Cost Optimization Using Regenerator Site Concentration and Routing in ROADM Networks

Balagangadhar G Bathula, Rakesh K. Sinha, Angela L. Chiu, Mark D. Feuer, Guangzhi Li, Sheryl L. Woodward, Weiyi Zhang, Robert Doverspike, Peter Magill, and Keren Bergman

***Abstract*—The advent of colorless and non-directional reconfigurable optical-add-drop multiplexers (ROADMs) will enable flexible pre-deployment of optoelectronic regenerators in future optical networks. Compared to the current practice of installing regenerators only when a circuit needs them, pre-deployment will allow service providers to achieve rapid provisioning and restoration. The pre-deployed regenerators should be concentrated in a selected subset of ROADM sites in order to attain high utilization and to reduce operational costs. We prove that the resulting optimization problem is NP-hard and present an efficient heuristic for this problem that takes into account both the cost of individual circuits (regenerator cost and transmission line system cost) and the probability of a given circuit request, as well as the number of regenerator sites. We provide various methods to reduce the number of regenerator sites, if low probability demands are allowed to have slightly costlier paths. Specific network examples show that the proposed heuristic has near optimal performance under most studied scenarios. We present results for several different cost models. We have also evaluated the heuristic for survivable optical networks, in which a second, disjoint path must be supported for each circuit. An extended version of this paper containing proofs, pseudo-codes and additional experimental results is available online [1].**

***Index Terms*—ROADM, regenerator placement, network optimization, and all-optical networks.**

I. Introduction

Traffic on backbone networks has increased by four orders of magnitude over the past 12 years and estimates on the growth rate going forward vary from 30% to 40% per year [2]. In order to sustain this traffic growth, the cost-effectiveness of optical networks must continue to improve. To date, advances in optical networks have been mainly in three areas: 1) improvements in fiber capacity; 2) improvements in optical reach (the distance over which a wavelength signal can be transmitted with adequate fidelity; and 3) the development and widespread adoption of the reconfigurable optical-add-drop multiplexers (ROADMs). If a circuit's path is longer than the system's optical reach (typically 1500-2800 km for modern long-haul systems), then one or more optoelectronic regenerators must be used to restore the signal, and each regenerator adds a cost comparable to a pair of endpoint transceivers. ROADMs enable any wavelength to either bypass a node, terminate at the node, or be regenerated to maintain signal quality. Optical bypass allows operators to deploy far fewer transponders and regenerators, yielding significant cost savings [3]–[5]. At junction nodes where more than two fiber pairs meet, ROADMs also allow wavelength routing from any link to any other link, forming an optical mesh network.

The next generation of backbone networks will be deployed with ROADMs that are both colorless (any add/drop port can serve any wavelength) and non-directional (any add/drop port can be routed to any inter-node path) [5]. Without these capabilities, each regenerator connected to the ROADM would be tied to a specific inter-node fiber pair, and to a pre-determined wavelength. In contrast, colorless and non-directional ROADMs make it economically feasible to pre-deploy regenerators, enabling recovery from network failures without the need for manual intervention and dramatically reducing mean time to repair (MTTR) [6]. Regenerator pre-deployment also supports more rapid provisioning [7] and improved network efficiency through traffic grooming at the optical layer [8]. Finally, it is a necessary step on the path to more dynamic optical networks [9]. However, considering regenerator cost, it behooves system operators to deploy them as efficiently as possible. Reducing the number of sites (we use "site" and "location" interchangeably) at which regenerators are pre-deployed can be an effective means of accomplishing this, provided one picks the regenerator sites wisely [10].

During the network design and planning process, selected network nodes (typically ROADMs) are designated as regenerator sites. The regenerator sites subset is chosen to ensure that regenerators can be placed among the set to satisfy the optical reach constraint for every possible demand. In some cases, additional requirements may also be placed on the routes allowed for a demand. The problem of picking a minimum set of regenerator sites RS is defined as regenerator location or placement problem (RLP) [11], [12]. The problem has been studied in two flavors: (a) the unconstrained-routing regenerator location problem (URLP); and (b) the explicit-routing regenerator location problem (ERLP). URLP does not limit circuit routing in any way. While the solution may be optimal in number of regenerator locations, individual circuits may incur high cost as a result of using longer routes or using more regenerators. ERLP constrains each route to a specified (typically min-distance) path, but individual circuits may use more regenerators than necessary. Paper [13] proves that URLP is NP-hard. They also propose and compare three heuristics for URLP. Paper [14] uses a biased random-key

Balagangadhar G. Bathula, Rakesh K. Sinha, Angela L. Chiu, Mark D. Feuer, Guangzhi Li, Sheryl L. Woodward, Weiyi Zhang, Robert Doverspike, and Peter Magill are with AT&T Labs - Research, Middletown, New Jersey, USA. (e-mail:{balab, sinha, chiu, mdfeuer, gli, sheri, maxzhang, rdd, pete}@research.att.com). Keren Bergman is with Department of Electrical Engineering, Columbia University, New York, NY (e-mail: bergman@ee.columbia.edu).

genetic algorithm to solve the URLP with improved results. Paper [15] proves hardness of several different variants of the regenerator placement problem and gives approximation algorithms with worst-case performance guarantees.

Most previous RLP have focused on minimizing the number of regenerator sites while still being able to route a circuit between any node pair [16], [17]. While minimizing the number of sites is an important consideration, it should be balanced with the sum of the cost of individual circuits. So there are two additional major considerations to the overall network cost: (1) cost of an individual circuit depends on the number of regenerators used as well as its transmission distance, and (2) traffic demands in production networks tend to be highly non-uniform, meaning that probability of a demand varies greatly among node pairs.

In this paper, we take a holistic view of minimizing overall network cost by considering all three factors. In previous work [18], we have proposed a new routing-constrained regenerator location problem that constrains the connections to use least-cost paths (including the cost of regeneration). The cost model can be extended to include the transmission distance (which affects the cost of fiber, buildings, and amplifiers), as well as the regenerator count. For some customers with strict latency requirements, shorter transmission paths are needed, above and beyond the cost of the optical path. In this paper, we build upon the routing-constrained regenerator location problem in [18] to take into account these additional practical considerations. As in [18], the regenerator site choice is based upon a cost model, so that the number of regenerator sites ($|RS|$) is not the minimal number for the URLP. Different from our previous work [18], the heuristics in this paper incorporates a projected traffic matrix to explore the trade-off between the cost of each circuit and the number of $|RS|$. For example, if a node-pair is unlikely to have significant traffic, its path could be slightly more costly than the minimum cost path, if that enables $|RS|$ to be further reduced. In our heuristic, the allowable path cost deviation for each node pair is roughly inversely proportional to the probability of a connection between that node-pair. The overall goal remains to minimize the total network cost, defined as the sum of a per-site cost and the total cost of active circuits. This paper presents more details on the heuristic, proof that the problem is NP-hard, and $|RS|$ for survivable networks (i.e., those supporting diverse paths).

The rest of the paper is organized as follows. In Section II, we start by defining a simple version of the problem where the cost of a circuit is given solely by the number of regenerators being used. We illustrate the problem with a small example network, and provide a proof of NP-hardness. In Section III, we explain the proposed heuristic approach for reducing the regenerator sites and network costs and describe several variations. A lower bound on the optimal solution is provided. In Section IV, we propose a trade-off for reducing number of regenerator sites by allowing small increase in the cost of low probability circuits. We determine the extra regenerator sites needed for link-disjoint paths in Section V. Section VI contains the experimental evaluation results of our algorithms on large scale network topologies. Finally, in Section VII we summarize the work and outline possible extensions.

II. Problem definition

We start by defining a version of the problem where the cost of a circuit is given only by the number of regenerators being used (in most cases, this is the dominant variable component of the circuit cost). The goal is to minimize the number of regenerator sites subject to the constraint that each circuit uses minimum possible number of regenerators. This allows us to present the main ideas of the heuristic in a simple setting. Then as we add more parameters to the problem, we outline the necessary changes to the heuristic.

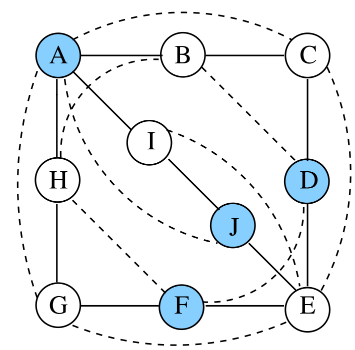

Fig. 1: An example network consisting of 10 nodes and 11 links. Solid lines represent original edges and dotted lines represent augmented edges for reach=2.5 times the link length. Filled circles represent regenerator locations obtained for min-regeneration CRLP.

Let $minregen(u, v)$ be the minimum number of regenerators needed on any route between nodes u and v assuming that regenerators are available at *all* nodes. A route P_{uv} between nodes u and v is called a *constrained* route, if it uses $minregen(u, v)$ regenerators. Constrained-routing regenerator location problem (CRLP) can be formally defined as follows: *Given network topology with link distances, and maximal optical reach distance, find a minimum set of RS such that between each node pair, at least one* constrained *route is reachable using a subset of regenerators in RS.*

Fig. 1 shows an example network with 10 nodes and 11 links. For simplicity, we assume that the link lengths are equal and the optical reach is 2.5 times the link length. For URLP, we only need to place regenerators at 3 locations: $RS = \{A, E, I\}$. Using just these 3 regenerator sites, every node pair has a valid path, but some of these paths (e.g. A-D) use more regenerators and more distance than they would if regenerator sites were unrestricted. For ERLP, we require shortest distance routing between any two nodes, and find that 5 locations are needed: $RS = \{A, C, E, G, I\}$. For CRLP using min-regeneration routes, where we have a little bit more freedom to select routes, the RS set is reduced to 4 locations: $RS = \{A, J, D, F\}$. In this example, we found that URLP, with the most freedom to select routes, has the fewest regenerator locations in the solution, but involves more costly circuits, while ERLP, with no freedom to select routes, requires more regenerator locations, and CRLP lies in between.

We next formally prove the hardness of CRLP problem. We define a *decision version* of the CRLP problem (DCRLP) as following: *Given network topology with link distances, and optical reach distance, is there a set of K or fewer regen sites, RS, such that between each node pair, at least one constrained route is reachable using regenerators in RS.*

Theorem 1. *DCRLP is NP-hard.*

Proof. Our proof uses a reduction from the Vertex Cover Problem (VCP) [19] to DCRLP. Given an undirected graph $G = (N, E)$, where N and E represent set of vertices and set of edges respectively, and a positive integer $K \leq |N|$, the VCP asks if there is a subset $N_0 \subset N$ of cardinality at most K such that N_0 contains at least one of the two end points of each edge in E.

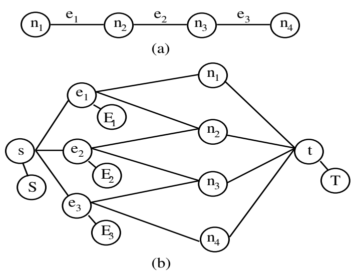

Fig. 2: VCP to DCRLP transformation (a) G (b) G'

Given an instance I_1 of VCP (N, E, K), we construct the corresponding instance I_2 of DCRLP (N', E', K) by the following transformation. (An example is shown in Fig. 2.) (1) for any $n_i \in N$, create a network node $n_i \in N'$; (2) for any $e_i \in E$, create one node $e_i \in N'$, and add following links in E': (e_i, n_k), (e_i, n_l), where n_k, n_l are two end nodes of e_i; (3) add two extra nodes s and t to N', and add following links to E': (s, e_i) for any node e_i and (t, n_j) for any node n_j; (4) add two nodes S and T in N', and links (s, S) and (t, T) to E'; (5) for any node $e_i \in E$, we create another node E_i in N', and add links (e_i, E_i) to E'. Now we have graph $G'(N', E')$. Clearly, I_2 can be constructed from I_1 in polynomial time.

If we assume the optical reach to be one hop, we have the following observations on I_2.

- If $d(x, y)$ denotes the minimum hop distance between nodes x and y, we have $d(s, e_i) = 1$, $d(t, n_j) = 1$, $d(s, n_j) = 2$, $d(t, e_i) = 2$, $d(s, t) = 3$. If n_j is end node of e_i then $d(e_i, n_j) = 1$ else $d(e_i, n_j) = 3$. Moreover, $d(e_i, e_j) = 2$ and $d(n_i, n_j) = 2$.
- It is straightforward to see that S, T, E_i will not be selected as regeneration sites. On the other hands, s, t, and e_i must be in RS because S, T, and E_i have to use s, t, e_i, respectively, as regeneration sites for their traffic.
- To find a feasible solution of I_2, we need to examine possible regenerator sites (from nodes n_i) for node pairs between t and e_i, E_i, s or S.

Based on the observation above, the reduction follows from

- If I_2 has a feasible solution, say set R of regenerator sites then node set R-$\{s, t, e_i\}$ is a feasible solution for I_1.
- Conversely, suppose the instance I_1 of VCP has a feasible solution of node set R' then node set $R' \cup \{s, t, e_i\}$ is a feasible solution of I_2.
- Since VCP is NP-hard, we conclude that DCRLP is also NP-hard. ■

An algorithm for finding an optimal solution for CRLP can solve DCRLP by checking if its solution has cardinality at most K. Thus we get

Corollary 1. *If $P \neq NP$ then no polynomial time algorithm can find an optimal solution to CRLP.*

III. Greedy CRLP heuristic

We first transform the CRLP problem into a graph problem of finding a set of nodes covering min-hop paths. Many other papers [20] and [21] have considered similar transformations. We augment the network graph by adding edges (i, j), whenever nodes i and j are within reach distance. Now all node pairs within reach distance have a direct edge between them. An example transformation is shown in Fig. 1. Let us call the resulting graph G_A; then we claim a 1:1 correspondence between min-regeneration paths in the network and min-hop paths in G_A. (This proof is omitted for brevity, but can be found in [1].)

Lemma 1. *If P is a min-hop path in G_A then placing a regenerator on each internal node of P results in a valid min-regeneration path in the network. Moreover for every valid min-regeneration path in the network, we have a min-hop path in G_A with direct links between adjacent pair of regenerator sites.*

So we have reduced the CRLP problem to picking a subset of nodes, RS, s.t. each pair of nodes has a min-hop path in G_A with all internal nodes in RS. The greedy heuristic maintains the following data structures:

- A set C of candidate regenerator sites (for future placement).
- A binary path matrix P s.t. P_{ij} is 1, iff we have a valid min-hop path between nodes i and j using the reach distance and regenerator site selected so far. The heuristic stops when all entries in P are 1.
- Min-hop matrix D s.t. D_{ij} gives the min-hop distance in G between i and j. The matrix D lets us check whether a node v belongs to min-hop path from i to j as: $D_{iv} + D_{vj} = D_{ij}$? (Pf: $D_{iv} + D_{vj}$ is the length of the min-hop path from i to j that includes node v. By definition it is at least the length of the min-hop path, and the equality will happen when there is indeed a min-hop path going through node v.)

This generic greedy heuristic picks what appears to be the next best site from among nodes in C (to be elaborated shortly), updates its data structures, and repeats these steps until all source-destination pairs have valid min-hop paths (in G_A) using existing regenerator sites. In the next few subsections, we give a number of customized enhancements to this heuristic, including a way of estimating how far it is from the optimal.

We use $O(|E|+|N|)$ breadth-first-search to compute the single-source shortest min-hop path and $O(|N|^3)$ Floyd-Warshall algorithm for computing all-pairs shortest path.

A. Seeding the greedy algorithm

For any node-pair (a,z), we have to place a regenerator on all intermediate nodes in a min-hop path. Apriori we do not know which min-hop path will be used. However if node v is in *all* min-hop paths between a and z then a regenerator must be placed on node v. Let us call this set of nodes R^+. Seed the algorithm by placing a regenerator on all nodes in R^+.

Conversely if node u is not in any min-hop path then an optimal regenerator assignment will not place a regenerator on node u. Let us call this set of nodes R^-.

We start by placing a regenerator at every node v in R^+. If this regenerator placement creates a valid (i,j) path then we update $P_{ij}=1$. Finally, we initialize the set of candidate regenerator sites $C = N\backslash(R^+ \cup R^-)$.

The set R^+ can be computed in $O(|E|\times|N|^2)$ time by iterating over all nodes v and checking if deleting v changes the min-hop distance for at least one pair. The set R^- can be computed in $O(|N|^3)$ time by iterating over all nodes v and checking that for each pair (a,z), the shortest $a-z$ path *through* v is longer than the shortest $a-z$ path.

B. Rank function for selecting best candidate

We use two different rank functions for picking the "best" candidate in C. One reasonable choice is to select the node that belongs to min-hop paths of highest number of pairs from those that don't already have a valid path. Mathematically, we can define this rank of a candidate node as

$$\text{rank}_1(v) = |\{(i,j)|(P_{ij}=0)\wedge(D_{iv}+D_{vj}=D_{ij})\}| \quad (1)$$

The first term in the Boolean AND expression says that (i,j) doesn't already have a path and the second term says that v is in a min-hop path between i and j.

Another possibility is to only count source-destination pairs which get a valid path as a result of placing a regenerator on node v.

$$\begin{aligned}\text{ramp}(v) \quad = \quad & |\{(i,j)|(P_{ij}=0)\wedge(P_{iv}=1)\wedge(P_{vj}=1) \\ & \wedge(D_{iv}+D_{vj}=D_{ij})\}| \qquad (2)\end{aligned}$$

While (2) results in a rapid convergence, it has a problem. During the initial phase of the heuristic, no single regenerator placement may result in a valid path, thereby assigning rank of zero to all candidate nodes. We get around this by considering a weighted combination of the two rank rules, with a large weight assigned to ramp as:

$$\text{rank}_2(v) = \text{rank}_1(v) + (|N|-1)\times\text{ramp}(v) \quad (3)$$

The overall effect of using (3) is the ramp(v) has the dominant effect in picking the "best" candidate but if it cannot distinguish among multiple candidates, then the rank$_1(v)$ acts as a tie-breaker.

C. Post processing (PP) to improve the solution

The greedy algorithm never deletes a site after it gets selected. So it is possible that it may select node v_1, but then a set of nodes selected later cover all source-destination pairs that v_1 was originally covering, rendering v_1 superfluous.

We add a simple post-processing step. For each regenerator site v in the output, we check if deleting v still enables all source-destination pairs to have valid paths. If yes, we delete v and then repeat the check for the next node. We stop if we cycle through all nodes without being able to delete any of them.

Checking if v can be deleted is similar to checking its membership in R^+ and takes $O(|E|\times|N|)$ time. In the worst case, each deletion may require checking all nodes in the output so altogether it can take time $O(\#$ of RS $\times$ (# deletions in PP) $\times |E|\times|N|)$.

D. Time and space complexity

The running time is given by,

$$O\left(\left(\frac{|RS|\times(\text{\# of deletions in PP})}{d_a}+d_a\right)\times|N|^3\right),$$

where d_a is average degree in G. The space complexity is the size of data structures, $O(|N|^2)$.

It is possible to improve these running times by use of more sophisticated dynamic graph algorithms and (for computing R^+) by adapting algorithms for the related problem of 'most vital nodes'. We chose the simplest algorithms for ease of implementation and because the running times (less than 2 seconds, on a generic PC with 2.3 GHz CPU, for our 75 node topology) suffice for our purpose. We will show later that we get near optimal results with this heuristic. In contrast, our attempt at ILP based solutions (without any special customization) ran out of memory even with commercial packages.

E. Comment on R^+, R^-, and lower bound

The heuristic can be implemented without R^+ and R^- and will still yield a solution where all node pairs have valid paths. They serve different purposes.

Any reasonable ranking algorithm would leave out nodes from R^- so not having R^- does not change the behavior of the algorithm. The usefulness of R^- is that rather than computing the rank of the nodes in R^- in each iteration, we exclude them after a one-time computation of $O(|N|^3)$.

R^+ serves a more critical role. By seeding the algorithm with this set, it can hopefully lead to a better quality solution. Moreover, as the following theorem shows, it can also be used to get a bound on how far the heuristic is from the optimal solution.

Theorem 2. *If the heuristic gives a solution of cardinality $|R^+|$ then it is optimal. Otherwise, $1+|R^+|$ is a lower bound on the cardinality of any solution.*

Proof: By definition of R^+, any solution must contain all nodes in R^+ so a heuristic solution of R^+ is certainly optimal. To see the claim on $|R^+|+1$, notice that the only way the heuristic does *not* stop at R^+ is if those regenerators are not sufficient so we need at least one more regenerator. ■

For the networks we have tested so far, the solution produced by the heuristic turns out to be within at most 1 or 2 of the lower bound. There are many advantages of deriving lower bound: it lets us assess how far we are from the optimal without having to compute the optimal. It also gives us an efficient way of computing the optimal. E.g, if we know that the heuristic solution is (say) within 3 of $|R^+|$. We can try a brute-force approach of adding one site (all $|N|$ possibilities) to R^+ and then checking if any of them cover all node pairs. If not, we try all possible pairs of sites and finally all possible triplets of sites. Because of the upper bound given by the heuristic, we know that at least one of these $O(|N|^3)$ possibilities would succeed and give us the optimal set of regenerator sites.

It is also possible to improve this lower bound by treating the path matrix as the adjacency matrix of a graph and realizing that the resulting graph's diameter reduces by at most 50% as a result of a single regenerator placement. This will give a $\log$(diameter of path matrix) in the lower bound.

F. Minimum cost paths

A min-regeneration path is not necessarily min-distance and vice versa. As an example, consider nodes a and z connected by two disjoint paths $R_1 = a - v1 - v2 - v3 - z$ and $R_2 = a - v4 - v5 - z$. If the reach distance is 2000 km and the length of each link in R_1 is 1050 km and the length of each link in R_2 is 1950 km, then R_1 is the shortest (= min-distance) path. But note that R_1 requires three regenerators, whereas R_2 requires only two and so is the min-regeneration path. Thus, we see that min-regeneration paths may incur distance penalty, also described as excess wavelength-km penalty, and min-distance (shortest) paths may incur a regeneration penalty. By incorporating wavelength-km as well as number of regenerators in the cost model of the path (circuit), one can achieve the required trade-offs. We can define the cost of a path R as:

$$c_r \times \text{number of regenerations in } R + c_m \times \text{length of } R, \quad (4)$$

where c_r (c_m) is unit regeneration (wavelength-km) cost. The two extreme cases are: setting $c_r = 1, c_m = 0$ reduces to the CRLP problem; whereas setting $c_r = 0, c_m = 1$ becomes the ERLP problem with min-distance paths.

We modify the heuristic to find min-cost paths instead of min-regeneration paths as follows:

1) Change the definition of R^+ to the set consisting of nodes that belong to *all min-cost* paths between a and z. A similar change applies to the definition of R^-
2) Change the definition of path matrix, P: P_{ij} is 1, if we have a valid *min-cost* path between nodes i and j using the given regenerator placement.
3) The condition $D_{iv} + D_{vj} = D_{ij}$ now applies to min cost paths.

IV. Number of regenerator sites vs. Cost of paths

So far, we have applied a rigid constraint to min-cost paths only. However, the number of regenerator sites can be further reduced if we allow the heuristic to pick paths that are slightly more costly for rarely used routes. We extend the heuristic to use a *latitude matrix*, L, such that for node pair (i, j), we are allowed to pick any path that is of cost within $1 + L_{ij}$ of the min cost (i, j) path. If we have a traffic projection, we will typically assign small or zero latitude to node-pairs with heavy traffic demand and larger latitude to node-pairs with low traffic demand. The necessary changes to the heuristic are as follows:

1) Change the definition of R^+ to the set consisting of nodes that belong to *all* paths between i and j of cost less than $1 + L_{ij}$ of the min-cost (i, j)-path. To test membership of node v in R^+, we check if deleting v changes the min-cost path for at least one pair (i, j) by more than a multiplicative factor of $1 + L_{ij}$. A similar change applies to the definition of R^-
2) Change the definition of path matrix, P: P_{ij} is 1, if we have a valid path between nodes i and j using the given regen placement s.t. the cost of the path is within $1 + L_{ij}$ of the min-cost (i, j)-path.
3) In post-processing, for each node v in the output, we check if deleting v still enables all source-destination pairs (i, j) to have valid paths of cost within $1 + L_{ij}$ of the min-cost (i, j)-path. We would like to point out a subtlety here with an example. If latitude is 5% and say the path with the original output is within 102% of the min-cost and deleting node v raises the cost to 104% of min-cost then we still delete node v as the cost is within the threshold of latitude. So this is unlike the case without any latitude, where any increase suggests that node v can't be deleted from the output.

We present simulation results for several possible choices of latitude matrix in Section VI.

V. Regenerator sites for Diverse Routes

Survivable optical networks can reconfigure and set-up a connection upon failure, as discussed in prior work [22]. Fast reconfiguration for survivability can be achieved using *link-disjoint* primary and backup (secondary) paths that are precomputed for each request. For a given request, we use the path from CRLP as the primary path, which carries the traffic under normal operation, while the backup path is used upon a link failure. Backup paths are usually unconstrained; here we assume that they can share nodes with primary paths, as long as there are no shared links.

In this section we evaluate the proposed CRLP heuristic described in Section III for survivable optical networks. We determine the set of additional regenerator sites (Δ_{RS}) required for the node-pairs to have a valid reachable secondary path (as long any disjoint path exists). We define $RS_D = RS_{CRLP} \cup \Delta_{RS}$, where RS_{CRLP} refers to the set of regenerator sites obtained using the CRLP heuristic.

First, let us return to the example network shown in Fig. 1. Note that for this topology, all node-pairs have a disjoint paths for the primary CRLP path. We have a CRLP solution for min-regeneration paths as shown in Fig. 1. For every node-pair there exists a min-regeneration primary path, which we denote as $(\mathcal{R})$. We designate the disjoint path $(\mathcal{R}')$ as valid iff it is reachable using the existing set of regenerator sites. For instance the min-regeneration path for node-pair (B, E) is $\mathcal{R}_{BE}$=B−C−D−E and the disjoint path, $\mathcal{R}'_{BE}$=B−A−I−J−E, is valid via regeneration at locations A and J. As an example of a node-pair without a valid disjoint path, consider (A, D) with $\mathcal{R}_{AD}$=A−I−J−E−D and $\mathcal{R}'_{AD}$=A−B−C−D. We define the diverse percentage (p_d) as the percentage of node-pairs with a valid disjoint path. For the network in Fig. 1, using to $RS_{CRLP} = \{A, D, J, F\}$, there are $10 \times 9/2$=45 node-pairs and only 26 node-pairs have valid $\mathcal{R}'$, so $p_d \approx 58\%$. If we require $p_d = 100\%$ for the Fig. 1 network, then we must add sites in addition to the set RS_{CRLP}.

An iterative process is used to choose these extra sites from candidate set C_D, defined as the set of intermediate nodes present in all the invalid disjoint paths, but $\notin RS_{CRLP}$. In each iteration, we select a node $c_d \in C_D$ that belongs to the largest number of invalid disjoint paths $\mathcal{F}(c_d)$. For the network in Fig. 1, we have $\mathcal{F}(B) = 14$. Adding node B as a regenerator site increases p_d to 84.4%. Finally, including with node H (or node G) as well enables all node-pairs to have a valid disjoint path, so that $(p_d = 100\%)$. Thus we select $\Delta_{RS} = \{B, H\}$. Due to space limitations, we present additional details of the algorithm in [1]. For cost reasons, operators usually design optical networks to provide disjoint backup paths for most, but not all, possible node-pairs. The few node-pairs without valid disjoint paths might have backup paths transported over a disjoint path on an alternate or pre-existing optical layer.

VI. Experimental Evaluation

In this section, we evaluate the performance of the proposed CRLP heuristic for various backbone network topologies. We have studied two networks: (1) a US mesh network (USMESH) with 24 nodes and 43 bi-directional links [20] with modified link distances [1] and (2) a continental US topology with 75 nodes and 99 bi-directional links shown in Fig. 3 (CONUS). CONUS is a fiber-optic backbone network developed for use in the research of large-scale DWDM networks [23].

We have compared the CRLP heuristic with an optimal integer linear programming (ILP) solution for min-regeneration on the USMESH topology. For reach distances 1400, 1800, 2000, 2400, and 2500 km, the numbers of regenerator sites were 17, 12, 11, 8, and 9, respectively, for both the CRLP heuristic and the ILP solution. The solutions from both methods also included exactly the same individual regeneration locations. We did not evaluate (optimal) ILP solutions for the CONUS topology because of the very long time to run on a regular desktop (it also ran out of memory in several instances), but the lower bound discussed in Section III-E allows us to assess how far our heuristic is from the optimal.

For the rest of paper we have evaluated the CONUS network because of its close resemblance to carrier backbone networks. In Fig. 3 we show the RS for min-regeneration $(c_r = 1, c_m = 0)$ CRLP.

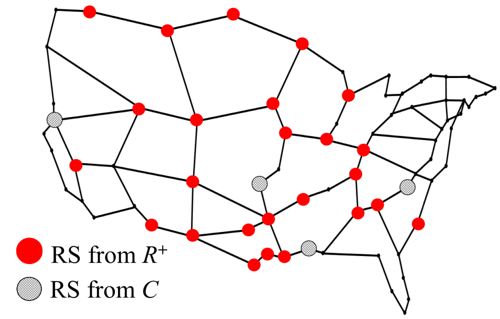

Fig. 3: The CONUS network topology. Circles indicate regenerator sites appearing in the solution set for reach distance of 1800 km, $c_r = 1$ and $c_m = 0$ (Min-regeneration).

A. Effect of seeding with R^+

For all reach distances we considered, the heuristic solution for min-regeneration $(c_r = 1, c_m = 0)$ was within 2 of R^+, meaning that it was at most 1 off from optimal. For a reach distance of 2000 km, we found that R^+ (without any other sites) turned out to be a solution. We were interested in seeing if not seeding with R^+ still leads to a good solution (albeit at slower convergence) so we ran some simulations with a generic greedy algorithm (without R^+) and found the solutions to be inferior in general. The reason may be that a greedy algorithm does *local* optimization, which may not lead to global optimization. Starting with an empty set of sites, there is a greater likelihood of it deviating farther from the global optimum. When the algorithm is seeded with R^+, it only has to pick a few additional sites so its scope of making wrong decisions is minimized. This suggests that our customized enhancement of seeding with R^+ leads to a better quality solution and speeds up convergence.

If the variables in (4) are set to $c_r = 0$ and $c_m = 1$, then CRLP reduces to a min-distance (or min-delay) routing. For the third scenario, min-cost routing, we set the parameters as $c_r = 1000$ and $c_m = 1$. In other words, the cost of one regeneration is equivalent to 1000 km of fiber distance. In Fig. 4(a) we compare the number of regenerator sites obtained for all three scenarios, which were also presented in [18]. We observe that the $|RS|$ is the lowest for min-distance case and highest for min-cost case. Lesser number of regenerator sites ($|RS|$) for min-distance routing, causes more regenerations along the path, and there by increasing the overall network cost [18]. We have also computed the sum of the costs of all circuits (assuming a uniform traffic matrix, where each node-pair has the same number of circuits) for each case, and observed that, as expected, this cost is lowest for the min-cost case [18].

In describing the algorithm (Section III), we suggested two reasonable ranking rules ((1), (3)) and a post-processing (PP) step as an added optimization. We have evaluated the performance of our heuristic with different ranking rules, both with and without PP for all the three scenarios: min-regeneration, min-distance, and min-cost. For min-distance CRLP, the PP step reduces the $|RS|$ especially for the lower reach distances (1500, 1800 and 2000 km). However for the min-regeneration CRLP the PP does not reduce the $|RS|$ for any reach distance and network topology. In the min-cost CRLP, PP improves the solution by one regenerator site for reach distances 2400 and

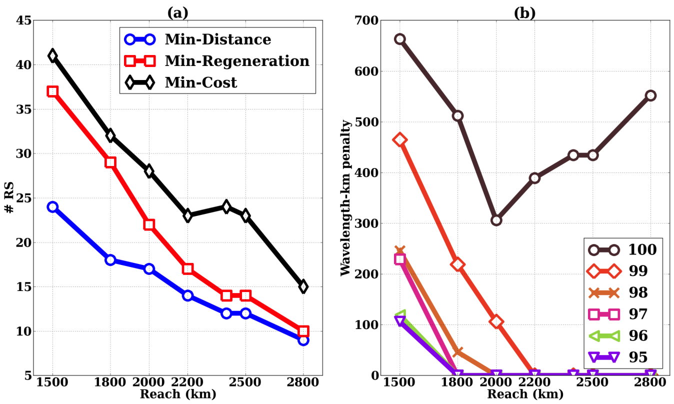

Fig. 4: (a) Number of regeneration sites for different CRLP. (b) Highest excess wavelength penalty among certain % node-pairs for min-cost CRLP.

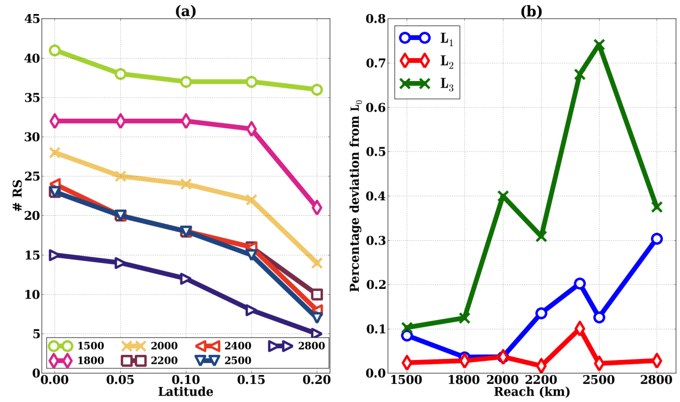

Fig. 5: (a) Number of regeneration sites for min-cost CRLP ($c_r = 1000; c_m = 1$) versus different latitudes for a given reach distance (km). (b) Percentage increase from L_0 for allowing small latitude from the min-cost paths.

2500 km. As explained earlier, we have two different rank rules, and neither was consistently better than the other. Therefore we have reported the minimum of the two solutions. (We can think of the final heuristic as invoking two heuristics in serial order and then picking the better one; this doubles the running time.)

As explained in Section III-F, a min-regeneration path is not necessarily a min-distance (shortest) path and vice-versa. We have evaluated the excess wavelength-km penalty incurred by min-regeneration paths compared to the min-distance paths. Note that these excess wavelength-km penalties translate to longer latencies for circuits using these paths. We observed that for approximately 85% of the node-pairs, the min-regeneration path coincides with the min-distance path, so there is no wavelength-km penalty. We also verified that min-cost paths (because they are trying to minimize a combination of number of regenerators and distance) consistently show lower excess wavelength-km penalty than min-regeneration CRLP [18]. For the rest of the discussion we consider min-cost CRLP.

Fig. 4(b) shows the highest excess wavelength-km penalty for each percentile of node-pairs for min-cost CRLP. That is, we computed the wavelength-km penalty for each node-pair and then computed the highest penalty for each percentile of node-pairs. For example, the 99% value at a reach distance of 2000 km is approximately 100 wavelength-km, meaning that when we pick the circuit routes so as to minimize their overall cost, 99% of the routes incur a distance penalty of less than 100 km (which translates into a latency penalty of less than 0.5 ms).

B. Trade-off between circuit cost and $|RS|$

As explained in Section IV, allowing the heuristic to pick paths that are more costly than min-cost paths can potentially reduce $|RS|$. Fig. 5(a) shows the effect of such latitude for min-cost CRLP, where we define $latitude$ as a parameter indicating the percentage by which a node pair's route is allowed to exceed the minimum cost for that pair.

1) Fixed latitude: We first consider the same latitude for all network node-pairs. The leftmost values in Fig. 5(a) indicate $|RS|$ for strict min-cost, i.e., zero latitude for all node-pairs. From Fig. 5(a) we observe that allowing a certain percentage of increase in the cost, will reduce the $|RS|$ for most of the reach distances. For instance, by allowing 5% (0.05) latitude we observe a reduction from 28 to 25 in the number of sites for reach distance of 2000 km.

2) Variable latitude: We next consider applying different latitude to different node pairs. The basic concept is that if we have a demand matrix, then latitude for a node-pair should be inversely proportional to the amount of demand. Our techniques apply to any general demand matrix, but for the results in this section, we assume that the amount of demand between a node-pair is proportional to the product of their populations. (This is the well-studied gravity model of demand.)

We have studied several different latitude rules. For the first three, we classify node-pairs to be high (h_i, both nodes have a population exceeding 5 million), medium (m_i, both nodes have a population exceeding 1 million and at least one node has population less than 5 million), or low (l_i, all other node pairs). We define latitude scenarios as L_1, L_2, and L_3 as $L_1 = [h_1 = 5\%, m_1 = 10\%, l_1 = 15\%]$, $L_2 = [0\%, 5\%, 10\%]$, and $L_3 = [0\%, 10\%, 25\%]$ respectively. If no latitude was allowed to any node-pair (np) then it is defined as L_0, which is the baseline CRLP.

The next two latitude rules L_4 and L_5 (given in (6) and (7)) can be thought of as continuous versions of the first three latitude definitions. As is standard in gravity based models, we define the probability of a demand between a node-pair to be proportional to the product of the population of the cities at which they are located. If Pop(i) and Pop(j) are population of city at node i and j, then we define the probability of the node-pair (i, j) as:

$$Pr(i,j) = \frac{\text{Pop}(i)\text{Pop}(j)}{\sum\limits_{\forall (i,j), i \neq j} \text{Pop}(i)\text{Pop}(j)} \tag{5}$$

If Pop(i) and Pop(j) both are 5 million each, then we define the rule L_4 such that it allows a latitude of 3%. As the product of the populations decreases, we allow larger latitudes but we cap it at 20%.

$$L_4(i,j) = \min\left(20, \left\lfloor \frac{25}{\text{Pop}(i) \times \text{Pop}(j)} \right\rfloor \times 3\right) \quad (6)$$

In the definition of $L_5(i,j)$, we also cap the latitude at 20% but it allows smoother changes. The definition also guarantees that the expected increase is at most 5% as each node-pair contributes at most $Pr(i,j) \times L_5(i,j) \leq \frac{5 \times Pr(i,j)}{Pr(i,j) \times np} = \frac{5}{np}$ to this expectation.

$$L_5(i,j) = \min\left(20, \frac{5}{Pr(i,j) \times np}\right) \quad (7)$$

Latitude	Reach distance (km)						
	1500	1800	2000	2200	2400	2500	2800
L_0	41	32	28	23	24	23	15
L_1	37	31	24	18	17	18	11
L_2	38	32	24	21	18	19	14
L_3	36	27	21	18	15	16	11
L_4	39	26	21	20	16	15	12
L_5	39	28	23	20	18	17	14

TABLE I: $|RS|$ for various latitude scenarios considered in the proposed heuristic.

From Table I we see that the latitude rule L_3 has lowest $|RS|$, compared to L_1 and L_2. This is not surprising because L_3 allows latitude of up to 25%. However this reduction in $|RS|$ comes with cost penalty. In Fig. 5(b) we compare the percentage increase from the min-cost CRLP (L_0) for L_1, L_2, and L_3 and L_3 has the highest cost penalty.

In Fig. 6(a) we evaluate the excess wavelength-km penalty for various percentages of node-pairs based on latitude rule L_3. Here we observe that among 99% of the node-pairs, the highest deviation is approximately 300 km ($\approx$ 1.5 ms) at reach distance 2000 km, which is higher than L_0 in Fig. 4(b).

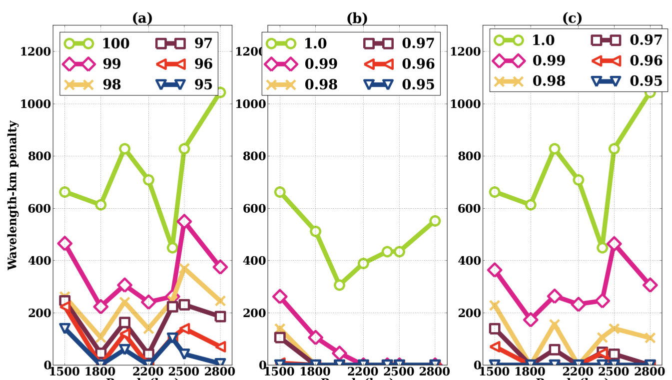

Fig. 6: (a) Highest excess wavelength penalty among certain % node-pairs for min-cost CRLP with latitude L_3. (b) Highest deviation using probability of a node-pair for min-cost CRLP with zero latitude. (c) Highest deviation using probability of a node-pair for min-cost CRLP with latitude L_3.

Similar to the percentile calculations, we compute the highest wavelength-km for various probabilities (p_r) 1.00, 0.99, 0.98 for both L_0 and L_3 as shown in Fig. 6(b) and Fig. 6(c). From Fig. 6(b) and Fig. 6(c) we observe that 99% of demands will have excess wavelength penalty of less than 50 km in the case of L_0 and 300 km for L_3 at optical reach 2000 km.

Each latitude rule presents a different point in the trade-off between minimizing cost of circuits and $|RS|$ (minimizing $|RS|$ can provide savings in shared infrastructure, operations, and pre-deployment of idle regenerators). For instance, latitude rule L_3 reduces the $|RS|$ but there is a comparatively higher wavelength-km penalty. From Table I we see that L_4 and L_5 have number of regenerator sites slightly higher than L_3, but lower than L_0. In Fig. 7(a) we compare the percentage of node-pairs that have excess wavelength-km penalty for latitudes L_0, L_3, L_4, and L_5. We observe in Fig. 7(a) that the percentage node-pairs for L_5 are near to that of L_0.

We next consider *expected deviation* $\bar{D}$ defined as $\bar{D} = \sum_{\forall(i,j),i\neq j} Pr(i,j)\delta(i,j)$, where $\delta(i,j)$ is the wavelength-km penalty for node-pair (i,j). In Fig. 7(b) we plot the expected deviation for latitudes L_0, L_3, L_4, and L_5. We conclude that $\bar{D}$ for L_5 is almost the same as of L_0 for most of the reach distances. Thus L_5 seems to provide a very attractive trade-off given the reduction in the $|RS|$ and minimal excess wavelength-km penalty.

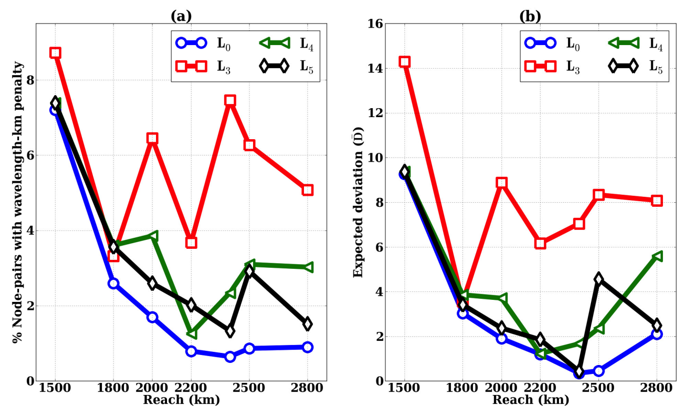

Fig. 7: (a) Percentage of node-pairs with wavelength-km penalty for zero latitude, L_3, L_4 and L_5. (b) Expected deviation ($\bar{D}$) for zero latitude, L_3, L_4 and L_5.

We define the *network cost* for a given latitude L_k, where $k \in \{1,2,3,4,5\}$ as $Cost_k = \sum_{\forall(i,j),i\neq j} Pr(i,j) \times Cost_p(i,j)$, where $Cost_p(i,j)$ is the cost of the circuit for a given node-pair (i,j) calculated using (4), with $c_r = 1000$ and $c_m = 1$. We define the expected cost deviation of various latitude rules from the min-cost CRLP (i.e., latitude is set to zeros and the paths are strictly min-cost) as:

$$\bar{D}_c = \frac{Cost_k - Cost_0}{Cost_0} \times 100, \ \ k \in \{1,2,3,4,5\} \quad (8)$$

From Fig. 8 we observe that L_2, L_4, and L_5 shows only small percentage in the cost-deviation compared to L_0.

Finally, we evaluate the Δ_{RS} as defined in Section V for supporting backup paths on the CONUS topology shown in Fig. 3. In Table II p_d is the percentage of node-pairs with a

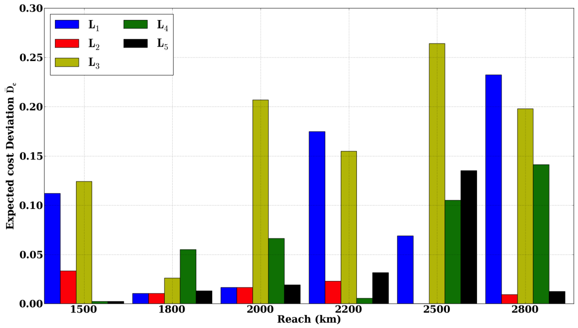

Fig. 8: Expected cost deviation ($\bar{D}_c$) for L_k where $k \in \{1, 2, 3, 4, 5\}$

valid disjoint path using RS_{CRLP} and Δ_{RS} is the number of additional regenerators required to provide backup paths to all node-pairs. In min-distance CRLP ($c_r = 0$ and $c_m = 1$), we find that using RS_{CRLP} all node-pairs have a valid disjoint path, i.e., $p_d = 100\%$ and hence $\Delta_{RS} = 0$ for most of the reach distances. For min-regeneration ($c_r = 1$ and $c_m = 0$) and min-cost ($c_r = 1000$ and $c_m = 0$) CRLP, $\Delta_{RS} \neq 0$.

Routing	Parameter	Reach (km)			
		1500	1800	2000	2500
Min-Distance	RS_{CRLP}	24	18	17	12
	p_d	100	100	97.55	100
	Δ_{RS}	0	0	1	0
Min-Regeneration	RS_{CRLP}	37	29	22	14
	p_d	88	87.96	87.56	83.27
	Δ_{RS}	3	4	4	4
Min-Cost	RS_{CRLP}	41	32	28	23
	p_d	97.37	96.61	95.06	97.55
	Δ_{RS}	2	2	2	1

TABLE II: Evaluation of Δ_{RS} for diverse routes on CONUS topology

VII. Conclusion

To plan pre-deployment of regenerators effectively in practical networks, we have introduced a new constrained routing regenerator location problem and presented a novel heuristic approach to address it. Unlike previous research work on regenerator placement, we pursue a holistic approach of minimizing overall network cost by considering a combination of number of regenerators used and wavelength-km of individual circuits, as well as the probability of a demand between each node-pair, and number of sites. We start with a basic heuristic and then present various enhancements which refine the trade-off between the number of sites and the cost of individual circuits. The heuristic also constructs a lower bound, thus letting us evaluate how closely we approach the optimal. Extensive simulations on large topologies show that the heuristic achieves near-optimal results. Our heuristic algorithm can be easily tuned for practical ROADM networks where instead of a fixed reach distance, a reach table is used to specify all reachable paths in the network (such tables are generated by vendor's planning tools using detailed network information). Finally our experiments indicate that a small additional number of regenerator sites allow survivable connections between most node-pairs. We further plan to extend this work to evaluate: (1) the usage of regenerators at each selected site for dynamic traffic; (2) complex requirements of disjointness on connection between multiple node-pairs.

Acknowledgment

We gratefully acknowledge many useful suggestions and comments from Dahai Xu, and Gil Zussman and support from the CIAN NSF ERC (subaward Y503160).

References

[1] B. G. Bathula *et al.*, "Cost optimization using regenerator site concentration and routing in ROADM networks," Columbia University and AT&T, Tech. Rep., July 2012. [Online]. Available: http://www.ee.columbia.edu/~bbathula/techrep/RegenCon.pdf

[2] A. Gerber *et al.*, "Traffic types and growth in backbone networks," in *OFC/NFOEC*, Mar. 2011, pp. 1–3.

[3] J. Simmons *et al.*, "On the value of wavelength-add/drop in WDM rings with uniform traffic," in *OFC/NFOEC*, Feb. 1998, pp. 361–362.

[4] W. Van Parys *et al.*, "Quantifying the benefits of selective wavelength regeneration in ultra long-haul WDM networks," in *OFC/NFOEC*, vol. 2, Mar. 2001, p. TuT4.

[5] M. D. Feuer *et al.*, *ROADMs and their system applications*, Optical Fiber Telecommunications V ed. Academic, 2008, pp. 293–343.

[6] A. Mahimkar *et al.*, "Outage detection and dynamic re-provisioning in GRIPhoN: A Globally Reconfigurable Intelligent Photonic Network," in *OFC/NFOEC*, Mar. 2012, p. NM2F.5.

[7] S. L. Woodward *et al.*, "Service Velocity: Rapid provisioning strategies in optical ROADM networks," *J. Opt. Commun. Netw.*, vol. 4, no. 2, pp. 92–98, Feb. 2012.

[8] X. Zhang *et al.*, "Bridge-and-roll demonstration in GRIPhoN (Globally Reconfigurable Intelligent Photonic Network)," in *OFC/NFOEC*, Mar. 2010, pp. 1–3.

[9] A. Chiu *et al.*, "Architectures and protocols for capacity efficient, highly dynamic and highly resilient core networks [Invited]," *J. Opt. Commun. Netw.*, vol. 4, no. 1, pp. 1–14, Jan 2012.

[10] M. D. Feuer *et al.*, "Simulations of a service velocity network employing regenerator site concentration," in *OFC/NFOEC*, Mar. 2012, p. NTu2J.5.

[11] M. Flammini *et al.*, "On the complexity of the regenerator placement problem in optical networks," in *Proc. of the 2st annual symposium on Parallelism in algorithms and architectures (SPAA)*, 2009, pp. 154–162.

[12] S. Chen *et al.*, "The generalized regenerator location problem," in *Proc. of International Network Optimization Conference*, 2009.

[13] S. Chen, I. Ljubi, and S. Raghavan, "The regenerator placement problem," *Networks*, vol. 55, no. 3, pp. 205–220, 2010.

[14] A. Duarte *et al.*, "Randomized heuristics for the regenerator location problem," *Optimization Online*, 2010. [Online]. Available: http://www.optimization-online.org/DB_HTML/2010/08/2706.html

[15] M. Flammini *et al.*, "On the complexity of the regenerator placement problem in optical networks," *IEEE/ACM Trans. on Netw.*, vol. 19, no. 2, pp. 498–511, Apr. 2011.

[16] G. Shen *et al.*, "Sparse placement of electronic switching nodes for low blocking in translucent optical networks," *J. Opt. Netw.*, vol. 1, no. 12, pp. 424–441, Dec. 2002.

[17] X. Yang *et al.*, "Sparse regeneration in translucent wavelength-routed optical networks: Architecture, network design and wavelength routing," *Photonic Network Communications*, vol. 10, pp. 39–53, 2005.

[18] B. G. Bathula *et al.*, "On concentrating regenerator sites in ROADM networks," in *OFC/NFOEC*, Mar. 2012, p. NW3F.6.

[19] M. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.

[20] S. Rai *et al.*, "On provisioning in all-optical networks: An impairment-aware approach," *IEEE/ACM Trans. on Netw.*, vol. 17, no. 6, pp. 1989 –2001, Dec. 2009.

[21] C. Gao *et al.*, "Survivable impairment-aware traffic grooming and regenerator placement with connection-level protection," *J. Opt. Commun. Netw.*, vol. 4, no. 3, pp. 259–270, Mar 2012.

[22] X. Yang *et al.*, "Survivable lightpath provisioning in WDM mesh networks under shared path protection and signal quality constraints," *J. of Light. Tech.*, vol. 23, no. 4, pp. 1556 – 1567, Apr. 2005.

[23] A. L. Chiu *et al.*, "Network design and architectures for highly dynamic next-generation IP-over-optical long distance networks," *J. of Light. Tech.*, vol. 27, no. 12, pp. 1878–1890, 2009.