Regenerator Placement and Traffic Engineering with Restoration in GMPLS Networks

Emre Yetginer, Ezhan Karasan*

Department of Electrical and Electronics Engineering, Bilkent University, Bilkent 06800, Ankara, Turkey

E-mail: {yetginer, ezhan@ee.bilkent.edu.tr}

Received July 29, 2002; Revised and Accepted April 8, 2003

Abstract. In this paper, we study regenerator placement and traffic engineering of restorable paths in generalized multiprotocol label switching (GMPLS) networks. Regenerators are necessary in optical networks in order to cope with transmission impairments. We study a network architecture where regenerators are placed only at selected nodes for decreasing cost of regeneration. We propose two heuristic algorithms for optimum placement of these regenerators. Performances of these algorithms in terms of required number of regenerators and computational complexity are evaluated. In this network architecture with sparse regeneration, off-line computation of working and restoration paths is studied for traffic engineering with path rerouting as the restoration scheme. We study two approaches for selecting working and restoration paths from a set of candidate paths and formulate each method as an integer linear programming (ILP) problem. A traffic uncertainty model is developed in order to compare these methods based on their robustness with respect to changing traffic patterns. Traffic engineering methods are compared based on number of additional demands resulting from traffic uncertainties that can be carried over the network. Proposed heuristic regenerator placement algorithms are also evaluated from a traffic engineering point of view.

Keywords: GMPLS, traffic engineering, regenerator placement, restoration.

1 Introduction

Multiprotocol label switching (MPLS) is primarily developed for Internet protocol (IP) networks so that virtual connections can be established between two points in an IP network. One of the most important applications of MPLS is traffic engineering which has the goal of using the network resources more efficiently [1,2]. The idea of extending MPLS as a control plane that can be used not only with IP routers, but also with other equipment such as optical cross-connects (OXC) is called generalized multiprotocol label switching (GMPLS) or multiprotocol lambda switching (MP λ S) [3]. The idea of a common control plane is essential in the evolution of open and interoperable optical networks, and has many advantages. First, a common control plane simplifies operations and management, thus reduces the cost of operation. Next, it provides a wide range of deployment scenarios ranging from overlay model to peer model. Besides, building the common control plane from a proven signaling and routing protocol

reduces the software development time. GMPLS provides a simpler and cost-effective network architecture which is capable of carrying a wide range of data-streams and very large volumes of traffic. GMPLS-based photonic multilayer routers have already been developed [4].

Naturally, to adopt to the non-ideal behavior of photonic switches, some modifications and additions to MPLS routing and signaling protocols are necessary, and these are being standardized by IETF [5,6]. The issues and challenges involved in developing a standardized optical network control plane have been addressed [3,7]. Signaling, routing, and management enhancements for GMPLS have been studied [6,8–10]. Different architectures and algorithms for deploying IP over optical networks have been considered [11,12]. Although GMPLS has many advantages, there are several issues that must be contemplated while applying MPLS to the optical layer, such as restoration performance [13] and the effects of physical layer limitations [14].

*Corresponding author.

There are major differences between routing in optical and IP networks. In conventional IP networks, packets are forwarded on a hop-by-hop basis while in optical networks, an end-to-end connection or light-path is established based on network topology and available resources. In optical networks, routing protocols are used to update network topology and resource status information, but are not involved in data forwarding. Another difference is the separation of the control plane from the data plane. In IP networks, control channels are embedded in the same data-bearing channels, i.e., in-band control signaling is used. On the other hand, in optical networks control information is carried in an out-of-band fashion, e.g., via an optical supervisory channel.

There are also physical layer constraints imposed by various analog transmission concerns that affect routing in optical networks. These impairments can be classified in two categories: linear and nonlinear. Linear effects are independent of the signal power and affect wavelengths individually. Amplifier spontaneous emission, polarization mode dispersion, and chromatic dispersion are examples of linear impairments. Nonlinear effects not only generate dispersion on each channel but they also produce crosstalk between channels.

Wavelength continuity is another constraint specific to optical networks. If wavelength conversion is not available at each node, wavelength continuity has to be preserved along each path or path-segment. This complicates the routing and wavelength assignment computation and increases the size of the link state information since wavelength resource information should also be considered in the routing process.

Since a higher degree of multiplexing is done and much more traffic is carried over a single link, failures can affect much more users in optical networks. Thus, survivability is a critical issue in optical networking. Diversity routing is a common technique which is used to provide fast protection or restoration capability. Diversity refers to the situation where two lightpaths have no single point of failure. For diversity routing, fiber, conduit, and right-of-way diversity requirements can be considered. For this aim, a link attribute called shared risk link group (SRLG) is introduced for supporting diversity routing [6]. SRLG information is used to denote all links subject to a similar type of failure at a lower layer. For

example, a fiber cut affects all the fibers in the same conduit, thus there is no point in using a recovery path over a fiber which is in the same conduit with the fiber carrying the working traffic.

In the optical networks currently being deployed, each link is optically isolated by transponders that are employing O/E/O conversions, and these networks are called opaque [15]. Since transponders increase the network cost and they are bit-rate and format specific, there is a tendency to introduce all-optical subnetworks called domains of transparency [16]. Although this architecture has some advantages such as multi-vendor operability, it also has some drawbacks. In each all-optical subnetwork, it is assumed that all routes have adequate signal quality and to ensure that, the geographic sizes of these all-optical islands are kept limited. With increasing bit rates, e.g., from OC-48 to OC-192 and OC-768, transmit powers are increasing that in turn increases transmission impairments and further limits the geographical extents of all-optical subnetworks.

In this work, we assume that the standardization of optical network elements and protocols will lead to optical networks where multi-vendor operability is not an issue. Therefore we envision the optical network as a single entity, and we deploy regenerators at selected nodes in order to maintain the connectivity of the network subject to link failures and optical layer impairments at the minimum cost of regeneration. With this approach, smaller number of regenerators are used compared to the domains of transparency architecture where optical signals are regenerated at the boundaries of each all-optical island. We also assume that optical wavelength conversion is available at all nodes.

In Section 2, we present two heuristic regenerator placement algorithms with the objective of using minimum number of regenerators considering physical layer constraints and all possible single link failures. We then develop the traffic engineering approach in Section 3 which is used for comparing the efficiencies of these algorithms with respect to routing. Performances of these algorithms are evaluated in terms of the number of required regenerators, computational complexity and capability of each network in handling uncertainties in traffic projections. In Section 4, we present the numerical results obtained on a 32-node mesh network topology.

2 Regenerator Placement

In an optical transport network (OTN), the length of any path segment is limited due to the physical layer impairments. Hence, regeneration of optical signals by either optical or electrical means is inevitable. Since cost of regenerators is an important portion of the total network cost, it is economically beneficial to have regeneration at some selected nodes instead of regeneration at all nodes. In this section, we study the problem of regenerator placement in an optical network. In our model, the requirement for regenerator placement is to have at least two feasible SRLG-disjoint paths (one for working path and the other for the restoration path) between each source-destination pair such that both paths satisfy optical transmission constraints.

In the regenerator placement problem, there is a trade-off between the number of regenerators and the average path length used by working and restoration paths. Having a smaller number of regenerators results in longer paths and higher fiber cost since some traffic should pass through regenerators which are not necessarily on shortest paths. On the other hand, in order to be able to use shorter paths, regenerators have to be placed at a larger number of nodes. In this work, the cost of regeneration is assumed to be the dominant factor in determining the total network cost, and our goal is minimizing the number of regenerators in the network.

The regenerator placement problem can be formulated as an integer linear programming (ILP) model. But the complexity of the problem limits the use of this formulation only to small networks. Therefore, heuristic approaches are needed. In the following subsections, we first present a method for determining a maximal set of all possible SRLG-disjoint paths between a node pair in an optical network with known regenerator locations, and we then develop heuristic regenerator placement algorithms.

2.1 Path Set Generation in the Optical Network

With known regenerator locations, finding the maximum number of SRLG-disjoint paths in an optical network with optical layer constraints is computationally more intensive compared to the case without optical constraints. The main reason is the existence of signal regenerators in the network that affect the path set generation process. The maximum

range constraint which limits the length of any path-segment between regeneration points must be taken into account in this process. Moreover, the available SRLG information should be considered for reliability.

The path set generation problem is formulated below as an ILP model. Suppose the network topology is represented by an undirected graph $G = (V, E)$ where V is the set of nodes and E is the set of links. Each link (i, j) has a length (or attenuation) d_{ij} . The locations of regenerators are known *a priori*, denoted by $r_i = 1$, if a regenerator is placed at node i .

Let D^{sd} correspond to the maximum number of SRLG-disjoint paths between s and d subject to the optical transmission constraints. In order to determine D^{sd} , we first solve the unconstrained maximum flow problem between s and d where optical transmission impairments are not considered [17]. Let $\tilde{D}^{sd}$ denote the solution of the unconstrained problem. We use $\tilde{D}^{sd}$ as the starting solution for the constrained ILP formulation given below, where $\tilde{D}^{sd}$ is decremented until the constrained path set generation problem has a feasible solution, and D^{sd} is given by that value of $\tilde{D}^{sd}$ for which the constrained problem is feasible.

Objective:

$$\text{Minimize } \sum_{\{(i,j) \in E\}} \sum_{k=1}^{\tilde{D}^{sd}} y_{ijk} + \alpha \sum_{\{(i,j) \in E\}} \sum_{k=1}^{\tilde{D}^{sd}} y_{ijk} r_i,$$

Subject to:

$$\begin{aligned} & \sum_{\{j:(i,j) \in E\}} y_{ijk} - \sum_{\{j:(j,i) \in E\}} y_{jik} \\ &= \begin{cases} 1, i = s \\ -1, i = d \\ 0, i \neq s, d \end{cases} \quad \forall i \in V, 1 \leq k \leq \tilde{D}^{sd}, \end{aligned} \quad (1)$$

$$\sum_{k=1}^{\tilde{D}^{sd}} \sum_{(i,j) \in S_m} (y_{ijk} + y_{jik}) \leq 1, 1 \leq m \leq L, \quad (2)$$

$$\begin{aligned} w_{ik}^+ - w_{jk}^- + y_{ijk}(d_{ij} + M) &\leq M, \quad \forall (i,j) \in E, \\ 1 \leq k &\leq \tilde{D}^{sd}, \end{aligned} \quad (3)$$

$$w_{ik}^+ = w_{ik}^-(1 - r_i), \quad \forall i \in V, 1 \leq k \leq \tilde{D}^{sd}, \quad (4)$$

$$w_{ik}^- \leq R_{\max}, \quad \forall i \in V, 1 \leq k \leq \tilde{D}^{sd}, \quad (5)$$

$$w_{sk}^+ = w_{sk}^- = 0, \quad 1 \leq k \leq \tilde{D}^{sd}, \quad (6)$$

$$y_{ijk} \in \{0, 1\},$$

where y_{ijk} is the decision variable defined as $y_{ijk} = 1$, if the k -th path uses link (i, j) , and $y_{ijk} = 0$, otherwise. The variables w_{ik}^- and w_{ik}^+ denote the path lengths for the k -th flow into and out of node i , respectively. The set of links belonging to SRLG m is given by S_m , and L denotes the total number of SRLGs. Finally, $R_{\max}$ denotes the maximum allowable length (attenuation) of a path segment subject to optical transmission constraints.

In this formulation, the objective is to minimize the total number of hops in the path set. The secondary objective is to minimize the total number of regenerators used by all paths. This is established by using a small number, α , as the coefficient for the second summation in the objective function, so that the minimization of the total number of links has the higher priority. As an example, α can be chosen as $\alpha = 1/(2\tilde{D}^{sd}|E|)$ where $|E|$ denotes the number of links in G . The number of regenerators used by the path set is minimized in order to avoid unnecessary use of regenerators. Constraint (1) is used to ensure path continuity for each path. Constraint (2) guarantees that at most one link from each SRLG can be used in the path set. The second term in the summation is needed since the links are bidirectional. Constraint (3) is used to determine the length of the path segment from the last regeneration or source node to any node on the path. M is a large number so that only links with positive flows are considered in the path segment length computations, and M can be chosen as any number larger than $R_{\max}$. Constraint (4) terminates a path segment at a node where regeneration occurs, and a new path segment is started with an initial length of zero. Constraint (5) limits the length of any path segment to be smaller than $R_{\max}$, and constraint (6) is used to initialize the path lengths at the source node.

If this problem turns out to be infeasible, $\tilde{D}^{sd}$ is decremented, and the same problem is solved until a feasible solution is obtained. The same procedure is applied separately for each demand pair (s, d) , and the set of all SRLG-disjoint feasible paths between all demand pairs is obtained. The traffic engineering methods to be presented in Section 3 select appropriate working and restoration path pairs from these path sets. The ILP formulation for the path set generation problem can be extended so that other additive optical transmission impairments can be incorporated into the formulation by adding extra constraints similar to (3)–(6).

Having presented the constrained path set generation, we are now ready to discuss the regenerator placement algorithms.

2.2 Maximum Infeasibility Reduction (MIR)

Algorithm

This method is similar to the maximum-descent algorithm. At each iteration, we first compute the total number of node pairs for which two SRLG-disjoint feasible paths cannot be found, and a regenerator is placed at the node which eliminates the maximum number of infeasible node pairs. MIR aims to place a minimum number of regenerators necessary to guarantee existence of at least two feasible SRLG-disjoint paths between each source-destination pair. MIR is described below where N corresponds to the number of nodes in the network and T is the set of all possible node pairs, i.e., $T = \{[i, j] \in V \times V \text{ and } i \neq j\}$.

1. Initialization: $r_i = 0 \forall i \in V$, $n_{reg} = 0$, $done = 0$, $g_{ij} = 0 \forall [i, j] \in T$.
2. Solve the ILP path set generation formulation given in Section 2.1 for each source-destination pair $[s, d]$ with initial number of paths $\tilde{D}^{sd} = 2$. Set $g_{sd} = 1$ for the source-destination pairs for which a feasible solution exists.
3. If $g_{ij} = 1 \forall [i, j] \in T$, set $done = 1$.
4. While not $done$ and $n_{reg} < N$,
 - 4.1. Set $f_i = 0 \forall i \in V$.
 - 4.2. For all nodes i with $r_i = 0$ do
 - 4.2.1. Set $r_i = 1$.
 - 4.2.2. Solve the path set creation formulation for all source-destination pairs $[s, d]$ for which $g_{sd} = 0$.
 - 4.2.3. Set $f_i = \text{number of feasible solutions}$.
 - 4.2.4. Set $r_i = 0$.
 - 4.3. Set $r_i = 1$ for i which maximizes f_i ; increment n_{reg} ; set $g_{sd} = 1$, for source-destination pairs (s, d) for which a feasible solution is obtained by setting $r_i = 1$.
 - 4.4. If $g_{ij} = 1 \forall [i, j] \in T$, set $done = 1$.
5. If not $done$, there is no feasible solution for this problem.

Else the solution is the set of nodes for which $r_i = 1$.

MIR requires the solution of the path set generation problem for each source-destination pair for possible

placement of each candidate regenerator, i.e., for placing each regenerator, the path set generation problem is solved $O(N^3)$ times. In the next section, a more efficient algorithm for solving the regeneration placement problem is presented.

2.3 Maximum Regeneration Demand (MRD)

Algorithm

MRD algorithm uses a different approach: Instead of trying all nodes for each regenerator placement, which is computationally inefficient, two paths for each node pair are determined so that the total number of required regenerators for both paths is minimized. After all paths are computed for all node pairs, a regenerator is placed at the most demanding node, i.e., the node where the number of paths requiring regeneration is maximized. With the new regenerator locations, for each node pair two paths are recomputed such that the total number of regenerators is minimized, and the same procedure is continued until two paths can be found for all node pairs such that no additional regenerators are required. The algorithm for this method is given below.

1. Initialization: $r_i = 0 \forall i \in V, n_{reg} = 0, done = 0$.
2. While not *done* and $n_{reg} < N$,
 - 2.1. For each source-destination pair $[s, d] \in T$, calculate two link disjoint paths, using the *ComputeBestPaths*(s, d).
 - 2.2. Determine $\{t_i\}$, where t_i is the number of regeneration points assigned to node i obtained by calling $t_i = \text{ComputeRegenerationPoints}(i)$.
 - 2.3. Set $t_{max} = \max_i \{t_i\}$.
 - 2.4. If $t_{max} = 0$, set *done* = 1,
 - Else, set $r_i = 1$, for node for which $t_i = t_{max}$, increment n_{reg} .
3. If not *done*, there is no feasible solution for this problem.
 - Else, the solution is the set of nodes i for which $r_i = 1$.

The formulation for determining the path set for node pair $[s, d]$ which is used by *ComputeBestPaths*(s, d) in step 2.1 tries to find two SRLG-disjoint paths using the minimum number of regenerators. The ILP formulation for this problem is given below.

Objective

$$\text{Minimize} \quad \sum_{i \in R_{sd}} \sum_{k=1}^2 m_{ik} + \beta \sum_{(i,j) \in E} \sum_{k=1}^2 y_{ijk},$$

Subject to:

$$\begin{aligned} & \sum_{\{j:(i,j) \in E\}} y_{ijk} - \sum_{\{j:(j,i) \in E\}} y_{jik} \\ &= \begin{cases} 1, & i = s \\ -1, & i = d \\ 0, & i \neq s, d \end{cases} \quad \forall i \in V, k = 1, 2 \end{aligned} \quad (7)$$

$$\sum_{k=1}^2 \sum_{(i,j) \in S_m} (y_{ijk} + y_{jik}) \leq 1, \quad 1 \leq m \leq L, \quad (8)$$

$$w_{ik}^+ - w_{jk}^- + y_{ijk}(d_{ij} + M) \leq M, \quad \forall (i,j) \in E, k = 1, 2 \quad (9)$$

$$w_{ik}^+ = w_{ik}^-(1 - r_i), \quad \forall i \in V, k = 1, 2 \quad (10)$$

$$\frac{1}{R_{max}} w_{ik}^- - m_{ik} \leq 1, \quad \forall i \in R_{sd}, k = 1, 2 \quad (11)$$

$$w_{sk}^+ = w_{sk}^- = 0, \quad k = 1, 2 \quad (12)$$

$$y_{ijk} \in \{0, 1\}, \quad m_{ik} \geq 0, \quad m_{ik} \in \mathbb{Z}.$$

In this formulation, y_{ijk} , w_{ik}^- , w_{ik}^+ , r_i , R_{max} , and M are the same as defined in Section 2.1. R_{sd} is the set of nodes defined as $R_{sd} = \{i : i = d \text{ or } r_i = 1\}$, and L is the total number of SRLGs. The auxiliary variable m_{ik} denotes the smallest number of regenerators required to make the segment of the k -th path between s and d up to node $i \in R_{sd}$ feasible. For instance, if a path length into some node i , w_{ik}^- , is smaller than R_{max} , $m_{ik} = 0$, which indicates that there is no need to place a regenerator on this path-segment. On the other hand, if $w_{ik}^- = 2.5 \times R_{max}$, then $m_{ik} = 2$, which implies that at least two regenerators have to be placed on this path-segment to make it feasible.

The objective of this formulation is to minimize the total number of regenerators needed to make both paths feasible. As a secondary objective, the total number of hops in the path set is minimized. This is accomplished by weighting the second term in the objective function by a small number, β . As an example, β can be chosen as $\beta = 1/(2|E|)$. Equations (7)–(10) and (12) are the same constraints as in the

formulation of Section 2.1. Constraint (11) is used to set m_{ik} to the minimum number of regenerators required to make the path feasible.

Upon calculation of the path set using the above formulation, the best node for regeneration is determined. For this aim, at step 2.2 each node i is assigned a t_i value which is initially zero. For each path, we start from the source node and increment t_i at each node where the length of the path just exceeds $R_{\max}$. This procedure is continued until destination node is reached. Once all paths are considered, the node with the maximum value of t_i is chosen as the best node for regeneration. Regenerator placement is continued until all source and destination pairs have at least two feasible SRLG-disjoint paths. The algorithm used by *ComputeRegenerationPoint(i)* for computing t_i in Step 2.2 is described below.

2.2 For each path

2.2.1. Set length $d = 0, k = 1$.

2.2.2. For the k -th link $l_k = (i_k, j_k)$ on the path

2.2.2.1. Set $d_{last} = d$, and $d = d + w_{l_k}$ where w_{l_k} is the length of link l_k .

2.2.2.2. If $d > R_{\max} > d_{last}$, set $t_{i_k} = t_{i_k} + 1$ and $d = w_{l_k}$.

2.2.2.3. If $r_{j_k} = 1$, set $d = 0$.

2.2.2.4. $k = k + 1$.

In placing each regenerator with the MRD algorithm, the path set generation problem is solved $O(N^2)$ times resulting in less computational complexity compared with the MIR algorithm. Numerical results for the comparison of the MIR and MRD algorithms in terms of number of regenerators and computation times are presented in Section 4 for a 32-node network topology.

3 Traffic Engineering with Restoration

Traffic engineering has the goal of using the available network capacity in an efficient manner in order to carry as much traffic as possible. This requires appropriate routing of all working paths and their corresponding restoration paths. In the following subsections, two traffic engineering methods for off-line computation of working and restoration paths are presented. These methods are formulated as ILP models. The performances of the heuristic regenerator

placement algorithms are compared in terms of capabilities of resulting networks in handling uncertainties in traffic projections.

These traffic engineering methods are not intended to be used for online calculation or for micro flows. Instead, these methods are suitable for routing aggregate demands in the core of the network which uses GMPLS as a means of fast forwarding. The computations are done in an off-line fashion using projected demand and traffic information.

For each working path, the corresponding restoration path is pre-established. We only consider single-event failures, and the capacity needed for restoration on each link is calculated taking into account possible capacity sharing between restoration paths corresponding to different SRLG-disjoint working paths. End-to-end rerouting (global repair) is used for restoration in which the restoration path is SRLG-disjoint from the working path.

Given a demand set consisting of K demands, $Z = \{(s_k, d_k, r_k)\}$, where the triple (s_k, d_k, r_k) denotes the k -th demand with source and destination nodes s_k and d_k , and bandwidth requirement of r_k . The path set $P_k = \{P_{ki}\}$ corresponding to the k -th demand is constructed by using the path set generation formulation presented in Section 2.1, where P_{ki} denotes the i -th path for the k -th demand. Using these sets, traffic engineering methods select a working and restoration path pair for each demand. The efficiency of each method in using network resources is determined using the traffic uncertainty model, which characterizes the discrepancies of the actual traffic demands from projections. The performance of traffic engineering methods are compared by calculating the amount of carried additional demands for each method where additional demands result from the traffic uncertainty model.

3.1 Traffic Engineering with Load Balancing (TELB)

In this section, a design method which jointly optimizes the working and restoration path design problems with load balancing, is discussed. The ILP formulation for this method is given below, where $|P_k|$ is the number of paths in P_k and C_l is the capacity on link l .

Objective:

$$\text{Maximize } z + \gamma \sum_{l \in E} z_l,$$

Subject to:

$$\begin{aligned} \sum_{i=1}^{|P_k|} \sum_{j=1}^{|P_k|} v_{kij} &= 1, \quad 1 \leq k \leq K, \\ v_{kij} &= 0, \quad \text{if } i = j, \quad 1 \leq i \leq |P_k|, \quad 1 \leq j \leq |P_k|, \\ &\quad 1 \leq k \leq K, \\ \sum_{k=1}^K \sum_{i=1}^{|P_k|} \sum_{j=1}^{|P_k|} v_{kij} r_k \delta_{ki}^{l'} + \sum_{k=1}^K \sum_{i=1}^{|P_k|} \sum_{j=1}^{|P_k|} v_{kij} r_k \delta_{kj}^{l'} \delta_{ki}^l &+ z_{l'} \leq C_{l'}, \quad \forall l \in E, \quad \forall l' \in E, \\ z &\leq z_l, \quad \forall l \in E, v_{kij} \in \{0, 1\}, \quad z \geq 0, \quad z_l \geq 0, \end{aligned} \quad (13)$$

where v_{kij} is the decision variable denoting the working and restoration paths chosen for demand k defined as, $v_{kij} = 1$, if P_{ki} and P_{kj} are chosen as working and restoration paths, respectively, for demand k and $v_{kij} = 0$, otherwise. Auxiliary variables z_l and z denote the residual capacity on link l , and the minimum residual capacity over all links in the network, respectively. The indicator function δ_{ki}^l is defined as $\delta_{ki}^l = 1$, if P_{ki} uses link l and $\delta_{ki}^l = 0$, otherwise.

The objective is to maximize the minimum residual link capacity while simultaneously maximizing the total residual capacity in the network in order to evenly distribute the traffic load. In the objective function, the parameter γ is chosen small so that the maximization of z takes higher priority, e.g., $\gamma < 1 / \sum_l C_l$ can be used. The first constraint ensures that one working and one restoration path is chosen for each demand. The second constraint states that working and restoration paths should be distinct for any demand. The third constraint is the capacity constraint on link l' stating that in the case of failure of link l , the capacity used for working (first term) and restoration paths (second term) on link l' cannot exceed its capacity $C_{l'}$. The last constraint sets z to the minimum of the residual link capacities.

3.2 Traffic Engineering with Weighted Load Balancing (TEWLB)

In a typical network, the traffic injected to the network from some nodes may be much more than the others.

Besides, demands between particular source-destination pairs may be higher than for other node pairs. As a result, some links in the network may face more traffic depending on the traffic distribution and network topology.

The goal of the TELB method is to distribute the residual link capacity as uniform as possible over the network, neglecting the relative importance of each link. This approach may cause some links to become bottlenecks since the capacity usage on links vary depending on the factors stated above. A better design approach may be to have more residual capacities on links that are candidates of being overloaded, e.g., links with high estimated utilization levels. This is accomplished by assigning each link a weight which is inversely proportional with the estimated utilization level on that link. The links with high probability of usage are given less weight, so that maximizing the minimum of the weighted residual link capacities ensures that these links will have more residual capacities. As another application, link weights can be used to increase the reliability of selected paths instead of load balancing. Links with better reliability can be assigned higher weights. By determining link weights based on a reliability measure, links with better reliability records will be utilized more after traffic engineering which in turn decreases the effects of failures on traffic over working paths.

TEWLB is similar to TELB except that in order to take into account the relative importance of each link, the constraint stated in (13) is replaced by

$$z \leq \omega_l z_l,$$

where ω_l denotes the relative weight of link l . TELB can be obtained from TEWLB by assigning unit weights to each link.

In this work, the link weight is determined based on the expected utilization level on the link. For each source-destination pair, a demand with one unit capacity requirement is assumed, and the corresponding path set comprising all SRLG-disjoint paths between each source-destination pair is obtained. These paths correspond to possible paths to be used by working and restoration traffic. One unit of bandwidth is then assigned to each such path, and the amount of bandwidth, B_l , used on link l is computed. The utilization level for link l is defined as the ratio $U_l = B_l / C_l$, where C_l is the capacity of link l . U_l corresponds to the expected utilization on link l .

when traffic is uniformly distributed between all node pairs. Then, each link is assigned a weight $w_l \sim 1/U_l$, i.e., the link weight is inversely proportional with the expected utilization level.

3.3 Traffic Uncertainty Modeling

The demands on a network are not deterministic quantities. They are typically obtained from some traffic measurements and forecasts, and link capacities are designed based on traffic projections. These capacities are expanded typically every few years in order to cope up with increasing traffic demand and to relieve bottlenecks in some part of the network occurring as a result of deviations from traffic projections. An important performance measure of any working and restoration path design methodology is its robustness against traffic uncertainty. The designed network should be able to delay the trivial and expensive solution of capacity expansion as much as possible by efficiently using the available capacity.

To compare the relative efficiencies of the two traffic engineering methods and the two regenerator placement algorithms developed in this work, traffic uncertainty is modeled as additional demands on top of the given demands. We compare TELB and TEWL B by calculating the number of additional demands that can be carried for each design. We assume that existing working paths are not allowed to be reconfigured in order to minimize the effect of reconfiguration on existing traffic. But the existing restoration paths can be re-optimized in order to maximize the number of carried new connection requests. The performance measure is taken as the number of additional demands the network can carry under each design.

The ILP formulation for traffic uncertainty modeling is given below. The subscript k is used for existing demands, and k_e is used to denote the additional demands. The number of additional demands is given by K_e . Path sets $\{P_k\}$ are updated so that the working paths for existing demands are discarded, and the reduced path sets $\{P_k^*\}$ are obtained. P_k^e corresponds to the path set for additional demand k_e . The capacity of each link is reduced by the total capacity used by all working paths on that link, so the set of modified link capacities, $\{C_l^*\}$, is obtained. In the following, $|\cdot|$ denotes the set cardinality.

Objective:

$$\text{Maximize} \quad \sum_{k_e=1}^{K_e} \sum_{i=1}^{|P_k^e|} \sum_{j=1}^{|P_k^e|} v_{k_e ij},$$

Subject to:

$$\begin{aligned} \sum_{i=1}^{|P_k^*|} y_{ki} &= 1, \quad 1 \leq k \leq K, \\ v_{k_e ij} &= 0, \quad \text{if } i = j, \quad 1 \leq i \leq |P_k^e|, \quad 1 \leq j \leq |P_k^e|, \\ &\quad 1 \leq k_e \leq K_e, \\ \sum_{i=1}^{|P_k^e|} \sum_{j=1}^{|P_k^e|} v_{k_e ij} &\leq 1, \quad 1 \leq k_e \leq K_e, \\ \sum_{k_e=1}^{K_e} \sum_{i=1}^{|P_k^e|} \sum_{j=1}^{|P_k^e|} v_{k_e ij} r_{k_e}^i \delta_{k_e i}^j &+ \sum_{k_e=1}^{K_e} \sum_{i=1}^{|P_k^e|} \sum_{j=1}^{|P_k^e|} v_{k_e ij} r_{k_e}^i \delta_{k_e j}^l \delta_{k_e i}^l \\ &+ \sum_{k=1}^K \sum_{i=1}^{|P_k^*|} \varepsilon_{lki} y_{ki} r_k \leq C_l^*, \quad \forall l \in E, \quad \forall l' \in E, \quad l \neq l' \\ v_{k_e ij} &\in \{0, 1\}, \quad y_{ki} \in \{0, 1\}, \end{aligned}$$

where $v_{k_e ij}$ is the decision variable denoting the working and restoration paths chosen for demand k_e defined as $v_{k_e ij} = 1$, if $P_{k_e i}^e$ and $P_{k_e j}^e$ are chosen as working and restoration paths, respectively, for demand k_e and $v_{k_e ij} = 0$, otherwise, and y_{ki} is the decision variable denoting the restoration path chosen for demand k defined as $y_{ki} = 1$, if P_{ki}^* is chosen as restoration path for demand k and $y_{ki} = 0$, otherwise. The indicator function δ_{ki}^l is the path-link incidence function defined as $\delta_{ki}^l = 1$, if P_{ki}^e uses link l , and $\delta_{ki}^l = 0$, otherwise. The indicator function $\varepsilon_{lki l'}$ is defined as $\varepsilon_{lki l'} = 1$, if the existing working path for k -th demand uses link l and P_{ki}^* uses link l' and $\varepsilon_{lki l'} = 0$, otherwise.

The objective is to maximize the number of additional demands that are carried. The first constraint ensures that a restoration path is selected for each existing demand. The second constraint states that the working and restoration paths for each additional demand should be distinct. The third constraint ensures that at most one working and restoration path pair is chosen for each additional demand k_e . The last constraint is the capacity constraint for link l' guaranteeing that in case of

Fig. 1. Network topology used in simulations.

failure of any link l , the capacity constraint on link l' is not violated. The first term on the left-hand side is the necessary capacity for working paths on link l' corresponding to additional demands, and the second and the third terms are the restoration capacities required for additional and existing demands, respectively, in case of failure of link l .

4 Numerical Results

For simulation purposes, the mesh network shown in Fig. 1 is used. The network has a planar topology with 32 nodes and 50 links. Links are bidirectional, and the length of each link is shown next to it in Fig. 1. We assume that source and destination nodes for each demand are randomly chosen from a uniform distribution. Paths for all source-destination pairs are found, and the amount of bandwidth usage on each link is determined. Each link is first assigned a capacity which is proportional with the expected link utilization. In addition to this capacity assignment, a fixed amount of capacity is added to each link. In our numerical results, we assume that each SRLG contains exactly one link, i.e., there are 50 SRLGs.

In this case, SRLG-disjoint paths generated by the formulation in Section 2.1 correspond to link-disjoint paths.

Optimization problems are solved using the CPLEX optimization software package. Using the heuristic methods developed for regeneration placement, numerical results are obtained for three different values of $R_{\max}$, namely $R_{\max} = 1500$, 2000, and 2500. The results are tabulated in Table 1, where the nodes selected for regeneration are written in the order they are selected by the algorithms. For a maximum range of $R_{\max} = 2500$, both algorithms find two regeneration points. For other $R_{\max}$ values, the second method results in one more regeneration node than the first method. Both algorithms find similar nodes for all $R_{\max}$ values.

The computation times (in seconds) for MIR and MRD are given in Table 2 for different values of $R_{\max}$. We observe that while MRD generates a solution with comparable amount of regenerators compared to MIR, the solution is obtained within a significantly smaller amount of time. Furthermore, the computation times with both algorithms increase as $R_{\max}$ decreases, primarily due to increasing number of required regenerators.

Table 1. Regeneration nodes obtained by heuristic algorithms.

$R_{\max}$	MIR	MRD
2500	11, 15	7, 15
2000	11, 8, 21, 9	8, 10, 16, 12, 15
1500	21, 11, 9, 5, 19, 7	11, 5, 9, 21, 7, 14, 19

Table 2. Computation times, in seconds, for MIR and MRD.

$R_{\max}$	MIR	MRD
2500	2319	255
2000	3028	363
1500	3711	412

Table 3. Percentage of additional demands blocked by each method.

Set	$R_{\max} = 1500$				$R_{\max} = 2000$				$R_{\max} = 2500$			
	TELB		TEWLB		TELB		TEWLB		TELB		TEWLB	
	MIR	MRD	MIR	MRD	MIR	MRD	MIR	MRD	MIR	MRD	MIR	MRD
1	9.50	2.25	9.00	1.50	7.25	5.75	6.75	5.00	4.50	3.00	4.00	2.25
2	11.50	1.75	10.00	1.25	10.25	7.75	9.25	6.50	14.25	6.00	13.00	4.50
3	10.25	1.75	6.50	1.00	9.75	1.50	8.75	0.75	4.75	2.50	2.75	2.00
4	19.50	9.75	17.25	8.50	14.00	9.00	12.50	7.25	4.50	7.25	3.75	6.50
5	4.00	3.25	2.75	2.00	12.50	12.50	12.00	11.50	21.25	4.75	18.75	3.50
6	9.00	0.75	6.50	0.25	3.25	2.75	2.50	2.25	5.50	1.50	3.00	0.75
7	16.00	7.75	14.50	6.00	2.50	1.50	2.00	0.50	12.75	2.75	11.25	1.50
8	6.00	1.00	4.75	0.75	6.50	1.50	4.75	0.25	5.50	2.50	4.75	1.75
9	3.00	0.75	1.50	0.25	7.75	0.50	5.00	1.00	11.00	0.00	10.00	0.00
10	6.50	2.75	5.00	1.50	12.00	3.50	10.75	2.25	3.75	0.50	3.50	0.00

The performance of the regenerator placement algorithms are evaluated from a traffic engineering perspective using the TELB and TEWLB methods. Traffic uncertainty modeling of Section 3.3 is used to compare the robustness of each method to uncertainties in traffic demands. Ten different demand sets, each consisting of 80 demands with randomly chosen source and destination nodes, are created. Each demand has a randomly selected capacity requirement of either 1, 2 or 3 unit bandwidth. Corresponding to each demand set, twenty additional demand sets, each having twenty randomly generated demands, are created. The average number of additional demands that can be carried is used as the performance measure. The same demand sets are used with each regenerator placement algorithm for comparison purposes.

The results obtained for the three $R_{\max}$ values are shown in Table 3. For each design method, the percentage of additional demands that cannot be carried is tabulated for both regenerator placement algorithms. From a path design point of view, the results demonstrate that TEWLB, outperforms TELB with both regenerator placement algorithms and for all the three maximum range values. TEWLB, results in blocking percentages (defined as the percentage of additional demands that can not be carried by the network) which are on the average 30–40% lower than the blocking percentages obtained by TELB.

Using these results, the efficiency of the regenerator placement algorithms can also be evaluated from a traffic engineering point of view. For $R_{\max} = 1500$, MRD results in one more regenerator than MIR. But

MRD is much better in terms of its robustness against traffic uncertainties. For each design method, MRD has blocking ratios which are nearly one third of the percentages obtained for MIR. Similarly, for $R_{\max} = 2000$, MRD uses one more regenerator than MIR, but the blocking percentages are nearly halved for both TELB and TEWLB with MRD. And finally for $R_{\max} = 2500$, although both algorithms require two regeneration points, MRD is more efficient from a traffic engineering point of view, since it decreases the blocking ratios nearly to one third of the percentages obtained by MIR for both TELB and TEWLB. In summary, the results show that MRD is much more efficient from traffic engineering perspective.

5. Conclusion

In this work, we studied regenerator placement and traffic engineering of restorable paths in GMPLS networks subject to optical physical layer constraints and diversity requirements using SRLGs. We proposed two heuristic algorithms, MIR and MRD, for the efficient placement of regenerators. We studied two traffic engineering approaches, namely TELB and TEWLB, and a traffic uncertainty model in order to compare these approaches based on their robustness with respect to changing traffic patterns. We compared MIR and MRD based on the resulting number of regenerators, traffic engineering aspects, and their computational complexities. While MRD algorithm which is computationally more efficient than MIR results in slightly larger number of

regenerators, it makes regenerator placement assignment such that traffic engineering methods perform significantly better. Among the two traffic engineering methods TEWLb achieves better results in handling the traffic uncertainty by using expected link utilizations in path computation. On the average, TEWLb achieves blocking ratios that are 30–40% lower than the ratios obtained by using TELb.

Acknowledgment

Work partially supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under the project 199E005.

References

- [1] D. O. Awduche, MPLS and traffic engineering in IP networks, *IEEE Communications Magazine*, vol. 37, no. 12, (Dec. 1999), pp. 42–47.
- [2] X. Xiao, A. Hannan, B. Bailey, L. M. Ni, Traffic engineering with MPLS in the Internet, *IEEE Network*, vol. 14, no. 2, (March/April 2000), pp. 28–33.
- [3] D. Awduche, Y. Rekhter, Multiprotocol lambda switching: combining MPLS traffic engineering control with optical crossconnects, *IEEE Communications Magazine*, vol. 39, no. 3, (March 2001), pp. 111–116.
- [4] K. Sato, et al., GMPLS-based photonic multilayer router (Hikari Router) architecture: an overview of traffic engineering and signaling technology, *IEEE Communications Magazine*, vol. 40, no. 3, (March 2002), pp. 96–101.
- [5] A. Chiu, J. Strand, R. Tkach, J. Luciani, Features and requirements for the optical layer control plane, Internet draft (draft-chiu-strand-unique-olcp-02.txt), Feb. 2001.
- [6] S. Chaudhuri, G. Hjalmtysson, J. Yates, Control of lightpaths in an optical network, Internet draft, (draft-chaudhuri-ip-olxc-control-00.txt), Feb. 2000.
- [7] G. M. Bernstein, J. Yates, D. Saha, IP-centric control and management of optical transport networks, *IEEE Communications Magazine*, vol. 38, no. 10, (Oct. 2000), pp. 2–8.
- [8] Z. Zhang, J. Fu, D. Guo, L. Zhang, Lightpath routing for intelligent optical networks, *IEEE Network*, vol. 15, no. 4, (July–Aug. 2001), pp. 28–35.
- [9] A. Banerjee, et al., Generalized multiprotocol label switching: an overview of signaling enhancements and recovery techniques, *IEEE Communications Magazine*, vol. 39, no. 7, (July 2001), pp. 144–151.
- [10] A. Banerjee, et al., Generalized multiprotocol label switching: an overview of routing and management enhancements, *IEEE Communications Magazine*, vol. 39, no. 1, (Jan. 2001), pp. 144–150.
- [11] S. Sengupta, R. Ramamurthy, From network design to provisioning and restoration in optical cross-connect mesh networks: an architectural and algorithmic overview, *IEEE Network*, vol. 15, no. 4, (July/Aug. 2001), pp. 46–54.
- [12] B. Rajagopalan, D. Pendarakis, D. Saha, R. S. Ramamoorthy, K. Bala, IP over optical networks: architectural aspects, *IEEE Communications Magazine*, vol. 38, no. 9, (Sept. 2000), pp. 94–102.
- [13] R. Doverspike, J. Yates, Challenges for MPLS in optical network restoration, *IEEE Communications Magazine*, vol. 39, no. 2, (Feb. 2001), pp. 89–96.
- [14] J. Strand, A. L. Chiu, R. Tkach, Issues for routing in the optical layer, *IEEE Communications Magazine*, vol. 39, no. 2, (Feb. 2001), pp. 81–87.
- [15] K. Bala, R. Cordell, E. Goldstein, The Case for Opaque Multiwavelength Lightwave networks, *Proc. IEEE/LEOS Summer Topical Meeting on Global Info. Infrastructure*, (Keystone, CO, Aug. 1995), pp. 58–59.
- [16] P. Green, Progress in optical networking, *IEEE Communications Magazine*, vol. 39, no. 1, (Jan. 2001), pp. 54–61.
- [17] R. K. Ahuja, T. L. Magnanti, J. B. Orlin, *Network Flows*, (Prentice-Hall, 1993).

Emre Yetginer was born in 1977 in Samsun, Turkey. He received his B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey in 1999 and 2002, respectively. From June 1999 to July 2000, he was with the Bilkent University Center for Communications and Spectrum Management. He is presently a research scientist at the Scientific and Technical Research Council of Turkey—National Research Institute of Electronics and Cryptology (TUBITAK-UEKAE). He is working toward his Ph.D. degree in the Department of Electrical and Electronics Engineering at Bilkent University. His current research areas include routing optimization in MPLS/GMPLS networks and mobile ad-hoc networking.

Ezhan Karasan received his B.S. degree from Middle East Technical University, Ankara, Turkey, a B.S. degree from Bilkent University, Ankara, Turkey, and his Ph.D. degree from Rutgers University, Piscataway, New Jersey, all in electrical engineering, in 1987, 1990, and 1995, respectively. During 1995–1996, he was a post-doctorate researcher at Bell Labs, Holmdel, New Jersey. From 1996 to 1998, he was a Senior Technical Staff Member in the Lightwave Networks Research Department at AT&T Labs-Research, Red Bank, New Jersey. Since 1998 he has been an assistant professor in the Electrical and Electronics Engineering Department at Bilkent University, Ankara, Turkey. His current research interests are in the application of optimization and performance analysis tools for the design and analysis of wireless ad hoc networks and optical burst/packet switching networks.