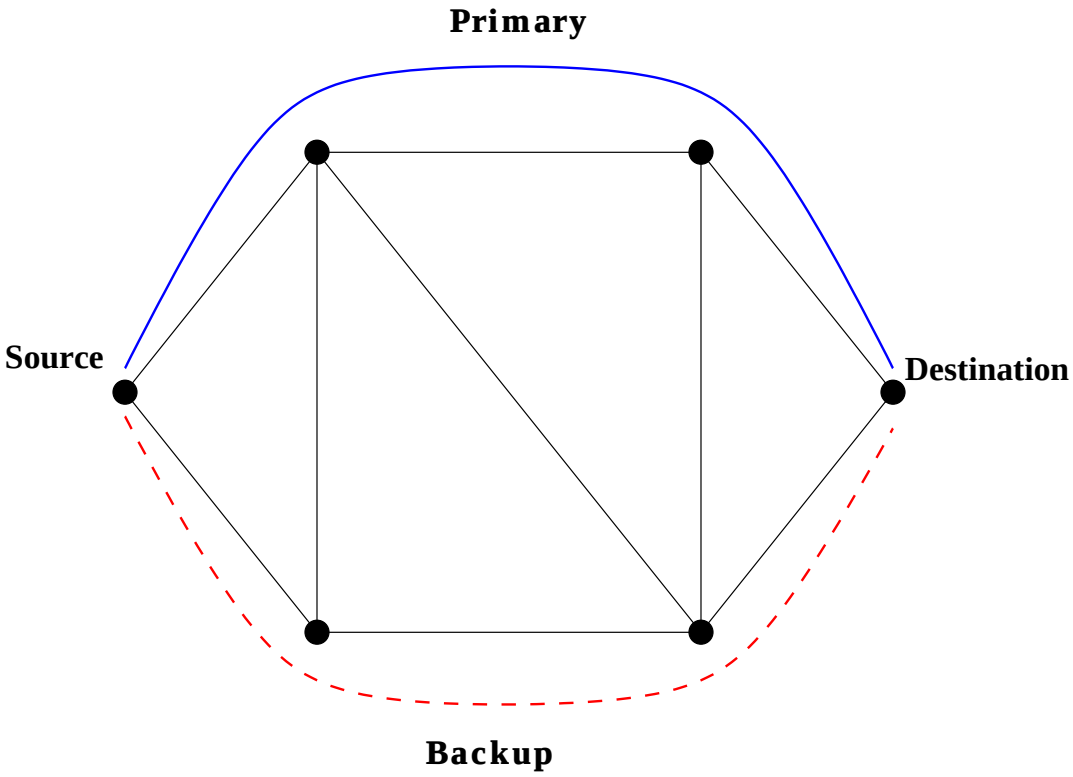

Figure 2-1: Path Protection

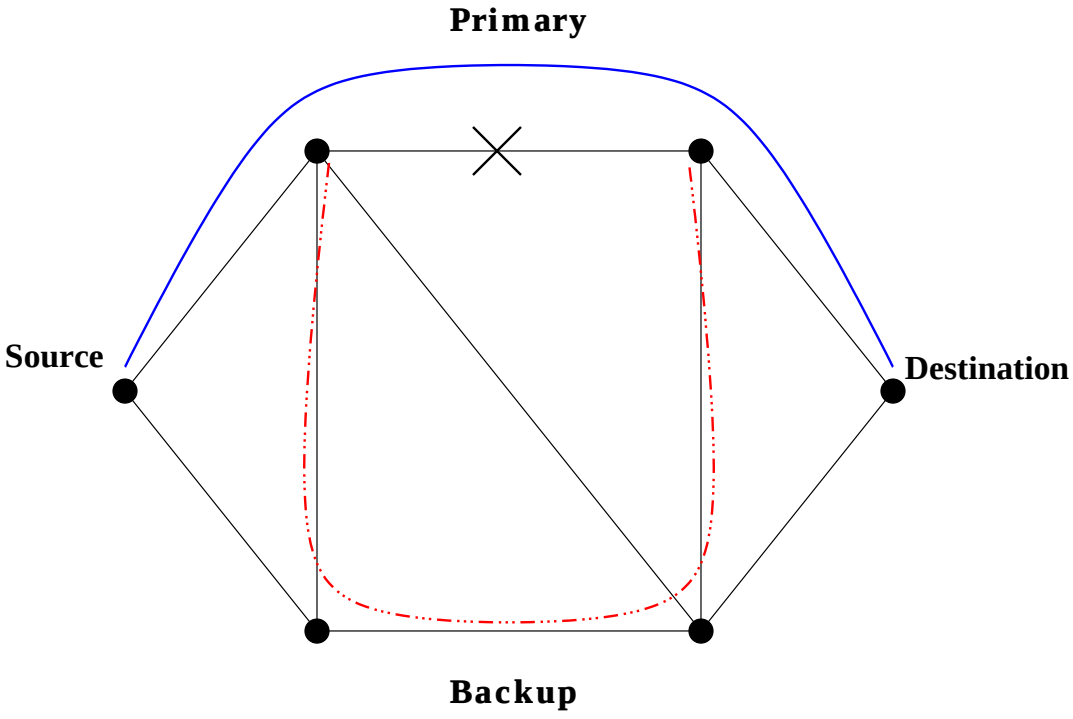

Figure 2-2: Prescheduled Link Protection

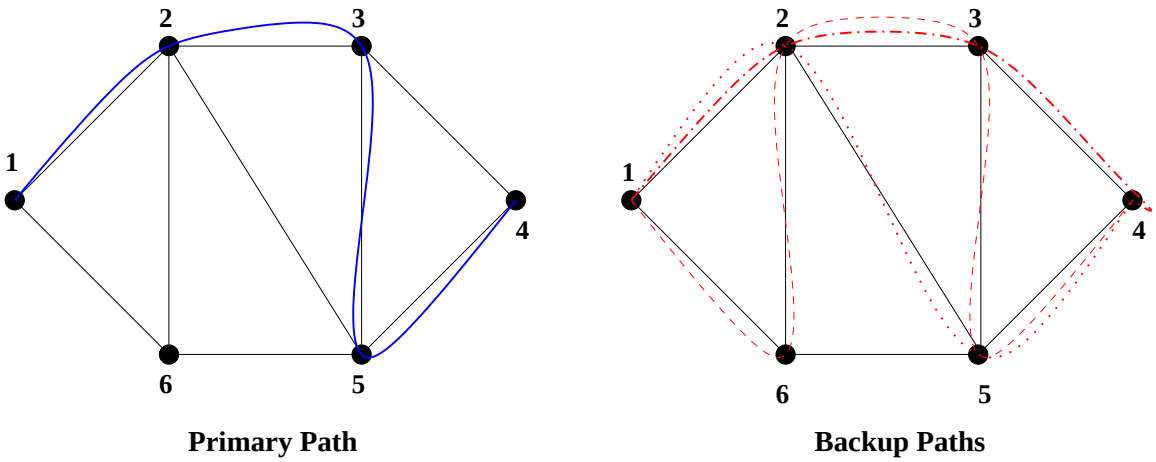

Figure 2-3: An Example for Partial Path Protection Scheme

2.2 Partial path protection scheme (PPP)

In PPP, the system reserves the protection resources while setting up a primary path. The major difference with path protection scheme is that the system now specifies a specific protection path for *each link* along the primary path. Thus, each protection path, rather than being associated with a single path as for the end-to-end path protection, or a single link as for link protection, is associated with a link/primary path pair. In the event of a link failure, the call is rerouted along the protection path corresponding to the failed link. For example, in Fig. 2-3, a call with source node 1 and sink node 4 has a primary path $1-2-3-5-4$. As illustrated in Table 2.1, the system applying PPP takes $1-6-2-3-5-4$ as the protection path against the failure of link $(1,2)$. Similarly, the network assigns $1-2-5-4$ to protect against the failures of links $(2,3)$ and $(3,5)$, and finally, $1-2-3-4$ to protect against the failure of $(5,4)$. Each of these protection paths needs only to be link-disjoint only from the link it protects.

2.3 Span protection

The *span protection* scheme considered in this paper is to provide different end-to-end protection paths to every segment, instead of every link, of a primary path. By

Link on Primary Path 1-2-3-5-4	Corresponding Protection Path
(1,2)	$1-6-2-3-5-4$
(2,3)	$1-2-5-4$
(3,5)	$1-2-5-4$
(5,4)	$1-2-3-4$

Table 2.1: Illustration of protection paths assigned by PPP for the primary path in Fig. 2-3

Segment on Primary Path $1-2-3-5-4$	Corresponding Protection Path
$1-2-3$	$1-6-5-4$
$3-4-5$	$1-2-3-4$

Table 2.2: Illustration of protection paths assigned by the span protection scheme for the primary path in Fig. 2-3

assigning the *segmentation nodes* in a network, we segment the primary paths passing through these specified nodes. Then for each segment of the primary path, the span protection assigns an end-to-end protection path which has no common link with the segment protected. For example, consider the primary path in Fig. 2-3, and we assign the node 3 as a segmentation node. Then the primary path is segmented into two sections: one from the source node to the segmentation node $(1-2-3)$ and the other one from the segmentation node to the destination node $(3-4-5)$. The span protection then assigns the associated protection paths to the two segments. For the first segment, $1-2-3$, the system reserves $1-6-5-4$ for its protection path, and for the second segment, $3-4-5$, the associated protection path is $1-2-3-4$. Note that each protection paths is completely link-disjoint from the corresponding segments it protects. Table 2.2 summarizes the resource assignment.

Path protection and partial path protection rest on the two opposite spectrum of the span protection. The former refers to the span protection with no segmentation node assigned in the network, whereas the partial path protection is the span protection under the condition that every node in the network is a segmentation node. In

this paper, we are of special interest in the path protection and the proposed PPP, since the two schemes represents two practical implementation in current networks; the path protection is commonly employed in most optical networks in addition to the link protection, and the partial path protection can be viewed as a circuit switching with the provision of the end-to-end protection. Therefore, we will focus on the discussion of path protection and PPP in the following paragraphs, and simulate span protection for better understanding about our idea of using local information to improve the network resource utilization.

Comparing PPP with path protection, we see that the former is more flexible than the latter. Indeed, any path protection scheme is a valid PPP, whereas the reverse does not hold. We expect, therefore, that PPP will enhance our ability to provide protection over traditional end-to-end path protection. To illustrate this fact, consider Fig. 2-3. By applying traditional end-to-end path protection, the network cannot find a protection path for the primary path shown. However, by applying PPP, we can provide protection service to the primary path. Since link protection schemes generally have a worse performance than path protection, we do not seek to compare PPP with link protection but only with traditional path protection.

2.4 Protection sharing

For path protection, a system can allow primary paths with no link in common to share protection bandwidth against a link failure, because we assume a single link failure can occur at a time. In addition to this type of bandwidth sharing, PPP allows a protection path to share bandwidth with portions of the primary path that remain operational after link failure. The following example illustrates the different levels of protection sharing for path protection and PPP.

Example 1 *Consider the network in Fig. 2-4 and assume the network is initially empty. The network now serves two call requests,* $(1, 5)$ *and* $(5, 4)$*, in sequence. Table 2.3 shows the resource assignments for primary and protection paths under the path protection and the PPP respectively. As shown in Table 2.3, the two primary*

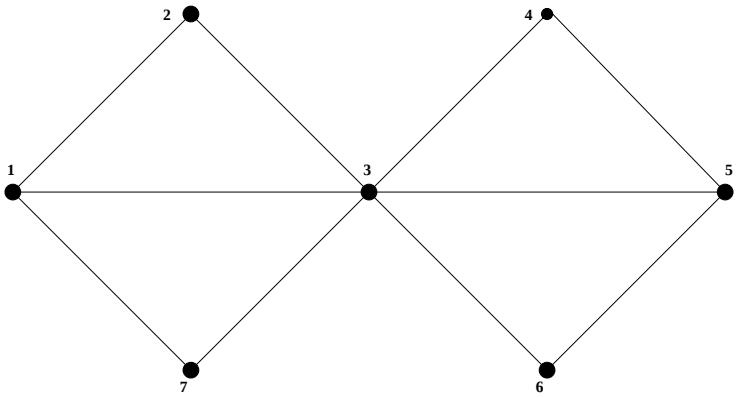

Figure 2-4: An example network for illustrating the partial path protection and path protection schemes in protection sharing

	SD Pair	Primary Path	Protection Path (protected link)	Total number of occupied λ's
Path Protection Scheme	(1,5)	1-3-5	1-2-3-4-5 (1-3)	6
			1-2-3-4-5 (3-5)	
	(5,4)	5-4	5-3-4 (5-4)	8 (share (3,4))
Partial Path Protection Scheme	(1,5)	1-3-5	1-2-3-5 (1-3)	6
			1-3-4-5 (3-5)	
	(5,4)	5-4	5-3-4 (5-4)	8 (share (3,4))

Table 2.3: Resource allocation for source destination pair $(1, 5)$ and $(5, 4)$ of the network in Fig. 2-4

paths, $1-3-5$ *and* $5-4$*, are completely link-disjoint from each other. By exercising protection sharing, the system reserves only one wavelength for protection on link* $(3, 4)$*, thus improving the network resource utilization.*

Example 1 illustrates the difference between path protection and PPP. Though the total number of occupied wavelengths to support the two requests is the same in both schemes, the protection wavelengths are used differently for path protection and for PPP. Consider, for example, link $(1, 2)$. In path protection, a wavelength on this link is assigned to protect link $(1, 3)$ and $(3, 5)$, while in PPP, the wavelength protects only the link $(1, 3)$. Hence, under PPP, this wavelength can be shared by a future call whose primary path includes link $(3, 5)$, but cannot be shared by using

path protection.

Chapter 3

Protection scheme formulation and path assignment approach

We model the path assignment problem for two arrival patterns in this chapter, the batch arrivals and the random call arrivals. For batch arrivals, we formulate models concerning path protection and PPP respectively. The formulation helps the system simultaneously assign primary as well as protection paths with minimum network resources to a batch arrival, a set of call requests. For the random call arrivals, the system serves only one call at a time. In this call-by-call model, the dynamics among call arrival not only reveals the dynamics in which the path assignment for current request is affected by previous resource allocation decisions but also shows how different protection schemes contribute different resource utilizations. A general formulation is relegated to Appendix A.

To practically attain the path assignment for batch arrivals, we formulate a mixed integer linear programming (MILP) problem. The reason we relax the integer constraints from the conventional integer linear programming (ILP) approach, which is common in this research area [4, 9, 10, 12, 14], is because it is impractical to solve an ILP problem of a large number of variables and constraints [8], and unfortunately the nature of a batch arrival problem inevitably yields a huge number of constraints. We relax such integer constraints by considering only a set of preselected primary and backup paths in the MILP for batch arrivals. That is, for each request in a batch

arrival, we preselect a few primary and corresponding backup paths and choose the one of the preselections as its resulting primary and backup paths by solving the formulated MILP problem.

Further, we consider two approaches to implement path protection, span protection, and PPP in the call-by-call model. In the case of path protection, there is a single protection path per primary path. In the case of PPP (span protection), there is a protection path for every link (segment resp.) in the primary path. The first approach we consider is the greedy approach, in which a system simultaneously allocates primary and protection paths to a new call by solving an ILP to minimize the use of previously unused wavelengths. The other approach is the *minimum hop* (MH) approach. The "minimum hop" implies the minimum number of wavelengths since a call occupies a whole wavelength for primary transmission when passes through a link. In MH, the system first assigns the path consist of minimum number of hops between a source and destination as a request's primary path. After having assigned the primary path, the system assigns each protection path (a single one in the case of path protection and possibly several ones in the case of PPP and span protection) using wavelengths that are free or shared with other protection paths. The cost of using each previously free wavelength is 1 and the cost of using a wavelength shared with other protection paths is 0. In the case of path protection, there is no sharing of wavelengths with the primary path. In the case of PPP and span protection, a path protecting a primary path can re-use at no cost a wavelength over an unfailed link in that primary path.

3.1 Problem Formulation for Batch Call Arrivals

To attain the optimum resource utilization, one can formulate an ILP problem to serve a set of R calls in the network using minimum number of wavelengths. Due to the difficulties of solving an ILP problem of a large scale, we formulate our problem into a mixed integer linear programming (MILP) problem by selecting the candidates of the primary and multiple corresponding backup paths for each source destination

requests in advance. This pre-selecting approach significantly reduces the demanding amount of computing power and machine memories, since it naturally eliminates huge amount of impractical path considerations. Solving the batch arrival problem thus becomes feasible. Note that, in order to concentrate on the topic of formulation, we ignore the wavelength continuity constraint at the beginning, and we will come back to this issue in Sec. 3.1.3.

3.1.1 Formulation for path protection

We first introduce the MILP formulation for the path protection. Let

$\mathcal{N}$	denote the number of nodes;
L	denote the set of all possible links;
R	denote the total number of connection requests;
W	denote the total number of wavelengths per link;
C	denote the total number of the *sufficient constraints*;
S_r	denote the source node of call r, $\forall r = 1, \ldots, R$;
D_r	denote the destination node of call r, $\forall r = 1, \ldots, R$;
$M(S_r, D_r)$	denote the total number of selections of primary path for call r, $\forall r = 1, \ldots, R$;
$N(M(S_r, D_r))$	denote the total number of backup path selections for the m^{th} preselected primary path of call r, $\forall r = 1, \ldots, R$.

For simplicity, we set $M(S_r, D_r) = M$ and $N(M(S_r, D_r)) = N$ in the following formulations. We also define

$$\beta_{i,j}^{w} = \begin{cases} 1, & \text{if wavelength } \lambda_w \text{ on link } (i,j) \text{ is reserved for backup restoration,} \\ 0, & \text{otherwise,} \end{cases}$$

$$x_{i,j}^{w,r} = \begin{cases} 1, & \text{if call } r \text{ reserves wavelength } \lambda_w \text{ on link } (i,j) \text{ for primary transmission,} \\ 0, & \text{otherwise,} \end{cases}$$

$$y_{i,j}^{w,r} = \begin{cases} 1, & \text{if call } r \text{ reserves wavelength } \lambda_w \text{ on link } (i,j) \text{ for backup restoration,} \\ 0, & \text{otherwise.} \end{cases}$$

$$\Phi_m^r = \begin{cases} 1, & \text{if call } r \text{ is served by its } m^{th} \text{ preselected primary path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$\Psi_{m,n}^r = \begin{cases} 1, & \text{if call } r \text{ is served by its } m^{th} \text{ primary path and the corresponding } n^{th} \text{ backup path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$P_{ij}^{m,r} = \begin{cases} 1, & \text{if call } r\text{'s } m^{th} \text{ primary path goes through link } (i,j), \\ 0, & \text{otherwise,} \end{cases}$$

$$Q_{i,j}^{m,n,r} = \begin{cases} 1, & \text{if the } n^{th} \text{ backup path of call } r\text{'s } m^{th} \text{ primary path goes through link } (i,j), \\ 0, & \text{otherwise.} \end{cases}$$

To know all selections of primary and backup paths is equivalent to know all link pairs, e.g. (l_p, l_b), in which whether a primary path passing through a common link (l_p) is possible to be protected by a wavelength on the other link (l_b). On the other hand, it is also equivalent to know, for each wavelength on a certain link, e.g. l_b, whether other links (say, l_p) may serve more than one primary paths which is protected by this considered wavelength on l_b. Once link l_p fails, then this link failure may cause the considered wavelength on l_b to support more than one connections, thus violating the protection sharing principle discussed in Sec. 2.4 . Therefore, in the MILP formulation for path protection, we enumerate all such link pairs in which a link reserved as backup purpose may protect more than one connections on the other link, and set up the *sufficient wavelength constraints* accordingly. The sufficient wavelength constraints guarantee that the number of wavelengths reserved on a certain link for protection is no less than the number of wavelengths for primary transmission on all the links protected by these restoration wavelengths. Note that this model does not specify which wavelength of a certain link on a preselected backup path to protect

the associated primary path, but only guarantees that the system reserves sufficient (and minimum) wavelengths for all possible link failures. In other words, this model assumes that, upon a link failure, the network management system will communicate to the nodes on the pre-selected backup paths, so as to allocate the wavelengths to the affected requests. Note that, since $\Psi^r_{m,n} = 1$ only if $\Phi^r_m = 1$ and $\Phi^r_m = 0$ guarantees that $\Psi^r_{m,n} = 0$, we replace Φ^r_m by $\Psi^r_{m,n}$ in the following formulation.

The MILP formulation is presented as follows.

$$\text{Minimize} \qquad \sum_{(i,j)\in L} \sum_{w=1}^{W} \sum_{r=1}^{R} x^{w,r}_{i,j} + \sum_{(i,j)\in L} \sum_{w=1}^{W} \beta^{w}_{i,j} \tag{3.1}$$

Eq.(3.1) is the sum of the number of wavelengths used as a primary transmission and the number of wavelengths reserved as backup resources, i.e., $\sum_{(i,j)\in L} \sum_{w=1}^{W} \sum_{r=1}^{R} x^{w,r}_{i,j}$ and $\sum_{(i,j)\in L} \sum_{w=1}^{W} \beta^{w}_{i,j}$ respectively. Clearly, to minimize the total resources assigned to serve a given batch arrival is to minimize the objective value of Eq.(3.1). Next, we present the constraint set.

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \Psi^r_{m,n} = 1, \quad \forall r = 1, \ldots, R \tag{3.2}$$

Eq.(3.2) requires the system to jointly assign a primary-backup selection pair for each call request.

$$\sum_{w=1}^{W} x^{r,w}_{i,j} = \sum_{m=1}^{M} (P^{m,r}_{i,j} \times (\sum_{n=1}^{N} \Psi^r_{m,n})), \quad \forall r = 1, \ldots, R, m = 1, \ldots, M, (i,j) \in L \tag{3.3}$$

$$\sum_{w=1}^{W} y^{w,r}_{i,j} = \sum_{m=1}^{M} \sum_{n=1}^{N} (Q^{m,n,r}_{i,j} \times \Psi^r_{m,n}), \quad \forall r = 1, \ldots, R, m = 1, \ldots, M, n = 1, \ldots, N, (i,j) \in L \tag{3.4}$$

Eq.(3.3) shows the relations between the choice of the primary-backup selection, $\sum_{n=1}^{N} \Psi^r_{m,n}$, and the actual wavelength assignment for primary transmission, $\sum_{w=1}^{W} x^{r,w}_{i,j}$. Through the given parameters $P^{m,r}_{i,j}$, which identify whether the m^{th} selection of primary path for request r rests on link (i,j), the system reserves a wavelength within

associated links for the primary transmission. Thus, if the m^{th} primary path is chosen to serve the request r, then one of the wavelengths on link (i,j) will be assigned to serve the primary transmission only if $P_{i,j}^{m,r} = 1$. Similarly, Eq.(3.4) practices the same function for backup paths.

$$\sum_{r=1}^{R} x_{i,j}^{w,r} \leq 1, \quad \forall w = 1, \ldots, W, (i,j) \in L \tag{3.5}$$

Eq.(3.5) prevents the system from assigning one wavelength to multiple primary paths.

$$\sum_{w=1}^{W} \beta_{i[c],j[c]}^{w} \geq \sum_{r=1}^{R} \sum_{m=1}^{M} \sum_{n=1}^{N} (\Psi_{m,n}^{r} \times P_{l[c],k[c]}^{m,r} \times Q_{i[c],j[c]}^{m,n,r}), \quad \forall c = 1, \ldots, C \tag{3.6}$$

Eq.(3.6) is the sufficient wavelength constraints discussed earlier. As discussed, for knowing all link pairs needed to take into consideration, we preselect and enumerate these link pairs as $(l_p[c], l_b[c]) = ((l[c], k[c]), (i[c], j[c]))$, for all $c = 1, \ldots, C$, and we thus have total C constraints. The importance of these constraints is that they provide the integer lower bounds for the variable $\beta_{i,j}^{w}$ which is not restricted to be an integer in the formulation. By formulating an MILP problem in this way, it is highly likely to obtain an integer optimal solution. Moreover, if the solution is non-integer, one adjacent integer vertex is guaranteed to reach the optimality [8].

$$\beta_{i,j}^{w} \geq y_{i,j}^{w,r}, \quad \forall w = 1, \ldots, W, r = 1, \ldots, R, (i,j) \in L \tag{3.7}$$

$$\sum_{r=1}^{R} x_{i,j}^{w,r} + \beta_{i,j}^{w} \leq 1, \qquad \forall w = 1, \ldots, W, \quad \forall (i,j) \in L \tag{3.8}$$

$$\sum_{w=1}^{W} \sum_{r=1}^{R} x_{i,j}^{w,r} + \sum_{w=1}^{W} \beta_{i,j}^{w} \leq W, \qquad \forall (i,j) \in L \tag{3.9}$$

Eq.(3.7) ensures that, for those links not considered in Eq.(3.6), sufficient wavelengths for restoration will be reserved. Eq.(3.8) shows that the primary and backup paths are completely link disjoint. Eq.(3.9) expresses the wavelength conservation constraint.

$$0 \leq x_{i,j}^{w,r}, y_{i,j}^{w,r}, \beta_{i,j}^{w} \leq 1, \qquad \forall w = 1, \ldots, W, \quad r = 1, \ldots, R, \quad (i,j) \in L \tag{3.10}$$

$$\Psi^r_{m,n} \in \{0,1\}, \qquad \forall m = 1,\ldots,M, n = 1,\ldots,N, r = 1,\ldots,R \quad (3.11)$$

At last, Eq.(3.10) and (3.11) denote the range and the integer constraints. Note that Ψ is the only variables restricted to be an integer in this formulation, while the other variables are relaxed from the integer constraints. By the nature of this formulation, we can always obtain a meaningful and optimal path assignments following the path protection for a batch arrival, i.e., an integer solution, by solving the MILP problem.

3.1.2 Formulation for partial path protection

In the batch arrival model, the formulation for PPP is similar to the one for path protection, except that the former assigns an end-to-end backup path to each link along each primary paths rather than to the entire primary path. Therefore, in addition to preselecting the primary paths for a given set of requests as done in the path protection formulation, we preselect end-to-end back paths for each link along the corresponding primary path. Consequently, in addition to the parameters defined in previous section, we have the following modified and new parameters.

$$y^{w,r,l}_{i,j} = \begin{cases} 1, & \text{if call } r \text{ reserves wavelength } \lambda_w \text{ on link } (i,j) \text{ to protect its } l^{th} \text{ link on} \\ & \text{the primary path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$\Psi^{r,l}_{m,n} = \begin{cases} 1, & \text{if call } r \text{ is served by its } m^{th} \text{ selection of primary path and the corresponding} \\ & n^{th} \text{ selection of backup path which protects its } l^{th} \text{ link along the primary path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$Q^{m,n,r,l}_{i,j} = \begin{cases} 1, & \text{if the } n^{th} \text{ backup path for protecting call } r\text{'s } l^{th} \text{ link on the } m^{th} \\ & \text{primary path goes through link } (i,j), \\ 0, & \text{otherwise.} \end{cases}$$

To identify the specific backup path for each link along each primary path, we modify the parameters y, Q and Ψ by introducing the *link* running variable l. Also note that,

in the definition of Q, we eliminate the link which is a portion of the primary path by setting its corresponding variable $Q_{i,j}^{m,n,r,l} = 0$. Additionally, we also let

$\mathcal{H}_m^r$ denote the total number of hops of the m^{th} primary path of call r,
$\forall m = 1, \ldots, M, r = 1, \ldots, R,$

ℓ_m^r denote the l^{th} node along the m^{th} primary path of call r,
$\forall m = 1, \ldots, M, \quad r = 1, \ldots, R.$

Such parameters are needed in PPP, since this scheme naturally distinguishes the backup path for each link on each primary path. In the formulation, we further use the variable $\hat{y}_{i,j}^r$ to denote whether the call r uses link (i, j) for protection in which the link (i, j) may be shared with other primary paths. The MILP formulation for placing the batch arrival in the PPP model is presented as follows.

$$\text{Minimize} \qquad \sum_{(i,j)\in L} \sum_{w=1}^{W} \sum_{r=1}^{R} x_{i,j}^{w,r} + \sum_{(i,j)\in L} \sum_{w=1}^{W} \beta_{i,j}^{w} \tag{3.12}$$

The objective function, Eq.(3.12) , is the same as the one in the formulation for path protection. The constraint set is presented as follows.

$$\textstyle\sum_{m=1}^{M} \Phi_m^r = 1, \qquad \forall r = 1, \ldots, R \tag{3.13}$$

$$\textstyle\sum_{n=1}^{N} \Psi_{m,n}^{r,l} = \Phi_m^r, \qquad \forall m = 1, \ldots, M,\ r = 1, \ldots, R,\ l = 1, \ldots, \mathcal{H}_m^r \tag{3.14}$$

Eq.(3.13) and (3.14) are for choosing the primary paths and corresponding backup paths on a link basis. Since the notations are complicated in this formulation, we explicitly represent Φ and Ψ for clarification, rather than replacing Φ by Ψ in the previous section.

$$\sum_{w=1}^{W} x_{i,j}^{r,w} = \sum_{m=1}^{M} (P_{i,j}^{m,r} \times \Phi_m^r), \quad \forall r = 1, \ldots, R, m = 1, \ldots, M, \quad (i, j) \in L \tag{3.15}$$

$$\sum_{w=1}^{W} y_{i,j}^{w,r} = \sum_{m=1}^{M} \sum_{n=1}^{N} Q_{i,j}^{m,n,r,l} \times \Psi_{m,n}^{r,l},$$

$$\forall l = 1, \ldots, \mathcal{H}_m^r, r = 1, \ldots, R, m = 1, \ldots, M, n = 1, \ldots, N, (i,j) \in L \quad (3.16)$$

$$\sum_{r=1}^{R} x_{i,j}^{w,r} \leq 1, \quad \forall w = 1, \ldots, W, \quad (i,j) \in L \quad (3.17)$$

Equations (3.15) to (3.17) are identical to the counterparts in the formulation of path protection. The *sufficient wavelength constraints* for PPP are presented as follows.

$$\hat{y}_{i[c],j[c]}^r \geq \Psi_{m,n}^{r,l} \times (P_{h[c],k[c]}^{m,r} + P_{\ell,\ell+1}^{m,r} - 1) \times Q_{i[c],j[c]}^{m,n,r,l}, \quad \forall c = 1, \ldots, C,$$
$$\forall l = 1, \ldots, \mathcal{H}_m^r, m = 1, \ldots, M, n = 1, \ldots, N, r = 1, \ldots, R \quad (3.18)$$

$$\sum_{w=1}^{W} \beta_{i[c],j[c]}^w \geq \sum_{r=1}^{R} \hat{y}_{i[c],j[c]}^r, \quad \forall c = 1, \ldots, C \quad (3.19)$$

Eq.(3.18) and (3.19) have the same effect as does Eq.(3.6), except that Eq.(3.18) and (3.19) further differentiate which links may be protected by the same wavelength on another link. As mentioned in the previous section, since knowing all possible primary paths and backup paths, we obtain the information concerning which pair of links for primary transmission and protection may violate the protection sharing principle. All the possible link pairs are enumerated as $(i[c], j[c])$, the link for protection, and $(h[c], k[c])$, the link for primary transmission, for all $c = 1, \ldots, C$. Thus we have total C sufficient constraints. However, other than knowing the link pairs, we must further ensure that the protected link exactly corresponds to the considered link for protection. Thus, $\hat{y}_{i[c],j[c]}^r$ equals to 1 for a certain number c and call r if and only if

(1) the link for primary transmission $(h[c], k[c])$ rest on the m^{th} selection of call r's primary paths;

(2) the link $(h[c], k[c])$ is the l^{th} link of the m^{th} primary path; in other words, the nodes $h[c]$ and $k[c]$ are equal to the nodes ℓ_m^r and $\ell_m^r + 1$;

(3) link $(i[c], j[c])$ rests on the n^{th} backup path of l^{th} link of call r's m^{th} primary path.

By above considerations, we obtain Eq.(3.18). By Eq.(3.19), we ensure that sufficient wavelengths are reserved to preserve the protection sharing principle.

$$\beta_{i,j}^{w} \geq y_{i,j}^{w,r,l}, \quad \forall w = 1, \ldots, W, r = 1, \ldots, R, l = 1, \ldots, \mathcal{H}_m^r, (i,j) \in L, \tag{3.20}$$

$$\sum_{r=1}^{R} x_{i,j}^{w,r} + \beta_{i,j}^{w} \leq 1, \qquad \forall w = 1, \ldots, W, \quad \forall (i,j) \in L, \tag{3.21}$$

$$\sum_{w=1}^{W} \sum_{r=1}^{R} x_{i,j}^{w,r} + \sum_{w=1}^{W} \beta_{i,j}^{w} \leq W, \qquad \forall (i,j) \in L, \tag{3.22}$$

$$0 \leq x_{i,j}^{w,r}, y_{i,j}^{w,r,l}, \beta_{i,j}^{w} \leq 1,$$

$$\forall (i,j) \in L, w = 1, \ldots, W, r = 1, \ldots, R, l = 1, \ldots, \mathcal{H}_m^r \tag{3.23}$$

$$\Phi_m^r, \Psi_{m,n}^r \in \{0,1\}, \quad \forall m = 1, \ldots, M, \quad n = 1, \ldots, N, \quad r = 1, \ldots, R. \tag{3.24}$$

Equations from Eq.(3.20) to Eq.(3.24) are identical to their counterparts in the formulation for path protection.

3.1.3 Wavelength continuity consideration

To highlight our work in solving the path assignment problem for the batch arrival, we did not include the wavelength continuity consideration in both formulations presented earlier. However, one can simply add Eq.(3.25) to the objective function Eq.(3.1) and Eq.(3.12) to consider the wavelength continuity constraint in the batch arrival models. The cost function for the wavelength continuity constraint is

$$\sum_{i \neq S_r, D_r}^{\mathcal{N}} \gamma_{1,i} \{ \sum_{j_1, j_2, j_1 \neq j_2}^{\mathcal{N}_i} \sum_{w=1}^{W} \sum_{r=1}^{R} (x_{j_1,i}^{w,r} - x_{i,j_2}^{w,r}) \} + \sum_{i \neq S_r, D_r}^{\mathcal{N}} \gamma_{2,i} \{ \sum_{j_1, j_2, j_1 \neq j_2}^{\mathcal{N}_i} \sum_{w=1}^{W} (\beta_{j_1,i}^{w} - \beta_{i,j_2}^{w}) \}, \tag{3.25}$$

where $\gamma_{1,i}$,$\gamma_{2,i}$ are two given constants, and $\mathcal{N}_i$ denotes the set of nodes connecting to node i by one link, for all $i \in \mathcal{N}$. By the principle of optimization, one can obtain solutions obeying wavelength continuity by solving the associated MILP problems, provided that the wavelengths are sufficient to support the given set of traffics and the constants $\gamma_{1,i}$,$\gamma_{2,i}$ are large enough [8].

One can also view Eq.(3.25) as a penalty function for violating the wavelength

continuity constraints. Then $\gamma_{1,i}$ represents the *shadow price* for placing a wavelength converter at node i for primary transmission, whereas $\gamma_{2,i}$ is the one for restoration.

3.2 Path assignment problem for random call arrivals

In many practical situations, calls do not arrive as a batch, but rather one at a time. Therefore, the formulations for arrivals on a call by call basis is more realistic and practical. We consider two approaches to implement path protection and PPP in this random call arrival model. The first approach we consider is the greedy approach, in which a system simultaneously allocates primary and protection paths to a new call by solving an ILP to minimize the use of previously unused wavelengths. The other approach is the MH approach. In MH, the system first assigns the shortest path, i.e., the path of minimum number of hops, connecting the source and destination nodes as a request's primary path. After having assigned the primary path, the system assigns each protection path (a single one in the case of path protection and possibly several ones in the case of PPP) using wavelengths that are free or shared with other protection paths. The cost of using each previously free wavelength is 1 and the cost of using a wavelength shared with other protection paths is 0. In the case of path protection, the is no sharing of wavelengths with the primary path. In the case of PPP, a path protecting a primary path can re-use at no cost a wavelength over an unfailed link in that primary path.

3.2.1 Greedy approach

To maximize network resource utilization, it is natural to seek to minimize the use of new resources for every call. We call this approach the greedy approach. We formulate the ILPs to realize path protection and PPP using the greedy approach. We first introduce the ILP formulation for path protection.

To begin with, we introduce the variables used in the formulation. Let

L denote the set of all possible links,

S denote the source node,

D denote the destination node,

$$c_{ij} = \begin{cases} 1, & \text{if at least one wavelength is available on link } (i,j) \in L, \\ \infty, & \text{otherwise,} \end{cases}$$

$$d_{ij}^{lk} = \begin{cases} 0, & \text{if at least one wavelength on link } (l,k) \text{ other than } (i,j) \text{ is already} \\ & \text{reserved to protect links other than } (i,j), \\ 1, & \text{else if at least one wavelength is available on link } (l,k) \in L, \\ \infty, & \text{otherwise,} \end{cases}$$

$$x_{ij} = \begin{cases} 1, & \text{if the primary path rests on an available wavelength in link } (i,j), \\ 0, & \text{otherwise,} \end{cases}$$

$$y_{ij} = \begin{cases} 1, & \text{if the system reserves a wavelength in link } (i,j) \text{ for protection,} \\ 0, & \text{otherwise,} \end{cases}$$

$$v_{ij}^{lk} = \begin{cases} 1, & \text{if a wavelength on } (l,k) \text{ is reserved to protect its associated primary path on } (i,j), \\ 0, & \text{otherwise.} \end{cases}$$

Note that, since we have no advance information about where the primary path will be placed, we need the variable d to indicate which links have wavelengths available to protect some specific link on which the primary path may reside. Furthermore, we also need the variable v to indicate the assignment of wavelengths to protection. The formulation of the ILP for a random call arrival is detailed below.

$$\text{Minimize} \qquad \sum_{(i,j)\in L} c_{ij}x_{ij} + \sum_{(i,j)\in L} y_{ij} \tag{3.26}$$

Eq.(3.26) represents the objective function, where c indicates whether a link has a free wavelength, x indicates the network resources for primary transmission and y

indicates the network resources reserved for protection. Notice that, in the ILP, the primary path and the protection path are considered concurrently. We next consider the constraint set.

$$\sum_{(S,j)\in L} x_{Sj} - \sum_{(j,S)\in L} x_{jS} = 1, \tag{3.27}$$

$$\sum_{(D,j)\in L} x_{Dj} - \sum_{(j,D)\in L} x_{jD} = -1, \tag{3.28}$$

$$\sum_{(i,j)\in L} x_{ij} - \sum_{(j,i)\in L} x_{ji} = 0, \quad \forall i \neq S, D, \tag{3.29}$$

$$\sum_{(S,l)\in L} v_{ij}^{Sl} - \sum_{(l,S)\in L} v_{ij}^{lS} \geq x_{ij}, \quad \forall (i,j) \in L, \tag{3.30}$$

$$\sum_{(l,D)\in L} v_{ij}^{lD} - \sum_{(D,l)\in L} v_{ij}^{Dl} \geq x_{ij}, \quad \forall (i,j) \in L, \tag{3.31}$$

$$\sum_{(l,k)\in L} v_{ij}^{lk} - \sum_{(k,l)\in L} v_{ij}^{kl} = 0, \quad \forall (i,j) \in L, \quad k \neq S, k \neq D, \tag{3.32}$$

Eq.(3.27) to Eq.(3.29) provide the flow conservation for the primary path. Similarly, Eq.(3.30) to Eq.(3.32) give the flow conservation for the protection path. Note that Eq.(3.30) to Eq.(3.31) are only active when the primary path passes through link (i,j), i.e., $x_{ij} = 1$.

$$v_{ij}^{ij} + v_{ji}^{ij} = 0, \quad \forall (i,j) \in L, \tag{3.33}$$

Eq.(3.33) enforces the path disjoint property.

$$y_{lk} \geq d_{ij}^{lk} v_{ij}^{lk}, \quad \forall (i,j),(l,k) \in L, \tag{3.34}$$

Eq.(3.34) indicates whether a unoccupied wavelength on link (l,k) will be reserved for protection. Notice that $v_{ij}^{lk} = 1$ and $d_{ij}^{lk} = 0$ together mean that sharing protection bandwidth is possible.

$$x_{ij} \geq v_{ij}^{lk}, \quad \forall (i,j),(l,k) \in L, \tag{3.35}$$

Eq.(3.35) prevents the possibility of assigning a protection path for a link that is not

used by the primary path.

$$v_{ij}^{lk} + x_{mn} \leq v_{mn}^{lk} + 1, \qquad \forall (i,j),(l,k),(m,n) \in L, \tag{3.36}$$

$$x_{ij}, y_{ij}, v_{ij}^{lk} \in \{0,1\}, \qquad \forall (i,j),(l,k) \in L. \tag{3.37}$$

Eq.(3.36) ensures that each link reserved for protection must also protect the whole primary path. For example, if a wavelength on link (l,k) is reserved to protect a primary path which passes through link (i,j), then we have $v_{ij}^{lk} = 1$. Since link (l,k) must also protect other links on the primary path, say link (m,n) ($x_{mn} = 1$), we need to set $v_{mn}^{lk} = 1$. If the primary path does not pass through link (m,n), i.e., $x_{mn} = 0$, then by constraint Eq.(3.35), $v_{ij}^{lk} = 0$ in this case. Hence, we assure the property that each link on a protection path protects every link of the associated primary path.

We next introduce the ILP formulation for PPP. Recall that, in this protection scheme, the system reserves a protection path for each link along the primary path and thereby the system reserves resources for one or multiple protection paths to protect the associated primary path.

The objective function for the path protection scheme remains the same for PPP. The constraint set of the formulation is as follows.

$$\text{Minimize} \quad \sum_{(i,j)\in L} c_{ij}x_{ij} + \sum_{(i,j)\in L} y_{ij}$$

$$\text{Subject to} \quad \sum_{(S,j)\in L} x_{Sj} - \sum_{(j,S)\in L} x_{jS} = 1, \tag{3.38}$$

$$\sum_{(D,j)\in L} x_{Dj} - \sum_{(j,D)\in L} x_{jD} = -1, \tag{3.39}$$

$$\sum_{(i,j)\in L} x_{ij} - \sum_{(j,i)\in L} x_{ji} = 0, \quad \forall i \neq S, D, \tag{3.40}$$

$$\sum_{(S,l)\in L} v_{ij}^{Sl} - \sum_{(l,S)\in L} v_{ij}^{lS} \geq x_{ij}, \quad \forall (S,l),(l,S),(i,j) \in L, \tag{3.41}$$

$$\sum_{(l,D)\in L} v_{ij}^{lD} - \sum_{(D,l)\in L} v_{ij}^{Dl} \geq x_{ij}, \quad \forall (D,l),(l,D),(i,j) \in L, \tag{3.42}$$

$$\sum_{(l,k)\in L} v_{ij}^{lk} - \sum_{(k,l)\in L} v_{ij}^{kl} = 0, \quad \forall (i,j) \in L, \forall k \neq S, k \neq D, \tag{3.43}$$

$$v_{ij}^{ij} + v_{ji}^{ij} = 0, \quad \forall (i,j) \in L, \tag{3.44}$$

$$y_{lk} \geq d_{ij}^{lk}(v_{ij}^{lk} - x_{lk}), \forall (i,j),(l,k) \in L, \tag{3.45}$$

$$x_{ij} \geq v_{ij}^{lk}, \quad \forall (i,j),(l,k) \in L, \tag{3.46}$$

$$x_{ij}, y_{ij}, v_{ij}^{lk} \in \{0,1\}, \forall (i,j),(l,k) \in L. \tag{3.47}$$

Note that the difference between the two formulations is that we transform Eq.(3.34) into Eq.(3.45), and we also remove Eq.(3.36) from the previous formulation. Eq.(3.45) considers the situation where a protection path overlaps part of its links with the links on its associated primary path. The overlap incurs no cost. We eliminate Eq.(3.36) from the formulation for the path protection scheme, because there is no need to force a link on a protection path to protect the entire primary path in PPP.

3.2.2 Minimum-Hop (MH) approach

Note that our system, with call arrivals and departures, is a discrete time system. The optimal solution can be obtained through dynamic programming, which would be prohibitively complex. The dynamic program takes into account the impact of present decisions on future system performance. The greedy algorithm only considers present resource usage, and thereby does not necessarily achieve optimality. The greedy approach can result in an inferior network resource utilization because it may choose paths with little opportunity for protection sharing (see Example 2). Therefore, we consider another implementation approach which encourages protection sharing as follows.

First, note that a request's primary path cannot be shared with other requests. Thus, it is natural to attempt to dedicate the fewest possible resources to a call's primary path. Therefore, we assign the path of minimum number of hops for a call request as its primary path. After the call's primary path is identified, we then seek the protection paths for it. To encourage protection sharing, we construct a new graph. In the new graph, the network topology remains intact but the link costs are updated according to the resource usage status. Wavelengths that are in use by other protection paths have a cost of 0. In the case of path protection, links used by the

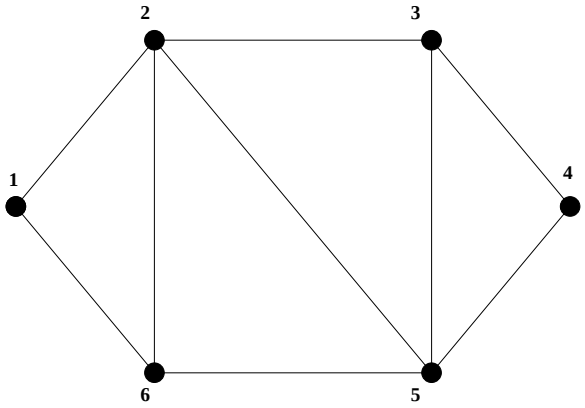

Figure 3-1: An example network

primary path are not available in the new graph. In the case of PPP, unfailed links in the primary path are available for no cost in the protection path.

3.2.3 Greedy versus MH

We may briefly compare our two implementation approaches. As mentioned, solving an ILP is computationally intensive. In contrast, since the algorithms for seeking the "shortest" paths, e.g. the Dijkstra's algorithm, are polynomial-time, the minimum-hop approach can place a new call rapidly. For static batch models, computational complexity is not very important, since decisions are not made in real time. For a dynamic call-by-call system, however, ease and speed of computation are more relevant.

Let us now consider resource efficiency. While the MH approach may at times require more resources for a given call, it is possible that over a number of calls, the MH approach may eventually result in more efficient bandwidth utilization. Example 2 illustrates this phenomenon.

Example 2 *Consider the network in Fig. 3-1 and assume that the network employs PPP. The network is initially empty and serves three call requests,* $(1, 4)$*,* $(6, 3)$*, and* $(3, 5)$*, in sequence. Table 3.1 shows the resource assignments for the greedy approach and the MH approach. In this example, the MH approach initially occupies more wavelengths to support the request* $(1, 4)$ *than does the greedy approach. However, as*

	SD Pair	Primary Path	Protection Path (protected link)	Total Number of Occupied λ's
Greedy approach	(1,4)	1-2-3-4	1-6-5-4 (1-2-3-4)	6 (no sharing)
	(6,3)	6-5-3	6-2-3 (6-5-3)	10 (no sharing)
	(3,5)	3-5	3-2-5 (3-5)	13 (no sharing)
Shortest path approach	(1,4)	1-2-3-4	1-6-2-3-4 (1-2) 1-2-5-4 (2-3) 1-2-5-4 (3-4)	7 (share (2-3-4)) (share (1,2))
	(6,3)	6-5-3	6-2-3 (6-5) 6-2-3 (5-3)	10 (share (6,2))
	(3,5)	3-5	3-2-5 (3-5)	12 (share (2,5))

Table 3.1: Resource usage for network employing partial path protection scheme implemented by different approaches in Fig. 3-1

the calls accumulate, the MH approach uses fewer number of wavelengths to support the same requests than the greedy approach.

In this example, the greedy approach endeavors to serve each request using the minimum number of previously unused wavelengths. However, in doing so, the greedy approach happens to choose paths with no protection sharing, harming network resource utilization. In contrast, though the MH is not optimal at first, it performs better over the call arrivals, by encouraging protection sharing.

Chapter 4

Simulation Results

To investigate the protection schemes, we not only solve the MILP for the batch arrivals within one step, but also simulate path protection, span protection and PPP schemes implemented using both the greedy approach and the MH approach in the random call model. We assume that the networks and the call requests have the following characteristics. First, all nodes in the network are equipped with wavelength converters. We therefore focus on the problem of whether an available wavelength exists on a link. Essentially, the network is regarded as a circuit-switched network. Second, the cost for placing a call refers to the aggregate link costs, as defined in Sec. 3.2.1. Third, in the random call model, we assume full knowledge of the network resource status in our search for primary and protection paths. Fourth, the acceptance of a call request is completed only after the system reserves the available network resources for both primary and protection paths. Otherwise, we regard the incoming request (a single request and a batch arrival) as being blocked. Fifth, we heuristically select the nodes of the highest degrees as the segmentation nodes in the simulation for span protection. Sixth, we assume that the arrival of call requests forms Poisson process and that calls have an exponentially distributed service time. The traffic load refers to the product of the arrival rate and the average service time. Finally, we assume uniform traffic, in which an arrival will choose one out of all possible source and destination pairs with equal probability.

4.1 Batch arrivals path assignment

The path assignment problem for a batch arrival is resolved by solving the MILP formulation presented in Chap. 3. Before solving this MILP problem, we preselect a set of possible primary paths to each requests and then choose a set of possible backup paths accordingly. Therefore, for each call request, we have a set of possible primary paths and corresponding backup paths, which is regarded as the input data for the formulation. We heuristically pick the paths of minimum hops as the possible primary paths. And for each selected primary path, obeying the associated protection schemes, we pick the corresponding backup paths which are of hops as minimum as possible. This way of selection not only eliminates the unrealistic long paths but also preserves our goal to boost the resource utilization.

We implement our formulation on the NSFNET (shown in Fig. 4-1) network and select three primary paths for each connection request. In the path protection, we select three backup paths for each primary path accordingly; in PPP, we select three backup path for each link along a primary path. Thus, we have $M = 3$ and $N = 3$ in the formulations. Though by preselecting the primary and backup paths allows us to practically allocate a batch arrival, the scalability problem still exists, and thus the solvable number of requests within a batch arrival is limited by the hardware. In the path protection, our formulation can solve the path assignment problem for a bath arrival of around 40 call requests in 6 hours and a batch of 20 requests within five minutes; in PPP, our formulation is able to deal with a batch arrival consist of 20 requests or so in 6 hours, and 10 calls in 5 minutes. We offer an example of a batch of 10 calls in the Table 4.1. The time consumed to solve the MILP is within five minutes, which is practical to place a batch arrival dynamically. Table 4.1 shows the number of wavelengths assigned to serve a certain batch of some 10 call requests. In addition to the result of considering 10 calls jointly in path assignment, we also show the results of the call-by-call model.

The result of allocating a batch arrival by our formulation highly depends on the selection of primary and backup paths. Further, the selection for backup paths is

Protection Scheme	Batch Arrival	Call-by-Call
Path Protection	44	49
Partial Path Protection	44	47

Table 4.1: Number of wavelengths consumed to serve a batch of 10 calls in the NSFNET network

more crucial than that for the primary paths because the protection sharing is the key of higher resource utilization. Therefore, we have the same primary selections for path protection and PPP, because the resource for primary transmission cannot be shared and it is fair to compare the two different schemes. Then we first solve the path assignment problem formulated for path protection, and obtain the optimal selection of primary-backup paths. Next, we put the optimal pair of primary and backup paths for path protection as a selected path for PPP, not only because it is fair to compare the two schemes but because the optimal pair for path protection should be a good selection for PPP.

We simulate many batch arrivals of different sizes. Most results have the following properties. First, for a batch arrival, both formulations for path protection and for PPP reach the same number of wavelengths which are used to serve a batch arrival. This property probably comes from the nature of the path assignment problem is to minimize the network resource to serve a certain number of connections. We will revisit this property in the random call model analysis. Second, the optimal selection of primary and backup paths in path protection is one of the optimal primary-backup pairs in PPP. This property is indeed a direct result of the previous property and also reflects that path protection is a valid PPP. Third, the result of batch arrival model is always no worse than that of the random call arrival model. This property comes from the fact that optimality of the call-by-call model is only a subset of that of the batch arrival model. Forth, as the number of requests within a batch arrival becomes large, the gain in network resource utilization from the random call model also becomes large, but not obvious. In the 10 call bath example shown in Table 4.1, the resource we gain by the batch call model is 3 to 5 wavelengths, and in most cases of

20 call bath, what we can gain is about 7 to 10 wavelengths. However, as the size of a batch arrival grows, the time needed to solve the formulated MILP problem increases greatly, partly because the number of constraints increases considerably. Since the time consumed to solve a batch arrival path assignment problem is relatively long and the gain in the efficiency is not relatively large, we focus on the random call model in the next section.

4.2 Random arrival path assignment

As discussed in Chapter. 3, in the random call arrivals model, the system serves only one call at a time. To investigate the protection schemes, we simulate path protection span protection, and PPP schemes implemented using both the greedy approach and the MH approach. We assume that the networks and the call requests have the following characteristics. First, all nodes in the network are equipped with wavelength converters. We therefore focus on the problem of whether an available wavelength exists on a link. Essentially, the network is regarded as a circuit-switched network. Second, in the simulation, the cost for placing a call refers to the aggregate link costs, as defined in Sec. 3.2.1. Third, we assume full knowledge of the network resource status in our search for primary and protection paths. Fourth, the acceptance of a call request is completed only after the system reserves the available network resources for both primary and protection paths. Otherwise, we regard the incoming request as being blocked. Fifth, we assume that the arrival of call requests forms Poisson process and that calls have an exponentially distributed service time. The traffic load refers to the product of the arrival rate and the average service time. Finally, we assume uniform traffic, in which an arrival will choose one out of all possible source and destination pairs with equal probability.

In our simulations, we consider three nation-wide US networks, NSFNET (shown in Fig. 4-1), Lata'X' (shown in Fig. 4-4, and Sprint's OC-48 network (shown in Fig. 4-2), and a regional network, the New Jersey LATA network (NJ LATA, shown in Fig. 4-3). Additionally, each link in the networks contains 16 bi-directional wavelengths.

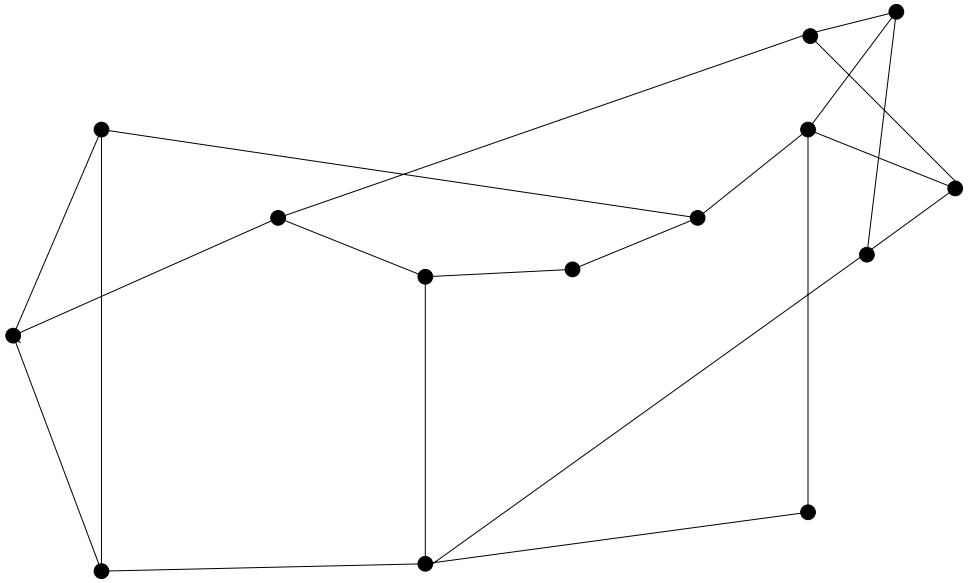

Figure 4-1: The NSFNET

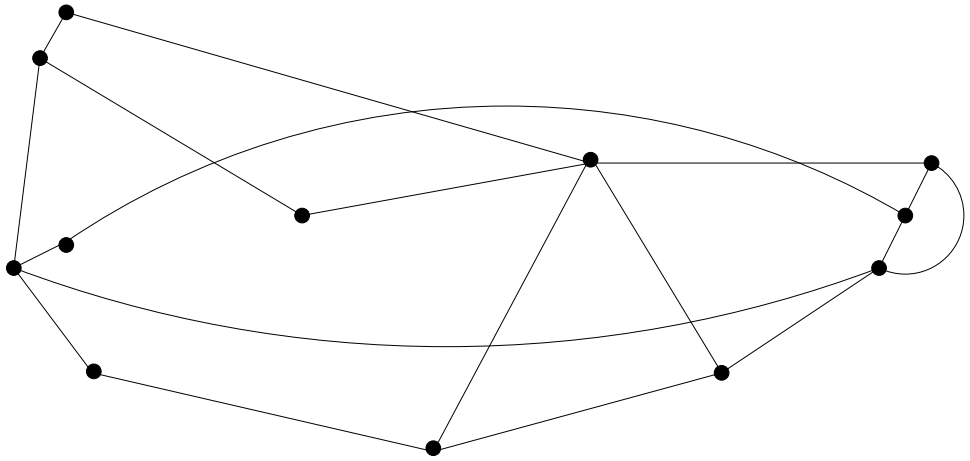

Figure 4-2: The Sprint OC-48 Network

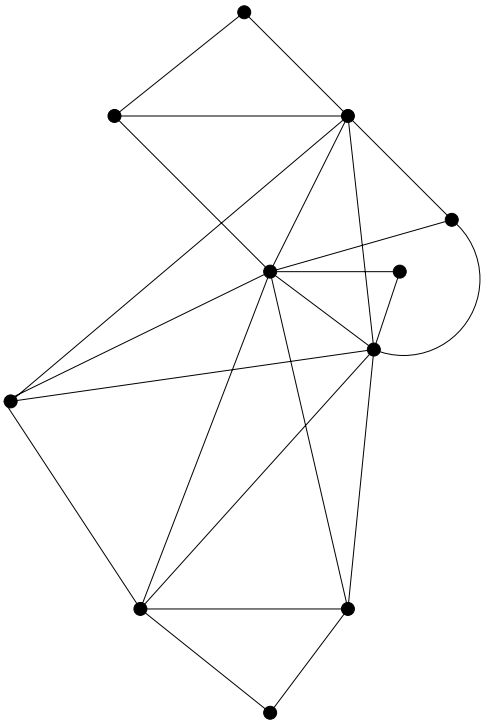

Figure 4-3: The New Jersey LATA Network

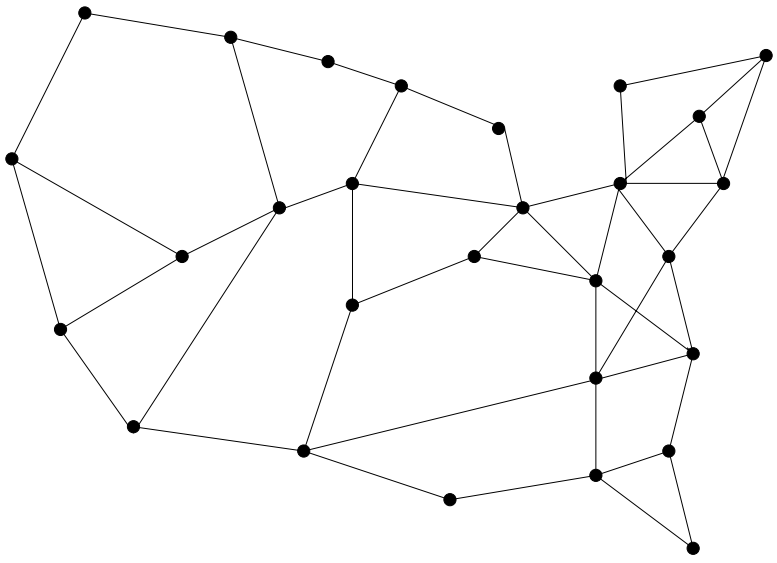

Figure 4-4: The LATA 'X' Network

Note that the nodes in both national networks usually have a lower degree than those in the NJ LATA network, i.e., the regional network is denser.

Two measurements are investigated in the simulations to evaluate the performances of the protection schemes. The first measurement is the steady state blocking probability. Blocking probability is related to opportunity cost, referring to the additional revenue available if certain customers were not turned away. The second measurement is the aggregate number of occupied wavelengths on each link to support connections in the network. This measurement reflects the network resource utilization. For simplicity, we denote PPP implemented by the greedy approach as Greedy-PPP and that implemented by the MH approach as PPP, respectively. We also denote span protection using the MH approach as SP. At last, we denote path protection using the greedy and the MH approaches as Greedy-PP and PP, respectively.

Fig. 4-5 to Fig. 4-14 present our simulation results and Table 4.2 summarizes the results. The results show that, with the same implementation approach, PPP is better than path protection. The results also state that the more segmentation nodes assigned in span protection, the better network resource utilization the span protection can improve from path protection. We thus demonstrate that using the local information does improve the network resource utilization. Still, for each of the protection schemes, the MH approach is better than the greedy approach as the calls accumulate. Our two major conclusions from our simulations are that, as shown in Table 4.2, the PPP scheme implemented using the SP approach has the best performance, that exploiting the local informations concerning link failure does enhance the network resource utilization, and that the other combinations, Greedy-PP, Greedy-PPP and PP in MH approach, perform worst. We discuss these conclusions below.

	Path protection	Span Protection	PPP
Greedy approach	close to PP	NA	close to PP
MH approach	Worst	Better from path protection	Best

Table 4.2: Summary of simulation results for the random call model

Our simulation result presented in Fig. 4-5 and 4-6 shows that the Greedy-PP performs slightly better than ,or even very close to, PP implemented by the MH approach but is still worse than PPP in the MH approach. This result is counter intuitive because the Greedy approach guarantees to serve one call request with minimum network resource while the MH approach does not. However, the Greedy approach neglects the dynamics among the resource allocation as call accumulates, so the Greedy scheme fails to promote the protection sharing, the key to improve utilization. Example 3 illustrates why MH-PP, Greedy-PP and Greedy-PPP perform almost the same. Owing to the nature of the greedy algorithm, the Greedy-PPP approach attempts to occupy the minimum number of wavelengths to serve a call. To this end, Greedy-PPP will find the smallest possible number of wavelengths to protect the corresponding primary path. As a result, one single protection path for a primary path occurs in most cases in the simulation, even though the partial path protection scheme does not require all the protection paths to be the same. Hence the Greedy-PPP has an extremely similar performance to MH-PP and Greedy-PP, which are restricted to assign one single protection path per primary path. Note that owing to the scale of the ILP for path protection, we provide the Greedy-PP result only for NJ LATA.

Example 3 *Consider Fig. 4-15 and a source destination pair* $(1, 4)$*. We have the resource allocation shown in Table 4.3 for MH-PP, MH-PPP, and Greedy-PPP. The table shows that the primary and protection paths for MH-PP are identical to those for Greedy-PPP. This is because Greedy-PPP attempts to fulfill the protection requirement with the minimum number of wavelengths. Note that MH-PPP has the worst performance in terms of network resource utilization in this case. This fact agrees with our simulation results showing that SP-PPP does not perform very well when the network is very lightly loaded. However, as calls accumulate, protection sharing becomes more important for resource utilization and thus MH-PPP is more efficient.*

Since PPP and span protection implemented by MH approach, denoted by MH-PPP and MH-SP respectively, are intrinsically more flexible than MH-PP in both

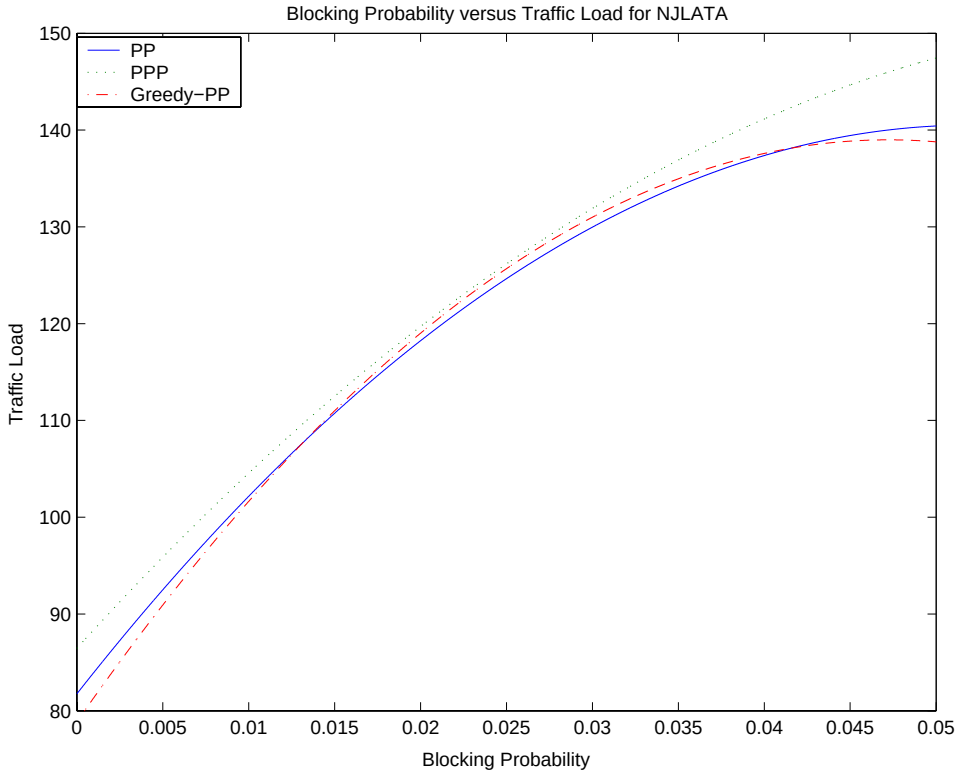

Figure 4-5: Traffic Load vs. Blocking Probability in New Jersey Lata Network

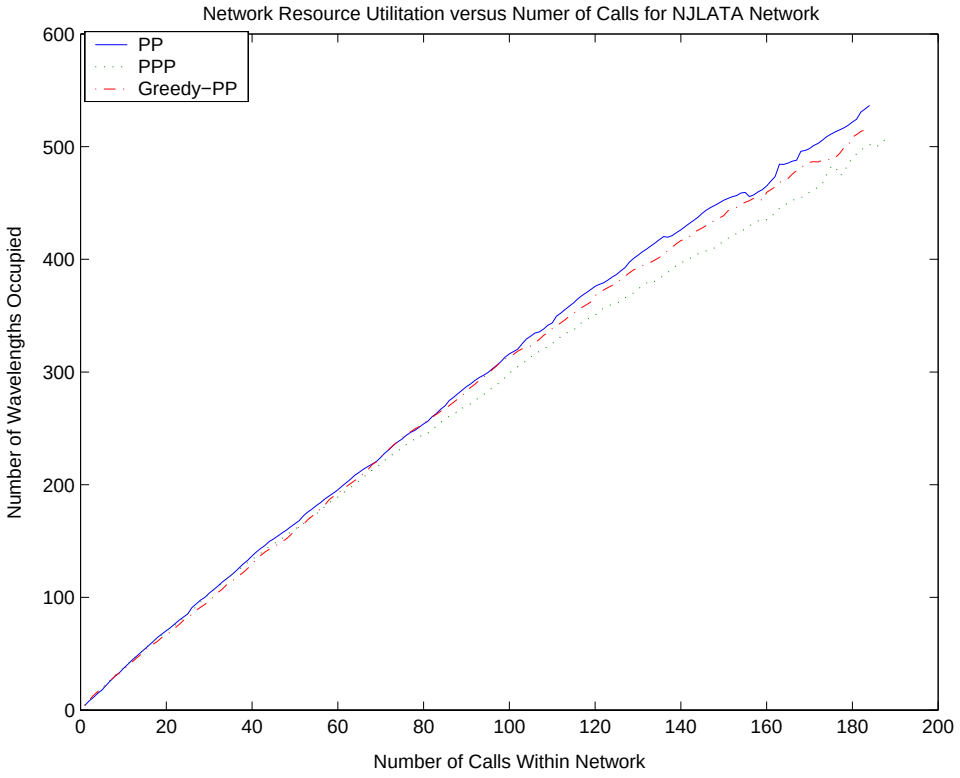

Figure 4-6: Network Resource Utilization in New Jersey Lata Network

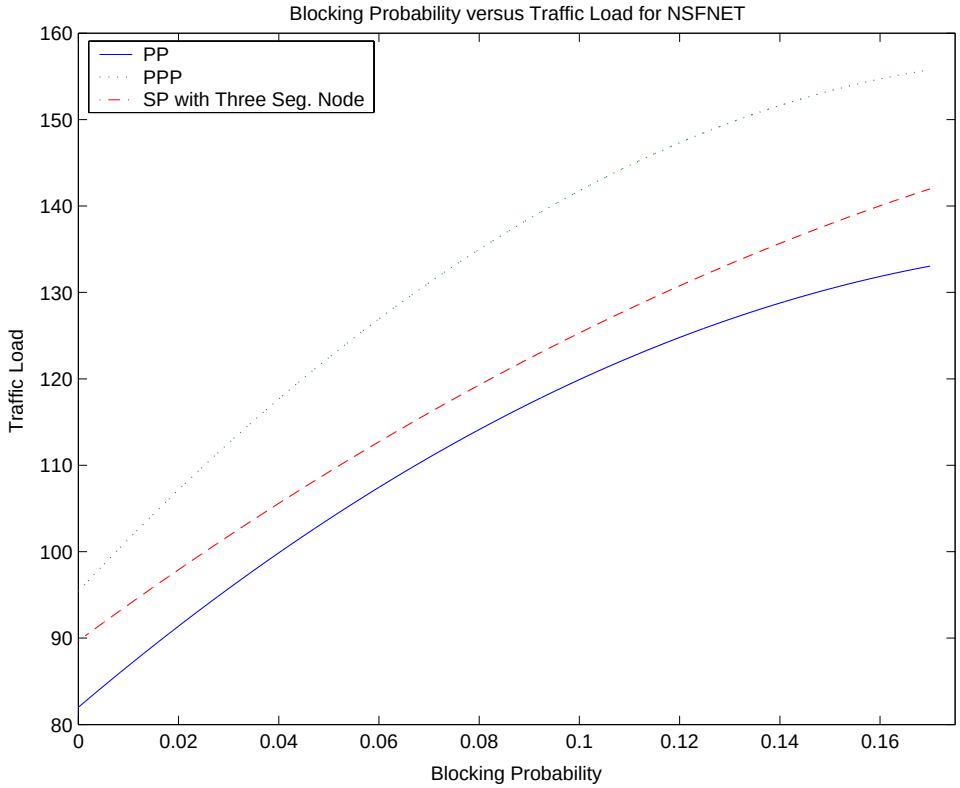

Figure 4-7: Traffic Load vs. Blocking Probability in NSFNET

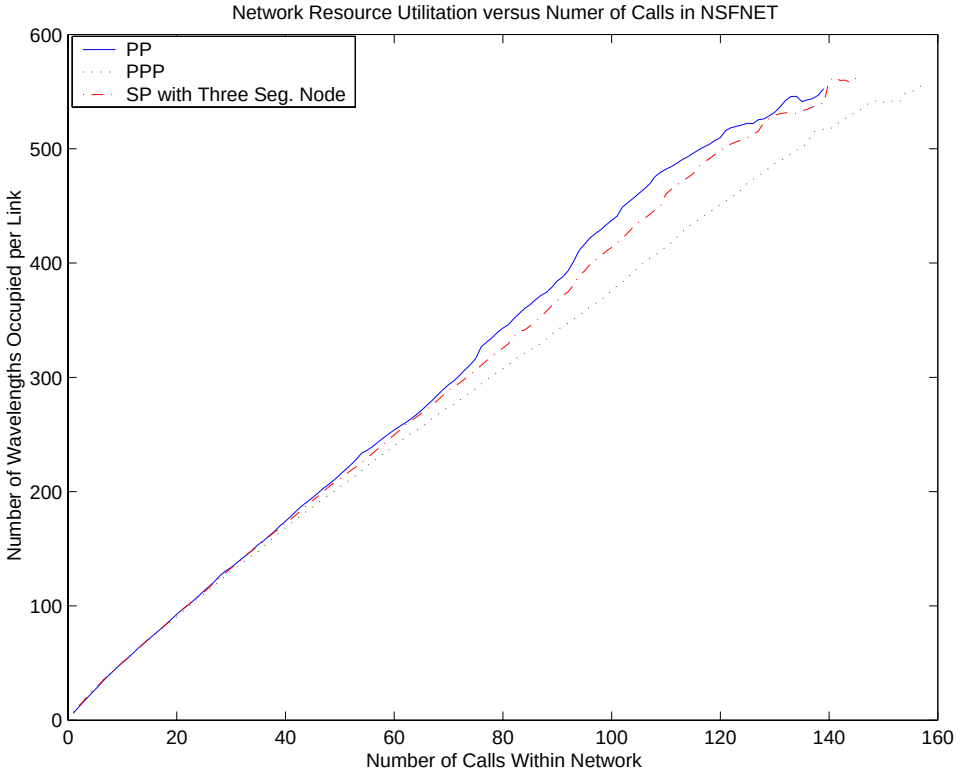

Figure 4-8: Network Resource Utilization in NSFNET

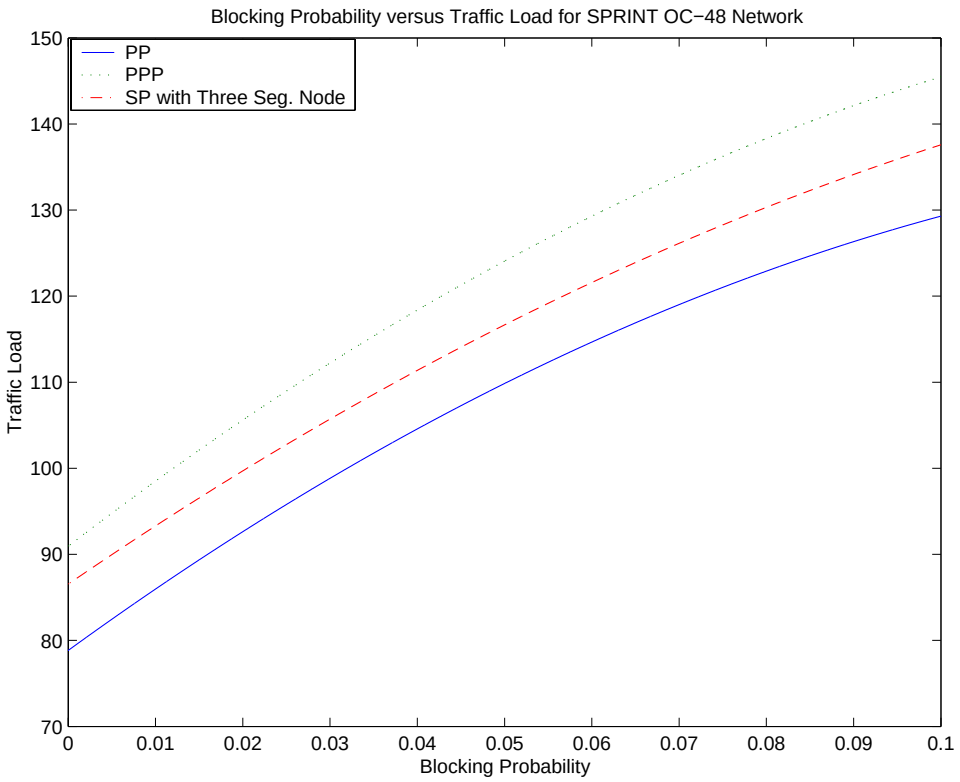

Figure 4-9: Traffic Load vs. Blocking Probability in SPRINT OC-48 Network

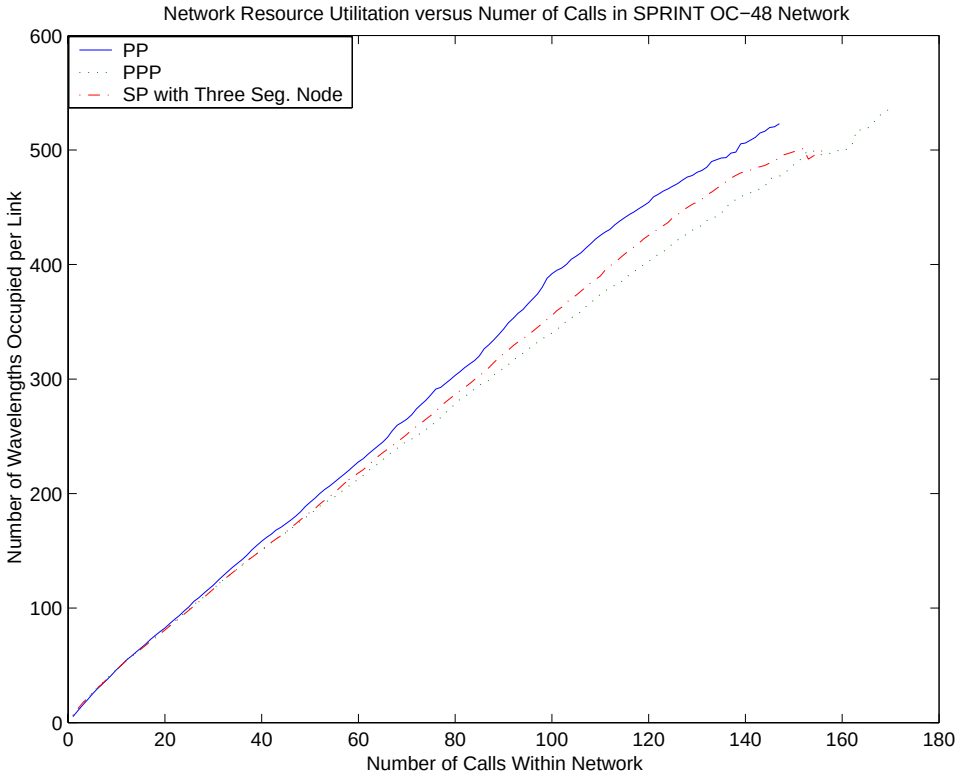

Figure 4-10: Network Resource Utilization in SPRINT OC-48 Network

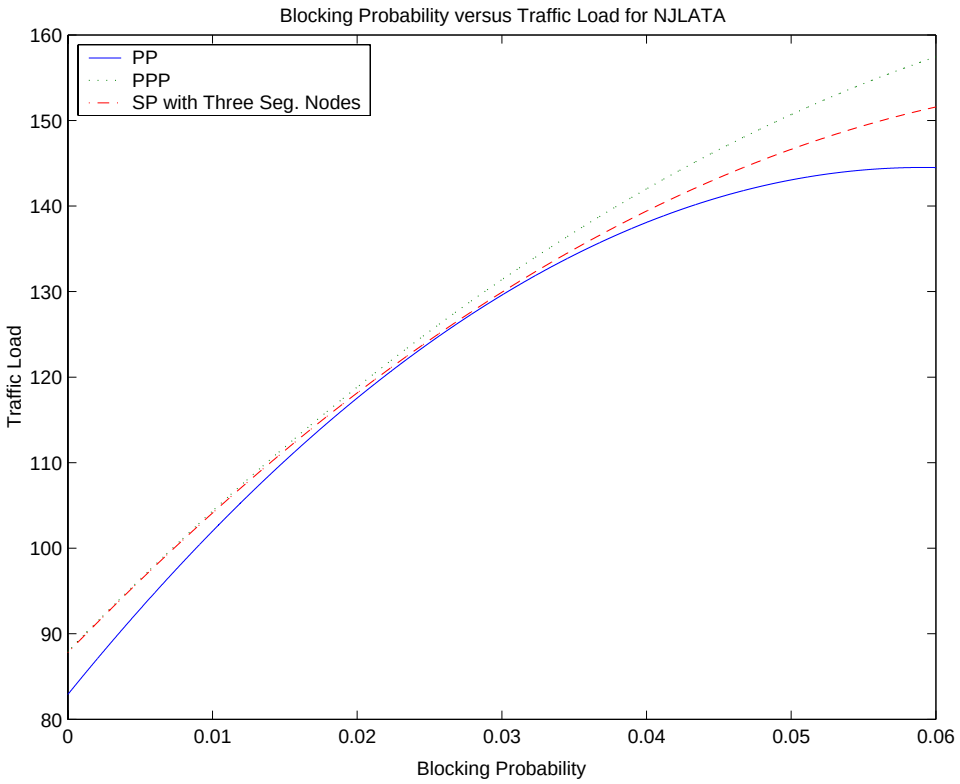

Figure 4-11: Traffic Load vs. Blocking Probability in New Jersey Lata Network

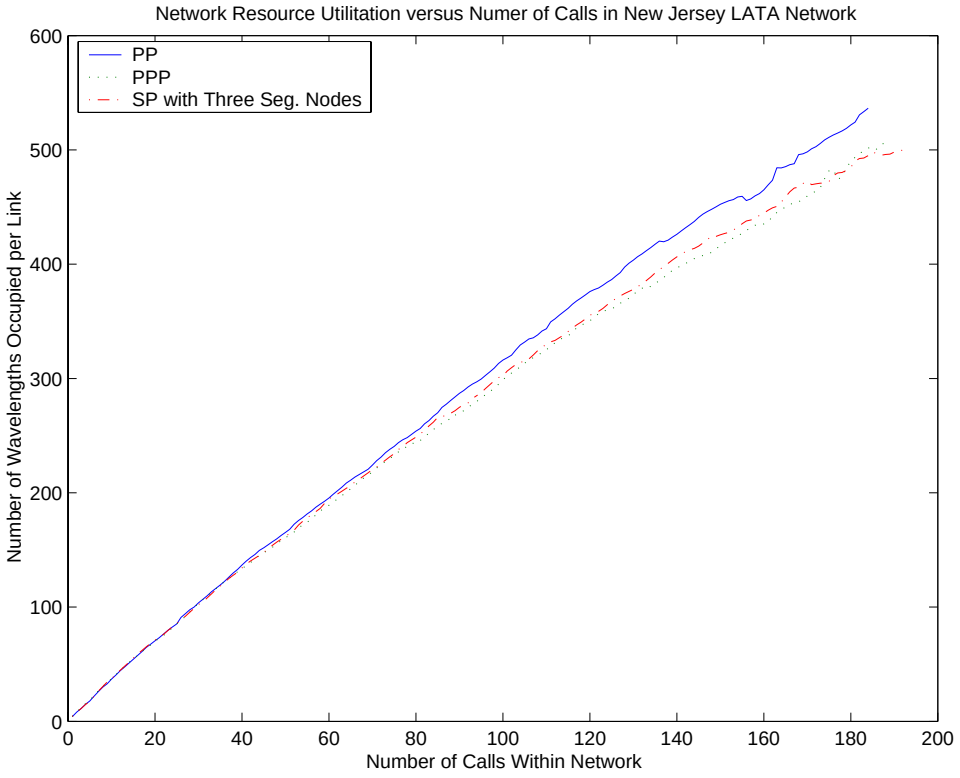

Figure 4-12: Network Resource Utilization in New Jersey Lata Network

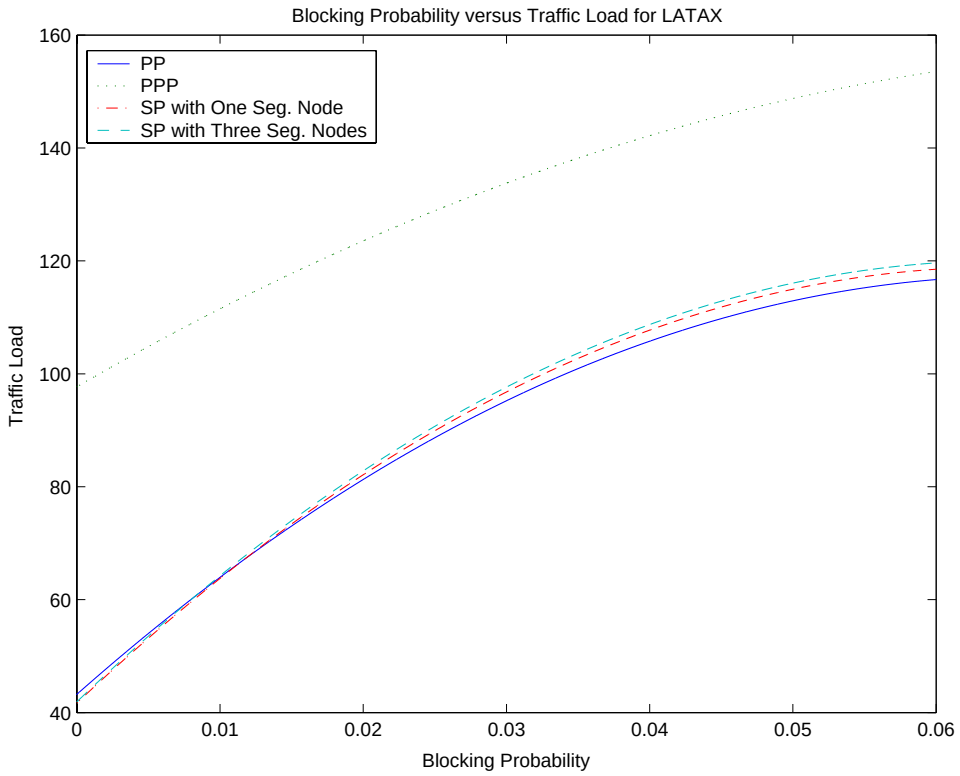

Figure 4-13: Traffic Load vs. Blocking Probability in Lata'X' Network

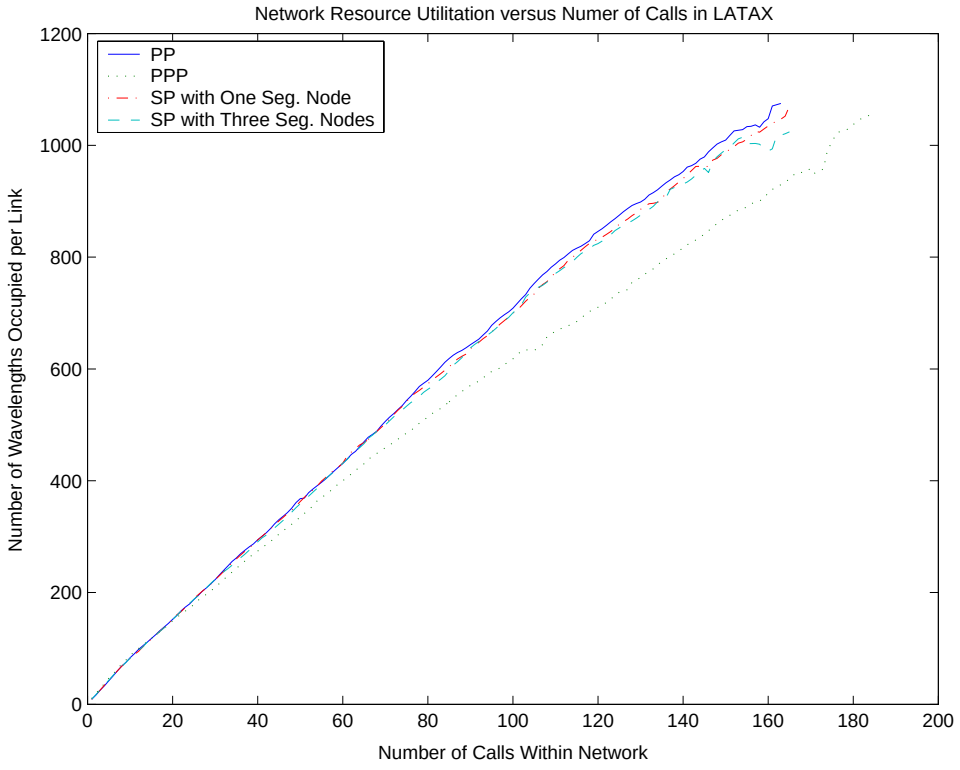

Figure 4-14: Network Resource Utilization in Lata'X' Network

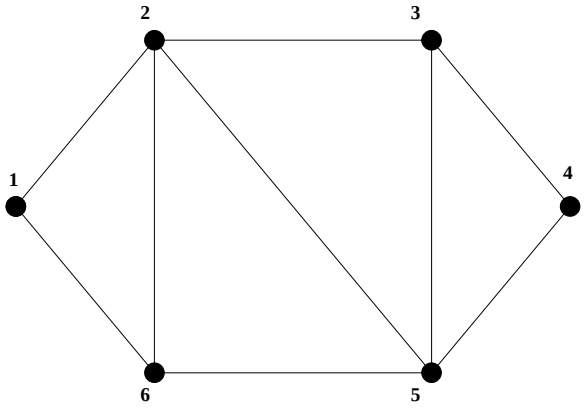

Figure 4-15: An example network

	Primary Path	Protection Path (protected link)	Number of Occupied λ's
MH-PP	1-2-3-4	1-6-5-4	6
MH-PPP	1-2-3-4	1-6-2-3-4 (1-2) 1-2-5-4 (2-3) 1-2-5-4 (3-4)	7
Greedy-PPP	1-2-3-4	1-6-5-4 (1-2) 1-6-5-4 (2-3) 1-6-5-4 (3-4)	6

Table 4.3: Resource allocation for source destination pair $(1, 4)$ of network in Fig. 4-15

the protection scheme and the implementation approach themselves, MH-PPP has the lowest blocking probabilities among all simulation networks and MH-SP is always the intermediate between the worst PP and the best PPP, shown in Fig. 4-7, 4-9, 4-11, and 4-13. These simulation results are consistent with our intuition that path protection and PPP are two special cases of span protection; the former is the case of no segmentation node and the latter is the case that all nodes within the network are segmentation nodes, as discussed in Sec. 2.3. This result points out that protection schemes exploiting the local informations which identify the failed link surely achieve better performance. Moreover, the more segmentation nodes assigned in a network, i.e. more intelligent hardware equipment, the higher resource utilization the system employing such protection schemes achieves.

Another important observation is that the performance of the protection schemes

are highly related to the network topology. For the highly connected regional network, i.e. the NJ LATA network, the blocking events are relatively rare. As shown in Fig. 4-11, when the blocking probability is set as 0.02, the achievable traffic load for NJ Lata network is far above 110. Conversely, from Fig. 4-7 and 4-9, the achievable traffic loads for the two relatively sparse national-wide networks are both of 100 or so, as the blocking probability is fixed at 0.02 as well. The main reason for this phenomenon is that there exist many more choices in the regional network to serve a call request than in the nation-wide networks. Hence, a blocking event is relatively rare in the highly connected network.

Comparing together for the two nation-wide networks of the same size, NSF and SPRINT OC-48 network, we first note that NSFNET has a lower degree than Sprint OC-48 network and has a better performance than Sprint OC-48 network. From Fig. 4-7 and Fig. 4-9, we observe that, when employing the flexible PPP, NSFNET achieves higher traffic load than does the Sprint network for a fixed blocking probability, whereas NSFNET has an inferior performance to the other network when employing the path protection. This indicates that, as the network becomes sparser, the benefits of MH-PPP over MH-PP and Greedy-PPP may increase. One conjecture to buttress this observation is that, since NSFNET is relatively sparse, there are more occurrences of long primary paths, which enhance the usefulness of capacity sharing and thus make the improvements due to SP-PPP noticeable.

Chapter 5

Policy Analysis

Providing reliable transmission against a network transmission failure is a significant quality of service (QoS) requirement in today's high-speed networks. In this chapter, we will address the importance of the network vulnerability from the technological perspective, analyze the influence on the economic market structure, and investigate the policy as well as the legislative requirements in terms of interconnections regarding the network reliability issues. The main concern in the technical section is the reason and the need to require equipment vendors and telecommunication operators to maintain the network robustness as a standard within the QoS. The economic analysis section includes the provision of network reliability with the competitive market environment. In the legislative section, we will discuss the network reliability issue from viewpoint of the Telecommunication Act of 1996. The cost and tariff structure section make a policy recommendation to break the bundle of the software and the hardware. And, finally, we propose a pricing scheme to overcome the problems of exchanging interconnection information among networks.

5.1 Technical Requirement

The feature to divide a tremendous amount of bandwidth in a fiber into a number of wavelengths enables a WDM optical network to transmit a huge volume of data simultaneously. A network failure, such as a link failure, would therefore lead to a

severe disruption in the traffic because of the high transmission rate. For example, consider a SONET of speed OC-96 (10 Gbps), and package size of 53-byte long. Suppose a link fails and the system takes 60 milliseconds to retrieve the traffic, then it will queue 1.43×10^6 53-byte packages in related nodes, apply high efficient elements to process package switching in the order of nanosecond, or suffer a huge amount of data loss. The first possibility would incur a system failure because of severe node congestion. The nodes on the affected paths cannot process any other request since the queues within the node is all occupied. When this negative effect propagates, the whole system would fail. The second probability is still under intensive research in the laboratory, and would incur high equipment cost in the near future. The third possibility would cause an operator to lose its customers, and hurts its profitability. These same incidents that hinder the restoration of the traffic in a SONET network is more damaging to a WDM network, simply owning to WDM networks' broader bandwidth. Thus, the network designers regard the network vulnerability as a crucial criterion for the quality of service (QoS). Therefore, in order to fulfill the QoS, the network management could adopt different schemes. To achieve interoperability and guarantee the QoS, the industry and government establish forums and institutions to set standards.

Conventionally, a SONET network is designed in a ring topology, though the ITU Recommendation does not require so. The ring topology enables the SONET network to perform the associated protection scheme, namely the *self healing ring* (SHR) or the *automatic protection switching* (APS). Whenever a node detects a network failure, the node automatically redirects the traffic into the reserved backup fiber in the opposite direction. This protection function is easy and performs well. Therefore, many networks are constructed on the assumption that network reliability is provided already. As a result, the network protection becomes a crucial part in current optical transmission standards. [**?**] provides an example for this argument.

To fulfill the standards and QoS, many researchers aim at developing protection schemes and protocols with wide variety to implement the network reliability requirement. Among all the protection schemes, the path and link protection schemes are

the ones commonly adopted. Constrained by the provision of the network traffic reliability as well as the restoration time limit required by standards, the WDM networks designed today adopt a static prescheduled strategy, instead of a dynamical network resource rearrangement.

According to the standards commonly accepted worldwide, the official bureau holds authorization of equipment test. Generally speaking, only the equipment passes the test could enter the market.

Driven by the practical need to prevent data loss and the commercial incentive to obtain the equipment authorization, the network protection function is now embedded into the standards for high speed networks, especially for the WDM networks. Thus, there is the reason for us to investigate the approaches that improve the network resource utilization in provision of network protection.

5.2 Economic Influence

In this section, we try to investigate the influence of the WDM network's provision of network reliability within the market structure. Considering a simplified model for n identical companies in the market, the revenue of a company is described by the following function:

$$R = \pi(n, s) - OC(e, c) - FC, \tag{5.1}$$

where R denotes the net revenue, OC denotes the operational cost, and FC denotes the fixed cost. Moreover, let variable $s = eW$, in which $0 \leq e \leq 1$ expresses the network utilization corresponding to different protection schemes, and W expresses the total available resources, e.g. the total number of wavelengths within the network counted on a link basis. The assumption of e will be discussed further. At last, the variable c is the complexity of the protection scheme considered.

This simplified revenue function is assumed to have the following properties.

1. The revenue would decrease as the number of companies increases, i.e., $\frac{\partial \pi}{\partial n} < 0$.

2. The revenue grows as the market size increases, i.e., $\frac{\partial \pi}{\partial s} > 0$.

3. If the network does not adopt any protection scheme, $e \cong 1$. However, if the network protection is provided, e would decrease to $0.5 \leq e \leq 0.75$. (see Chap. 4)

4. Operational cost is dominated by e and c.

   - If $e \cong 1$, the expected OC would be high because of the potential penalties of failing to achieve the QoS in the networks, and the low customer satisfaction.
   - If $0.5 \leq e \leq 0.75$, the OC is relatively low, due to lower penalties and higher customer loyalty.
   - If $c \cong 0$, no protection is offered. In this case, the OC is low because the algorithm of accepting a call is simple, introducing a low network management cost.
   - If $c > 0$, the QoS is guaranteed. Intuitively, we have $\frac{\partial OC}{\partial c} > 0$.

5. c usually increases with e, implies that the protection is offered.

6. The fixed cost FC refers to the cost of constructing an infrastructure network with the fixed network bandwidth W, including equipment cost, maintenance charges, license fee, interconnection fee, and so on. For an incumbent, the FC is relatively lower than that of the new entries, in part because it is exempt from the license fees of right-of-road.

In addition to the properties described above, one can find how the network reliability interacts with the following issues.

Regulation Many regulators limit the number of licenses issued to the operators in this industry. In most cases, a regulator would charge a license fee from the licensees. The higher the license fee is, the higher the entry barrier becomes, resulting a high FC. Moreover, a high FC puts the new investors at the risk of bankruptcy. The new entries confronting bankrupt are forced to sell the companies, especially when the economy/demand faces a down turn. The

incumbent companies with large amount of assets are encouraged to purchase other licenses. Therefore, in this case, not only n decreases but W increases. Note that both effects estimate the rate of return, resulting increased revenue.

Protection Scheme The decision whether to adopt the network protection mechanism would influence the market structure in the following ways. If a company decides not to apply any protection scheme ($e = 1, c = 0$,and fixed W presumptuously), then

1. The profit π would increase as the ability to maintain a larger market scale s grows.
2. The operational cost OC would increase because the occurrence of severe service disruption increases, causing the large penalties. In addition, the customer loyalty/satisfaction would shrink for the same reason, causing the loss in revenue.
3. The operational cost would instead remain low because the complexity of the admission algorithm becomes simple. However, the marginal effect would weaken in the long run;
4. The fixed cost FC would increase because the interconnection rate negotiated with other parties would be high. One of the reasons would be the company without protection mechanism does not meet its liability to maintain the robustness of the entire network. In other words, this company relies on networks belong to its competitors to prevent itself from network failure, which put the company in an unfavorable position during an interconnection negotiation. (See Sec. 5.3)

Technology From the revenue function(5.1), one can observe that the network resource utilization e for a network satisfying the protection QoS plays an important role in determining the net revenue. As long as e is high, one can find that (1) the profit π increases because the available resource increases, (2) the OC is reduced because the QoS is achieved and the additional cost caused by

higher complexity diminishes marginally, and (3) the FC would also increase from the equipment of higher quality performance. To sum up, the installation of a protection scheme to improve the network resource utilization would bring high net revenue in the long run.

From the second arguments above, the companies that do not adopt protection scheme could only survive in the market with very strong demand or when lack of competition, such as monopoly. Only in those markets can the companies without adopting protection schemes have their rate of return surpass the loss caused by service disruptions. This argument explains why the network fault management becomes so crucial in WDM networks. The technology argument provides the strong motivation for current research on protection schemes.

5.2.1 Company Scale

When one or two companies becomes large enough to dominate this industry, these companies become capable of offering a better quality of services at a lower price, partly because of the decreasing marginal cost and networking effects. On the other-hand, those big companies often have more bargaining power in negotiation with other smaller companies. It is not only because the large companies are the potential price maker in the market, but also because those small companies face weaker substitution in choosing backbones to get connected to the other networks. This phenomenon could partially explain the natural monopoly in this industry. Moreover, from the technical analysis in Chap. 4, the network resource utilization becomes efficient when the number of connections within a network grows. This result intensifies the scale effect and benefits the companies of large scale, which is consistent with the market structure today.

5.2.2 Service Class

In the sections above, we assume that all the connections require the same quality of service. However, this is not always true in practice. For example, to expand market

size ($s = eW$), the telecommunication operators usually transmit data asking lower QoS through backup fiber in $1 + 1$ protection scheme. For these data of lower QoS, once a link failure occurs, they will be dropped and replaced by the information being protected. This example points out two issues. The service type can be classified, and, according to the classification, different service types induce different costs/prices to operators/customers.

Basically, the service in a network can be roughly classified into three types: voice, video, and data transfer. The voice service has the highest quality requirement and thereby must be protected. Since the voice communication needs almost immediate response, it cannot be disrupted during the service period. Moreover, the charge for such a service is usually high. The video service requires the second highest service quality. Though the bandwidth needed for such a service is much broader than the voice service, yet such a one-directional communication with time limits could store partial information at the user end in advance. Once a link fails, such a service could thereby tolerate a longer restoration time. However, the large amount of data inflow should be controlled accordingly. Thus, for video broadcasting service, the price for connection would be a little bit cheaper on the bit basis. The data transfer service requires the lowest QoS, simply because the time limit imposed on file transmission is relatively loose. It would not be issues if a file is downloaded for a little longer, say three seconds. Sometimes, the option of reload is even allowable. This means that such a service may not need protection. However, since the local access market is intensively competitive today, the local carriers would choose to lease a broader bandwidth and sign a leased-line contract of the QoS having the protection mechanism involved with the long-haul carriers.

With the relation between the service classification and QoS in mind, one could naturally ask a question: who is the one to pay for the additional costs incurred by the protection requirement? From the economic analysis in this section and the observation of current market, one reasonable conjecture would be the common carriers and the end users, including the residential users and the local carriers, share the additional cost. For common carriers, they undertake the opportunity cost for reserving

the backup resource, the operation cost for network management, and the fixed cost on equipment purchase and interconnection charges. Nevertheless, the factors of scale and time would drive these costs marginal. For end users, they pay a flat rate for all services, i.e., the charge does not come with the QoS or the transmission volume. Therefore the marginal effect on the additional cost incurred by the provision of protection does not appear on the end user side. Moreover, whether the protection QoS is offered is not transparent to most residential users. Thus, in the long run, the end users could absorb a large portion of the additional cost in question.

5.3 Legislative Regulations

In this section, we discuss the incentives for a telecommunication operator to adopt protection schemes on its optical network. From the Section 251, 256, and 259 of the Telecommunication Act of 1996, regulations requires the incumbent local exchange carriers to be made available to any *qualifying* carrier infrastructure, telecommunication facilities, and functions as may be required by such qualifying carriers for the purpose of providing service. Due to the fact that the recovery function is regarded as a fundamental QoS in various standards for high-speed networks, for those new entries, they must prove their network robustness, so as not to jeopardize the service quality of its competitors' network and thereby become a qualifying carrier in definition. For the incumbents, according to the same section in Telecommunication Act of 1996, they must provide a *reasonable* level of service for interconnection. Thus, the provision of network reliability becomes a must. From the above legislative viewpoints, one could easily explain why the infrastructure networks need to implement network protection schemes.

5.4 Cost and Tariff Structures

In this section, we consider the influence of adopting different network management softwares on cost and tariff structures for optical network markets. For simplicity, we

refer the hardware to the network elements such as routers and switches, and refer the software to the network management system. Though various network manegement softwares are installed in different layers and domains within a network, to keep on the subject, we focus on the fault management softwares in this section. Furthermore, we simply consider two softwares. One is the simple fault management software, which basically performs the "$1 + 1$" protection scheme. The other is the complicated fault management system, which implements the protection sharing property. Apparently, the latter requires much more network management information and is more expensive yet more efficient than the former.

We will first consider the cost structure, which refers to the cost for an operator to build up an optical network. In most cases, the costs for hardware and software are bundled. That is, when an operator purchases network elements to construct a network, the prices offered by equipment vendors usually include the software. On the other hand, when an operator upgrades the network management software, it is very possible for the operator to renew the whole network elements to keep the software and hardware compatible. One can observe this phenomenon from the purchase specification published for an network operator. Such a tie of hardware and software brings a considerable profit to vendors. However, the tie is very possible to hamper the competition. In this section, we will discuss the impact of breaking the bundle of hardware and software on the cost structure.

We will next discuss the tariff structure, which is the mechanisms for the optical network operators to charge their customers. The tariff structure today is sensitive to the quality of service (whether requiring protection or not), but is insensitive to which kind of software is applied. This phenomenon mostly comes from the fact that the costs for software and hardware are bundled. We will focus on the incentive for an operator to adopt a complicated network management software in the viewpoint of the trade-off between the hardware and software marginal costs. Moreover, we will also consider the question "who will undertake the cost incurred by the request of network protection." From this question, we can realize the importance and meanings of freely adopting various network management softwares.

5.4.1 Cost Structure

The cost structure for an optical network operator is insensitive to the scheme applied for allocating backup paths, mainly because the hardware and software costs are bundled. Hence, what the operators confront is the bundled (total) cost instead of adopting appropriate protection schemes regarding to the physical networks. In the following, we will discuss the influence of breaking the bundle of costs on the incumbents and the new entries respectively.

Since incumbents own existing networks of large scales, there exist two basic solutions to solve the problem of bandwidth insufficiency. The first is to expand hardware, i.e., laying more fibers and increasing routers' and switches' bandwidth while applying the "$1+1$" protection scheme (simple software). The second approach is to adopt the sophiscated protection sharing scheme (sophisticated fault management software) to increase the network resource utilization, and hence virtually expand the available bandwidth. However, this approach would incur a high cost if the hardware and software costs are bundled, since upgrading software implies purchasing new network elements. As a result, the benefit for incumbents to adopt better protection schemes would weaken in this case. Therefore, only if the bundle is broken will the incumbents carefully consider which protection scheme to apply, and thereby adjust the associated cost structure to reflect hardware and software separately.

For a new entry, the optimal way to construct a network is to apply the software with regard to the properties of hardware. When the scale of a network is small, adopting a complicated protection scheme cannot attain a high network resource utilization, as shown in Chapter 4. Therefore, if the hardware and software costs are bundled together, then choosing a sophiscated fault management software at the beginning stage would increase the entries barrier. On the other hand, choosing a simple software with a low cost would weaken the competition ability of a new entry when its network becomes large. As a result, breaking the bundle of hardware and software allows the new entries to flexibly choose appropriate software, and thereby strengthen its competitive ability.

The time issue will influence both the incumbents and new operators. In practice, it would take an extremely long time to lay fibers and seek/build suitable spaces to locate network elements. However, upgrading the software costs a much shorter time than expanding the hardware. From this point of view, breaking the bundle of the costs would provide a reasonable motivation for operators to adjust the cost structure and thereby promote competition.

To make the unbundling feasible, one major concern is the hardware compatibility. Though the current standards cover the network management field, they are function-oriented rather than interface oriented. For example, the standard only requires the recovery time to be 60 milliseconds without specifying how the software can extract the necessary information from the hardware. Therefore, a vendor can sell equipment which is only compatible with its "proprietary" network management software, and thereby ties the hardware and software. Therefore, to break the bundle, we need to set more rigorous and reasonable standards for the interface of network management software and hardware.

5.4.2 Tariff Structure

The tariff structure in this section is the mechanism for the operators to charge their customers. In addition to the costs for constructing a network, an operator establishes its tariff structure according to the nature of tramission within the network. In the process of the call acceptance, as articulated in previous chapters, a system must assign a primary path to serve the call and, if required, a backup path for protection. Thus, the tariff structure could be designed based on the acceptance procedure.

For the primary paths, one can easily find that the resource reserved for primary transmission cannot be shared and hence one can apply the traditional tariff models for the "circuit switched" networks. The primary argument for this assertion is that, since the reserved wavelength for primary transmission is occupied during the whole transmission period no matter whether the wavelength actually carries the data, such a wavelength usage expels the opportunity for accepting other calls using the same network resource, which is the typical criteria of circuit-switched networks. There-

fore, the tariff for primary transmission is related to the length and the duration of the call considered. Note that this argument applies to both simple protection scheme/software and the complicated protection scheme/software. Hence, an operator could define the fee for the primary transmission as the access fee in the tariff structure in the leased line business.

For the backup paths, the nature of assigning backup resources depends on the protection scheme applied in athe network. If the network applies the simple $1 + 1$ protection scheme, i.e. the simple software, then an operator adopting this protection scheme could apply the traditional tariff structure in calculating the price for backup request. However, if the software applied is complicated, such as the partial protection scheme, then the resource for backup transmission can be shared with other call requests. In this case, the resource usage is close to that of package switching, in which a channel could be shared by multiple connections and the channel is released as soon as all the connections on this cannnel depart the system. Though the nature of sharing backup resources and that of package switch are different in a certain degree, their properties are similar in essence. Therefore, an operator could define the fee for offering the protection services as the fee for quality of service, and it could be evaluated by different models according to the software the operator applies.

We further consider the tradeoff between hardware and software. If an operator adopts a simple software to locate backup paths, then it would incur a high opportunity cost for hardware utilization, though the software cost is relatively low. On the other hand, if an operator applies a sophisticated software, though the opportunity cost for hardware utilization decreases, it is very expensive to apply such a complicated software. Since lacking the specific figures, we cannot process the quantative analysis such as the rate of return and the potential market in this section. However, it is clear that the cost of software is marginally decreasing while the cost incurred from the inefficient hardware utilization remains at a certain level without any margianl effect to occur. Thus our conjecture that an operator would adopt a sophisticated software to strengthen its competitive ability and thereby adjust its tariff structure is reasonable.

Finally, we consider the question of how to encourage an operator to lower the transferred costs incurred by the quality of service to the customers. From the principle of economics, if the substitution is complete, then the ratio of the transferred costs is low, and vice versa. Thus, in order to make the ratio low, we have to promote the competition within the market and maintain the substitution to a degree for customers. From the analysis mentioned above, breaking the bundle of software and hardware is helpful in promoting competition and hence avoid the situation that the operators shed the costs of building efficient networks to customers.

5.5 Interconnection Exchanged Information

From Chapter 4, we know the tendency that the larger the scale of a network is, the more efficient the network resource utilization a network can achieve. This observation motivates us to consider whether the different parties can share the network resources of the interconnected networks for backup resources. To simplify our discussions, we consider only the interconnection information exchange issues in this section.

As indicated in the Chapter **??**, implementing a sophisticated protection scheme which allows protection sharing needs a huge volume of exchanged information, since it needs the complete network resource utilization status. For example, to fulfill the protection sharing, the information must contain the entire protection status. Such a huge volume of information exchange would cause the following problems. First, since the carried information grows, the overhead of a package in the network will be longer and thereby lower the network throughput. Second, even though a network has its own network for transmitting management information, such a traffic volume will increase the traffic delay as well as the response time, especially when the network management system is implemented in a centralized manner. At last, to protect its own commercial secret and secure its network, an operator will not agree to make its infrastructure traffic information transparent to its competitors. However, this consideration would make the capacity sharing infeasible, simply because the protection sharing needs intensive resource allocation information.

To increase the network resource utilization as well as to overcome the difficulties mentioned above, we propose the pricing scheme and the distributed network admission control to achieve our goal. The pricing mechanism could go as follows. Let the source node of the call request broadcast the demand for the backup resource. Once the nodes at the gateways, which connect different networks, receive the broadcasting signal, they send back the requesting node a price signal associated with their network status, instead of the full information packages. Comparing all the prices, including the price offered by the network it belongs to, the source node chooses the resource with the lowest cost as the backup path. Then the tasks of physically allocating and recording the call request are handovered by the nodes among the different networks. By doing so, the amount of exchanged information is decreased and thereby the response delay diminishes. Moreover, the commercial secret is protected in this manner.

Chapter 6

Conclusions and Future Directions

We have introduced a novel protection scheme, the partial path protection (PPP), and investigated two arrival patterns, the batch arrival and the random call-by-call arrival. Moreover, in additional to considering traditional static capacity-efficiency measures for evaluating the efficiency of protection schemes, we considered the measurement of the blocking probability. To avoid the complexity of dynamic optimization we confront in the batch arrival model, we presented two heuristics to implementing path protection and PPP in the random call model. These approaches, which we termed greedy and MH, were compared to each other for both path protection, span protection and PPP. We have demonstrated that PPP and span protection, which exploiting the local information identifying the link failure location, is superior to path protection, which ignores such information. We also show that MH is superior to the greedy approach. As expected from the fact that PPP is more general and flexible than path protection, PPP outperforms path protection in terms of blocking probability. Moreover, the MH approach performs better than the greedy approach. It is the dynamic nature of our problem that renders MH superior to the greedy approach. Indeed, MH emphasizes reducing resource use among primary paths, since their bandwidth cannot be shared. The fact that MH may be less efficient than the greedy approach in its allocation of capacity for protection paths is mitigated by the fact that protection bandwidth can be shared.

The advantages of PPP as well as span protection over path protection have certain

implications in the area of network management. Path protection only requires that the source and destination node be aware that a failure occurred somewhere along the primary path. Localization of the failure is unimportant, since protection takes place in the same way regardless of where the failure occurs. Thus, once the protection path has been set up, the network management does not need to have detailed knowledge of the nature of the failure to effect protection. Path protection can then be handled by higher layer mechanisms. For link protection, local information is needed by the nodes adjacent to the failure, but there is no need to manage protection on a path-by-path basis. Lower layers can therefore ensure link protection. PPP and span protection, on the other hand, require on the part of the network management effecting protection both knowledge of the path and of the location of the failed link. Our results point to the fact that visibility by the network management system across layers may be useful for performing protection efficiently.

The policy analysis also opens a window of understanding the influence of protection schemes on market structure. To obtain the authorization for network equipment test, today's telecommunication operators and vendors regard the provision of network protection as an essential criterion for network performance. However, in the economic analysis, we find that corporations without robust networks cannot compete with other companies having reliable networks. Moreover, the protection scheme would favor the operators who have networks of large scale. This phenomenon intensifies when the applied network protection scheme improves the network utilization. In addition, once the provision of protection schemes is open in the standard, the IP router manufacturers and the optical switch vendors could cooperate to produce an vertically integrated node equipment for a WDM network. The vendors' dominant power in the standard establishment would thus become stronger.

There are several further research directions for our work. One direction is to develop a more efficient algorithm for the batch call arrivals. One possible approach is to formulate this problem in nonlinear programming (NLP) and analyze the geometry in the relating constraint set. Comparing the constraint set of the batch case with those of the dynamic system should yield insight into the effect of the dynamic assumption

upon the effectiveness of protection schemes. Another area of further research is the cost effective analysis of the cost for improving hardware capability to localize the link failure versus the associate reward gained from the improved resource utilization. Such an analysis would allow us to study the trend of future optical network industry.

Appendix A

Formulations

A.1 Problem Formulation for Batch Call Arrivals

A.1.1 Formulation for path protection

Varible Definition

Let

$\mathcal{N}$	denote the number of nodes;
L	denote the set of all possible links;
R	denote the total number of connection requests;
W	denote the total number of wavelengths per link;
C	denote the total number of the *sufficient constraints*;
S_r	denote the source node of call r, $\forall r = 1, \ldots, R$;
D_r	denote the destination node of call r, $\forall r = 1, \ldots, R$;
$M(S_r, D_r)$	denote the total number of selections of primary path for call r, $\forall r = 1, \ldots, R$;
$N(M(S_r, D_r))$	denote the total number of backup path selections for the m^{th} preselected primary path of call r, $\forall r = 1, \ldots, R$.

For simplicity, we set $M(S_r, D_r) = M$ and $N(M(S_r, D_r)) = N$ in the following formulations. We also define

$$\beta_{i,j}^{w} = \begin{cases} 1, & \text{if wavelength } \lambda_w \text{ on link } (i,j) \text{ is reserved for backup restoration,} \\ 0, & \text{otherwise,} \end{cases}$$

$$x_{i,j}^{w,r} = \begin{cases} 1, & \text{if call } r \text{ reserves wavelength } \lambda_w \text{ on link } (i,j) \text{ for primary transmission,} \\ 0, & \text{otherwise,} \end{cases}$$

$$y_{i,j}^{w,r} = \begin{cases} 1, & \text{if call } r \text{ reserves wavelength } \lambda_w \text{ on link } (i,j) \text{ for backup restoration,} \\ 0, & \text{otherwise.} \end{cases}$$

$$\Phi_{m}^{r} = \begin{cases} 1, & \text{if call } r \text{ is served by its } m^{th} \text{ preselected primary path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$\Psi_{m,n}^{r} = \begin{cases} 1, & \text{if call } r \text{ is served by its } m^{th} \text{ primary path and the corresponding } n^{th} \text{ backup path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$P_{ij}^{m,r} = \begin{cases} 1, & \text{if call } r\text{'s } m^{th} \text{ primary path goes through link } (i,j), \\ 0, & \text{otherwise,} \end{cases}$$

$$Q_{i,j}^{m,n,r} = \begin{cases} 1, & \text{if the } n^{th} \text{ backup path of call } r\text{'s } m^{th} \text{ primary path goes through link } (i,j), \\ 0, & \text{otherwise.} \end{cases}$$

Note that, since $\Psi_{m,n}^{r} = 1$ only if $\Phi_{m}^{r} = 1$ and $\Phi_{m}^{r} = 0$ guarantees that $\Psi_{m,n}^{r} = 0$, we replace Φ_{m}^{r} by $\Psi_{m,n}^{r}$ in the formulation. Additionally, for knowing all link pairs needed to take into consideration, we preselect and enumerate these link pairs as $(l_p[c], l_b[c]) = ((l[c], k[c]), (i[c], j[c]))$, for all $c = 1, \ldots, C$, and we thus have total C constraints.

Formulation

The MILP formulation is presented as follows.

$$\text{Minimize} \quad \sum_{(i,j)\in L}\sum_{w=1}^{W}\sum_{r=1}^{R} x_{i,j}^{w,r} + \sum_{(i,j)\in L}\sum_{w=1}^{W} \beta_{i,j}^{w} \tag{A.1}$$

$$\text{Subject to} \quad \sum_{m=1}^{M}\sum_{n=1}^{N} \Psi_{m,n}^{r} = 1, \quad \forall r = 1,\ldots,R \tag{A.2}$$

$$\sum_{w=1}^{W} x_{i,j}^{r,w} = \sum_{m=1}^{M}(P_{i,j}^{m,r} \times (\sum_{n=1}^{N} \Psi_{m,n}^{r})), \quad \forall r = 1,\ldots,R, m = 1,\ldots,M, (i,j) \in L \tag{A.3}$$

$$\sum_{w=1}^{W} y_{i,j}^{w,r} = \sum_{m=1}^{M}\sum_{n=1}^{N}(Q_{i,j}^{m,n,r} \times \Psi_{m,n}^{r}), \quad \forall r = 1,\ldots,R, m = 1,\ldots,M, n = 1,\ldots,N, (i,j) \in L \tag{A.4}$$

$$\sum_{r=1}^{R} x_{i,j}^{w,r} \leq 1, \quad \forall w = 1,\ldots,W, (i,j) \in L \tag{A.5}$$

$$\sum_{w=1}^{W} \beta_{i[c],j[c]}^{w} \geq \sum_{r=1}^{R}\sum_{m=1}^{M}\sum_{n=1}^{N}(\Psi_{m,n}^{r} \times P_{l[c],k[c]}^{m,r} \times Q_{i[c],j[c]}^{m,n,r}), \forall (i,j) \in L \tag{A.6}$$

$$\forall c = 1,\ldots,C, \beta_{i,j}^{w} \geq y_{i,j}^{w,r}, \quad \forall w = 1,\ldots,W, r = 1,\ldots,R, \tag{A.7}$$

$$\sum_{r=1}^{R} x_{i,j}^{w,r} + \beta_{i,j}^{w} \leq 1, \qquad \forall w = 1,\ldots,W, \quad \forall (i,j) \in L \tag{A.8}$$

$$\sum_{w=1}^{W}\sum_{r=1}^{R} x_{i,j}^{w,r} + \sum_{w=1}^{W} \beta_{i,j}^{w} \leq W, \qquad \forall (i,j) \in L \tag{A.9}$$

$$0 \leq x_{i,j}^{w,r}, y_{i,j}^{w,r}, \beta_{i,j}^{w} \leq 1, \forall w = 1,\ldots,W, r = 1,\ldots,R, (i,j) \in L \tag{A.10}$$

$$\Psi_{m,n}^{r} \in \{0,1\}, \quad \forall m = 1,\ldots,M, n = 1,\ldots,N, r = 1,\ldots,R \tag{A.11}$$

A.1.2 Formulation for Partial Path Protection

Varible Definition

Let

$\mathcal{N}$	denote the number of nodes;
L	denote the set of all possible links;
R	denote the total number of connection requests;
W	denote the total number of wavelengths per link;
C	denote the total number of the *sufficient constraints*;
S_r	denote the source node of call r, $\forall r = 1, \ldots, R$;
D_r	denote the destination node of call r, $\forall r = 1, \ldots, R$;
$M(S_r, D_r)$	denote the total number of selections of primary path for call r, $\forall r = 1, \ldots, R$;
$N(M(S_r, D_r))$	denote the total number of backup path selections for the m^{th} preselected primary path of call r, $\forall r = 1, \ldots, R$.
$\mathcal{H}_m^r$	denote the total number of hops of the m^{th} primary path of call r, $\forall m = 1, \ldots, M, r = 1, \ldots, R$,
ℓ_m^r	denote the l^{th} node along the m^{th} primary path of call r, $\forall m = 1, \ldots, M, \quad r = 1, \ldots, R$.

For simplicity, we set $M(S_r, D_r) = M$ and $N(M(S_r, D_r)) = N$ in the following formulations. We also define

$$\beta_{i,j}^w = \begin{cases} 1, & \text{if wavelength } \lambda_w \text{ on link } (i,j) \text{ is reserved for backup restoration,} \\ 0, & \text{otherwise,} \end{cases}$$

$$x_{i,j}^{w,r} = \begin{cases} 1, & \text{if call } r \text{ reserves wavelength } \lambda_w \text{ on link } (i,j) \text{ for primary transmission,} \\ 0, & \text{otherwise,} \end{cases}$$

$$y_{i,j}^{w,r,l} = \begin{cases} 1, & \text{if call } r \text{ reserves wavelength } \lambda_w \text{ on link } (i,j) \text{ to protect its } l^{th} \text{ link on} \\ & \text{the primary path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$\Phi_m^r = \begin{cases} 1, & \text{if call } r \text{ is served by its } m^{th} \text{ preselected primary path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$\Psi_{m,n}^{r,l} = \begin{cases} 1, & \text{if call } r \text{ is served by its } m^{th} \text{ selection of primary path and the corresponding} \\ & n^{th} \text{ selection of backup path which protects its } l^{th} \text{ link along the primary path,} \\ 0, & \text{otherwise,} \end{cases}$$

$$P_{ij}^{m,r} = \begin{cases} 1, & \text{if call } r\text{'s } m^{th} \text{ primary path goes through link } (i,j), \\ 0, & \text{otherwise,} \end{cases}$$

$$Q_{i,j}^{m,n,r,l} = \begin{cases} 1, & \text{if the } n^{th} \text{ backup path for protecting call } r\text{'s } l^{th} \text{ link on the } m^{th} \\ & \text{primary path goes through link } (i,j), \\ 0, & \text{otherwise.} \end{cases}$$

To identify the specific backup path for each link along each primary path, we modify the parameters y, Q and Ψ by introducing the *link* running variable l. Also note that, in the definition of Q, we eliminate the link which is a portion of the primary path by setting its corresponding variable $Q_{i,j}^{m,n,r,l} = 0$. Additionally, we also let

$\mathcal{H}_m^r$ denote the total number of hops of the m^{th} primary path of call r,
$\forall m = 1, \ldots, M, r = 1, \ldots, R,$

ℓ_m^r denote the l^{th} node along the m^{th} primary path of call r,
$\forall m = 1, \ldots, M, \quad r = 1, \ldots, R.$

Such parameters are needed in PPP, since this scheme naturally distinguishes the backup path for each link on each primary path.

Formulation

The MILP formulation is presented as follows.

$$\text{Minimize} \quad \sum_{(i,j)\in L}\sum_{w=1}^{W}\sum_{r=1}^{R} x_{i,j}^{w,r} + \sum_{(i,j)\in L}\sum_{w=1}^{W} \beta_{i,j}^{w} \tag{A.12}$$

$$\text{Subject to} \quad \sum_{m=1}^{M} \Phi_m^r = 1, \quad \forall r = 1,\ldots,R \tag{A.13}$$

$$\sum_{n=1}^{N} \Psi_{m,n}^{r,l} = \Phi_m^r, \quad \forall m = 1,\ldots,M,\ r = 1,\ldots,R,\ l = 1,\ldots,\mathcal{H}_m^r \tag{A.14}$$

$$\sum_{w=1}^{W} x_{i,j}^{r,w} = \sum_{m=1}^{M} (P_{i,j}^{m,r} \times \Phi_m^r), \quad \forall r = 1,\ldots,R, m = 1,\ldots,M, \quad (i,j) \in L \tag{A.15}$$

$$\sum_{w=1}^{W} y_{i,j}^{w,r} = \sum_{m=1}^{M}\sum_{n=1}^{N} Q_{i,j}^{m,n,r,l} \times \Psi_{m,n}^{r,l},$$
$$\forall l = 1,\ldots,\mathcal{H}_m^r, r = 1,\ldots,R, m = 1,\ldots,M, n = 1,\ldots,N, (i,j) \in L \tag{A.16}$$

$$\sum_{r=1}^{R} x_{i,j}^{w,r} \le 1, \quad \forall w = 1,\ldots,W, \quad (i,j) \in L \tag{A.17}$$

$$\hat{y}_{i[c],j[c]}^{r} \ge \Psi_{m,n}^{r,l} \times (P_{h[c],k[c]}^{m,r} + P_{\ell,\ell+1}^{m,r} - 1) \times Q_{i[c],j[c]}^{m,n,r,l}, \quad \forall c = 1,\ldots,C,$$
$$\forall l = 1,\ldots,\mathcal{H}_m^r, m = 1,\ldots,M, n = 1,\ldots,N, r = 1,\ldots,R \tag{A.18}$$

$$\sum_{w=1}^{W} \beta_{i[c],j[c]}^{w} \ge \sum_{r=1}^{R} \hat{y}_{i[c],j[c]}^{r}, \quad \forall c = 1,\ldots,C \tag{A.19}$$

$$\beta_{i,j}^{w} \ge y_{i,j}^{w,r,l}, \quad \forall w = 1,\ldots,W, r = 1,\ldots,R, l = 1,\ldots,\mathcal{H}_m^r, (i,j) \in L \tag{A.20}$$

$$\sum_{r=1}^{R} x_{i,j}^{w,r} + \beta_{i,j}^{w} \le 1, \qquad \forall w = 1,\ldots,W, \quad \forall (i,j) \in L, \tag{A.21}$$

$$\sum_{w=1}^{W}\sum_{r=1}^{R} x_{i,j}^{w,r} + \sum_{w=1}^{W} \beta_{i,j}^{w} \le W, \qquad \forall (i,j) \in L, \tag{A.22}$$

$$0 \le x_{i,j}^{w,r}, y_{i,j}^{w,r,l}, \beta_{i,j}^{w} \le 1,$$
$$\forall (i,j) \in L, w = 1,\ldots,W, r = 1,\ldots,R, l = 1,\ldots,\mathcal{H}_m^r \tag{A.23}$$

$$\Phi_m^r, \Psi_{m,n}^r \in \{0,1\}, \quad \forall m = 1,\ldots,M, \quad n = 1,\ldots,N, \quad r = 1,\ldots,R \tag{A.24}$$

A.2 Problem Formulation for Random Call Arrivals

A.2.1 Formulation for path protection

Varible Definition

Let

L denote the set of all possible links,

S denote the source node,

D denote the destination node,

$$c_{ij} = \begin{cases} 1, & \text{if at least one wavelength is available on link } (i,j) \in L, \\ \infty, & \text{otherwise,} \end{cases}$$

$$d_{ij}^{lk} = \begin{cases} 0, & \text{if at least one wavelength on link } (l,k) \text{ other than } (i,j) \text{ is already} \\ & \text{reserved to protect links other than } (i,j), \\ 1, & \text{else if at least one wavelength is available on link } (l,k) \in L, \\ \infty, & \text{otherwise,} \end{cases}$$

$$x_{ij} = \begin{cases} 1, & \text{if the primary path rests on an available wavelength in link } (i,j), \\ 0, & \text{otherwise,} \end{cases}$$

$$y_{ij} = \begin{cases} 1, & \text{if the system reserves a wavelength in link } (i,j) \text{ for protection,} \\ 0, & \text{otherwise,} \end{cases}$$

$$v_{ij}^{lk} = \begin{cases} 1, & \text{if a wavelength on } (l,k) \text{ is reserved to protect its associated primary path on } (i,j), \\ 0, & \text{otherwise.} \end{cases}$$

Formulation

$$\text{Minimize} \quad \sum_{(i,j)\in L} c_{ij}x_{ij} + \sum_{(i,j)\in L} y_{ij} \tag{A.25}$$

$$\text{Subject to} \quad \sum_{(S,j)\in L} x_{Sj} - \sum_{(j,S)\in L} x_{jS} = 1, \tag{A.26}$$

$$\sum_{(D,j)\in L} x_{Dj} - \sum_{(j,D)\in L} x_{jD} = -1, \tag{A.27}$$

$$\sum_{(i,j)\in L} x_{ij} - \sum_{(j,i)\in L} x_{ji} = 0, \quad \forall i \neq S, D, \tag{A.28}$$

$$\sum_{(S,l)\in L} v_{ij}^{Sl} - \sum_{(l,S)\in L} v_{ij}^{lS} \geq x_{ij}, \quad \forall (i,j) \in L, \tag{A.29}$$

$$\sum_{(l,D)\in L} v_{ij}^{lD} - \sum_{(D,l)\in L} v_{ij}^{Dl} \geq x_{ij}, \quad \forall (i,j) \in L, \tag{A.30}$$

$$\sum_{(l,k)\in L} v_{ij}^{lk} - \sum_{(k,l)\in L} v_{ij}^{kl} = 0, \quad \forall (i,j) \in L, k \neq S, k \neq D, \tag{A.31}$$

$$v_{ij}^{ij} + v_{ji}^{ij} = 0, \quad \forall (i,j) \in L, \tag{A.32}$$

$$y_{lk} \geq d_{ij}^{lk} v_{ij}^{lk}, \quad \forall (i,j),(l,k) \in L, \tag{A.33}$$

$$x_{ij} \geq v_{ij}^{lk}, \quad \forall (i,j),(l,k) \in L, \tag{A.34}$$

$$v_{ij}^{lk} + x_{mn} \leq v_{mn}^{lk} + 1, \quad \forall (i,j),(l,k),(m,n) \in L, \tag{A.35}$$

$$x_{ij}, y_{ij}, v_{ij}^{lk} \in \{0,1\}, \quad \forall (i,j),(l,k) \in L. \tag{A.36}$$

A.2.2 Formulation for partial path protection

Variable Definition

L denote the set of all possible links,

S denote the source node,

D denote the destination node,

$$c_{ij} = \begin{cases} 1, & \text{if at least one wavelength is available on link } (i,j) \in L, \\ \infty, & \text{otherwise,} \end{cases}$$

$$d_{ij}^{lk} = \begin{cases} 0, & \text{if at least one wavelength on link } (l,k) \text{ other than } (i,j) \text{ is already} \\ & \text{reserved to protect links other than } (i,j), \\ 1, & \text{else if at least one wavelength is available on link } (l,k) \in L, \\ \infty, & \text{otherwise,} \end{cases}$$

$$x_{ij} = \begin{cases} 1, & \text{if the primary path rests on an available wavelength in link } (i,j), \\ 0, & \text{otherwise,} \end{cases}$$

$$y_{ij} = \begin{cases} 1, & \text{if the system reserves a wavelength in link } (i,j) \text{ for protection,} \\ 0, & \text{otherwise,} \end{cases}$$

$$v_{ij}^{lk} = \begin{cases} 1, & \text{if a wavelength on } (l,k) \text{ is reserved to protect its associated primary path on } (i,j), \\ 0, & \text{otherwise.} \end{cases}$$

Formulation

$$\text{Minimize} \quad \sum_{(i,j)\in L} c_{ij}x_{ij} + \sum_{(i,j)\in L} y_{ij}$$

$$\text{Subject to} \quad \sum_{(S,j)\in L} x_{Sj} - \sum_{(j,S)\in L} x_{jS} = 1, \tag{A.37}$$

$$\sum_{(D,j)\in L} x_{Dj} - \sum_{(j,D)\in L} x_{jD} = -1, \tag{A.38}$$

$$\sum_{(i,j)\in L} x_{ij} - \sum_{(j,i)\in L} x_{ji} = 0, \quad \forall i \neq S, D, \tag{A.39}$$

$$\sum_{(S,l)\in L} v_{ij}^{Sl} - \sum_{(l,S)\in L} v_{ij}^{lS} \geq x_{ij}, \quad \forall (S,l),(l,S),(i,j) \in L, \tag{A.40}$$

$$\sum_{(l,D)\in L} v_{ij}^{lD} - \sum_{(D,l)\in L} v_{ij}^{Dl} \geq x_{ij}, \quad \forall (D,l),(l,D),(i,j) \in L, \tag{A.41}$$

$$\sum_{(l,k)\in L} v_{ij}^{lk} - \sum_{(k,l)\in L} v_{ij}^{kl} = 0, \quad \forall (i,j) \in L, \forall k \neq S, k \neq D, \tag{A.42}$$

$$v_{ij}^{ij} + v_{ji}^{ij} = 0, \quad \forall (i,j) \in L, \tag{A.43}$$

$$y_{lk} \geq d_{ij}^{lk}(v_{ij}^{lk} - x_{lk}), \forall (i,j),(l,k) \in L, \tag{A.44}$$

$$x_{ij} \geq v_{ij}^{lk}, \quad \forall (i,j),(l,k) \in L, \tag{A.45}$$

$$x_{ij}, y_{ij}, v_{ij}^{lk} \in \{0,1\}, \forall (i,j),(l,k) \in L. \tag{A.46}$$

Bibliography

[1] T. Wu, *Fiber Network Service Survivability,* Norwood, MA: Artech House, 1992.

[2] O. Gerstel, *Opportunities for Optical Protection and Restorations,* Proc., OFC '98, San Jose, CA, vol. 2, pp. 269-270, February 1998.

[3] P. Bonenfant, *Optical Layer Survivability: A Comprehensive Approach,* Proc., OFC '98, San Jose, CA, vol. 2, pp. 270-271, February 1998.

[4] Rajiv Ramaswami and Kumar N. Sivarajan, *Routing and Wavelength Assignment in All-Optical Networks,* IEEE/ACM Transactions on Networking, vol. 3, no. 5, pp. 489-500, October 1995.

[5] Rajiv Ramaswami and Adrian Segall, *Distributed Network Control for Optical Networks,* IEEE/ACM Transactions on Networking, vol. 5, no. 6, pp. 936-943, Dec. 1997.

[6] Rajiv Ramaswami and Kumar N. Sivarajan, *Optical Networks: A Practical Perspective,* Morgan Kaufman, 1998.

[7] G. Ellinas and T.E. Stern, *Automatic Protection Switching for Link Failures in Optical Networks with Bi-Directional Links,* Proceedings of Globalcom, pp. 152-156, 1996

[8] Dimitris Bertsimas and John N. Tsitsiklis, *Introduction to Linear Optimization,* Athena, 1997.

[9] R.R. Iraschko, M.H. MacGregor, and W.D. Grover, *Optimal Capacity Placement for Path Restoration in Mesh Survivable Networks*, Proceedings of ICC, pp. 1568-1574, 1996.

[10] Murali Kodialam and T.V. Lakshman, *Dynamic Routing of Bandwidth Guaranteed Tunnels with Restoration,* Infocom 2000, April 2000.

[11] M. Kuznetsov, N. M. Froberg, Eytan Modiano, and et. al., *A Next-Generation Optical Regional Access Network*, IEEE Communications Magazine, vol. 38, Issue 1, pp. 66-72, Jan. 2000.

[12] S. Ramamurthy and B. Mukherjee, *Survivable WDM mesh networks, Part I - Protection,* INFOCOM '99, Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 2, pp. 744-751, 1999.

[13] S. Ramamurthy and B. Mukherjee, *Survivable WDM mesh networks. II. Restoration,* Communications, 1999. ICC '99. 1999 IEEE International Conference on, vol. 3, pp. 2023-2030, 1999.

[14] Eytan Modiano and Aradhana Narula, *Survivable Routing of Logical Topologies in WDM Networks,* Infocom 2001, Anchorage, Proceedings, IEEE, vol. 1, pp. 348-357, 2001.

[15] Bala Rajagopalan, Dimitrios Pendarakis, Debanjan Saha, Ramu S. Ramamoorthy, and Krishna Bala, *IP over Optical Networks: Architectural Aspects,* IEEE Communications Magazine, pp. 94-102, Sep. 2000.

[16] Muriel Médard, Steven G. Finn, Richard A. Barry, and Robert G. Gallager, *Redundant Trees for Prepalnned Recovery in Arbitrary Vertex-Redundant Graphs,* IEEE/ACM Transactions on Networking, vol. 7, no. 5, PP. 641-652, Oct. 1999.

[17] Steven S. Lumetta, Muriel Médard, and Yung-Ching Tseng, *Capacity versus Robustness: A Tradeoff for Link Restoration in Mesh Networks,* Journal of Lightwave Technology, vol. 18, Issue 12, pp. 1765-1775, Mar. 2000.

[18] Salman Z. Shaikh, *Span-Disjoint Paths for physical Diversity in Netowrks",* Computers and Communications, Proceedings., IEEE Symposium on, pp. 127-133 1995.