Maximizing Reliability in WDM Networks through Lightpath Routing

Hyang-Won Lee, *Member, IEEE,* Kayi Lee, *Member, IEEE,* Eytan Modiano, *Fellow, IEEE*

***Abstract*—We study the reliability maximization problem in WDM networks with random link failures. Reliability in these networks is defined as the probability that the logical network is connected, and it is determined by the underlying lightpath routing, network topologies and the link failure probability. By introducing the notion of lexicographical ordering for lightpath routings, we characterize precise optimization criteria for maximum reliability in the low failure probability regime. Based on the optimization criteria, we develop lightpath routing algorithms that maximize the reliability, and logical topology augmentation algorithms for further improving reliability. We also study the reliability maximization problem in the high failure probability regime.**

I. Introduction

Modern communication networks are constructed using a layered approach, with one or more electronic layers (e.g., IP, ATM, SONET) built on top of an optical fiber network. The survivability of such networks under fiber failures largely depends on how the logical electronic topology is embedded onto the physical fiber topology. In the context of WDM networks, this is known as *lightpath routing*. However, finding a reliable lightpath routing is rather challenging because it must take into account the sharing of physical fibers by logical links and its impact on the connectivity of the logical topology. Hence, the survivability of a layered network is a complex function of logical topology, physical topology, lightpath routing, and link failure probability. In this paper, we study reliable layered network design assuming that physical links fail at random with some probability, where multiple links may fail simultaneously.

The probabilistic failure model represents a snapshot of a network where links fail and are repaired after a certain time as in many practical scenarios [1]. Hence, the link failure probability can be viewed as the average fraction of time that a link is in a failed state. This random failure model is somewhat general in that it can be used to model both networks with rare link failures as well as more frequent failures. It thus enables thorough understanding of network survivability in various failure regimes. For this reason, several works in the literature study survivable network design under the random failure model [1]–[4].

In the context of layered networks with random physical link failures, a natural survivability metric is the probability that given a lightpath routing, the logical topology remains connected; we call this probability the *cross-layer (network) reliability*. The cross-layer reliability reflects the survivability performance achieved by the layered network. Hence, it is desirable to design a layered network that maximizes the reliability. **Although the single-layer network design problem has been extensively studied [5]–[12], the layered network reliability problem remains largely unexplored. Existing work in the area [13]–[22] has mostly focused on finding a lightpath routing that survives a single physical link failure, rather than finding the one with maximum reliability. Our work in [32] was the first study to *maximize* the tolerance of such physical failures for a lightpath routing, and cross-layer reliability was introduced in [23] to generalize this notion. In particular, we extended the polynomial expression for single-layer network reliability to the layered setting, and developed approximation algorithms for reliability computation. We also demonstrated a positive correlation between the reliability and Min Cross Layer Cut (MCLC; The precise definition of MCLC is presented in Section II) in the low failure probability regime, and experimented with MCLC as the objective in our lightpath routing algorithm to approximate reliability maximization.**

Our goal is to fully characterize the structures that contribute to the reliability in a layered network. This gives us the precise optimization criterion for maximizing the reliability. Although optimizing the exact criterion is infeasible in practice, the insight allows us to develop a new objective that better approximates reliability maximization.

Typically, real-world networks experience very low link failure probabilities, and are designed accordingly. For instance, the failure probability of a 1000-mile cable in the Bellcore network is estimated to be about 0.006 [24]. However, in recent years there has been an increased concern about the impact of natural disasters or physical attacks on network survivability. Natural disasters, such as earthquakes and hurricanes or floods can lead to a large number of (possibly localized) link failures that cannot be survived by networks designed to deal only with isolated failures [25], [26]. Worse yet, a physical attack on the network by weapons of mass destruction, such as an Electromagnetic Pulse (EMP), can lead to widespread failures throughout large geographical areas [26]–[28]. Such an attack can have a disastrous effect on telecommunication links that rely on electronic components from fiber amplifiers to regenerators, switches and routers for their operation. Worse yet, such an attack is likely to

Hyang-Won Lee is with Konkuk University, Seoul, Republic of Korea. Kayi Lee is with Google, Inc., Cambridge, MA. Eytan Modiano is with the Massachusetts Institute of Technology, Cambridge, MA 02139 (e-mail: {hwlee, kylee, modiano}@mit.edu).

This work was supported by NSF grants CNS-0830961 and CNS-1017800, and by DTRA grants HDTRA1-07-1-0004 and HDTRA-09-1-0050.

disrupt the power grid [29], [30], which can in-turn lead to significant additional (cascading) failures of communication links, as was recently observed during a blackout event in Italy [31]. Thus, while typically one may expect extremely low failure probabilities, and design networks accordingly, such designs may not be robust to widespread failures that may result from a natural disaster or attack. Furthermore, it may be worthwhile to strengthen networks of critical importance so that they can withstand such scenarios.

Our primary focus in this work is on the low failure probability regime, as that is the regime that networks are typically designed for. However, to account for the increasing concerns with large scale failures, we also characterize network survivability in higher failure probability regimes. While such designs may not be applicable to most networks, they may prove valuable to the design of networks with stringent survivability requirements.

One of the major challenges in the area of cross-layer survivability is the inherent complexity of the problems. For example, in [32], we proved that the MCLC, a critical component in layered network reliability, is NP-hard to compute and approximate with within a $O(\log n)$ factor. Therefore, problems for maximizing cross-layer reliability is likely to be intractable. The common approach in existing lightpath routing algorithms involves finding the physical routes of all logical links jointly, typically by solving an ILP that captures the routing decision of all the logical links, which is often infeasible for large networks. In this paper, we consider a different approach by incrementally improving the layered network, one logical link at a time. Such an approach has the advantages over the existing algorithms:

1) **Scalability: Routing the logical links incrementally reduces the problem space significantly. As a result, it is more applicable to large networks.**
2) **Solution Quality: The incremental approach allows us to use a more sophisticated objective function that better approximates the cross-layer reliability. As a result, the lightpath routings given by the new algorithm result in much higher reliability than existing algorithms.**

We also apply a similar idea to a different setting where the logical topology can be augmented to improve reliability. We develop an augmentation algorithm to find a good placement of a new logical link, and observe that reliability can be improved significantly, especially when the augmentation increases the MCLC.

Our contributions can be summarized as follows:

- We show that in general the optimal lightpath routing depends on the link failure probability.
- We show that for given logical and physical topologies, if there exists a uniformly optimal lightpath routing, then any locally optimal lightpath routing is uniformly optimal.
- We develop a novel "lexicographical ordering" for lightpath routing and derive precise optimality conditions in both the low and high failure probability regimes.
- We develop lightpath rerouting algorithms for maximizing reliability in the low failure probability regime.
- We develop a logical topology augmentation algorithm for improving the reliability of a given layered network.

The rest of the paper is organized as follows: In Section II, we present the network model, and introduce the polynomial expression for the cross-layer reliability and important connectivity parameters related to reliability. In Section III, we study the properties of optimal lightpath routings in the low failure probability regime. In Section IV, we develop lightpath rerouting and logical topology augmentation algorithms for reliability maximization, and in Section V, we present extensive simulation results. In Section VI, we discuss the optimality conditions for maximum reliability in the high failure probability regime.

II. Model and Background

We consider a layered network $\mathcal{G}$ that consists of the logical topology $G_L = (V_L, E_L)$ built on top of the physical topology $G_P = (V_P, E_P)$ through a lightpath routing, where V and E are the set of nodes and links respectively. In the context of WDM networks, a logical link is called a *lightpath*, and each lightpath is routed over the physical topology. This *lightpath routing* is denoted by $f = [f_{ij}^{st}, (i,j) \in E_P, (s,t) \in E_L]$, where f_{ij}^{st} takes the value 1 if logical link (s,t) is routed over physical link (i,j), and 0 otherwise.

Each physical link fails independently with probability p[1]. This probabilistic failure model represents a snapshot of a network where links fail and are repaired according to some Markovian process. Hence, p represents the steady-state probability that a physical link is in a failed state. This model has been adopted by several previous works [1]–[4].

If a physical link (i,j) fails, all of the logical links (s,t) carried over (i,j) (i.e., (s,t) such that $f_{ij}^{st} = 1$) also fail. A set S of physical links is called a *cross-layer cut* if the failure of the links in S causes the logical network to be disconnected. We also define the *network state* as the subset S of physical links that failed. Hence, if S is a cross-layer cut, the network state S represents a *disconnected* network state. Otherwise, it is a *connected* state.

A. Failure Polynomial and Connectivity Parameters

Assume that there are m physical links, i.e., $|E_P| = m$. The probability associated with a network state S with exactly i physical link failures (i.e., $|S| = i$) is $p^i(1-p)^{m-i}$. Let N_i be the number of cross-layer cuts S with $|S| = i$, then the probability that the network is disconnected is simply the sum of the probabilities over all cross-layer cuts, i.e.,

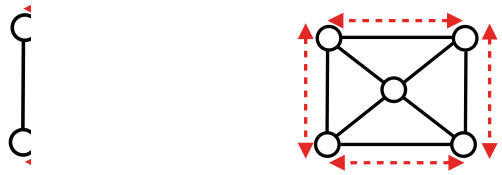
$$F(p) = \sum_{i=0}^{m} N_i p^i (1-p)^{m-i}. \quad (1)$$

Therefore, the failure probability of a multi-layer network can be expressed as a polynomial in p. The function $F(p)$ will be

[1]**Although we assume uniform link failure probability throughout the paper, our results can be readily extended to the case of non-uniform link failure probability by replacing each link with multiple links in series that fail with the same probability. See [23] for more details.**

called the *cross-layer failure polynomial* or simply the *failure polynomial*. The coefficients N_i's contain the information on the structure of a layered graph, determined by the underlying lightpath routing. Below we introduce some important coefficients related to connectivity.

Each N_i represents the number of cross-layer cuts of size i in the network. Define a *Min Cross Layer Cut (MCLC)* as a smallest set of physical links needed to disconnect the logical network. Denote by d the size of an MCLC, then d is the smallest i such that $N_i > 0$, meaning that the logical network will not be disconnected by fewer than d physical link failures. The MCLC is a generalization of single-layer min-cut to the multi-layer setting [32]. It was shown in [23] that maximizing

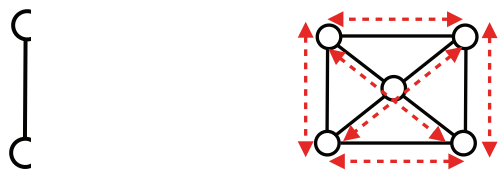


(b) After rerouting

Fig. 1. Example showing that lightpath rerouting can improve the reliability. Physical topology is solid line, logical topology is the rectangle formed by the 4 corner nodes and 4 edges, and lightpath routing is dashed line.

Although the MCLC criterion is useful for finding a lightpath routing with better reliability, it is not sufficient for fully characterizing reliable lightpath routings. For example, consider the two lightpath routings in Fig 1. The two lightpath routings have the same MCLC value of 2. However, for every value of p, the routing in Fig. 1(b) yields better reliability than the one in Fig. 1(a). This example shows that there are more precise conditions for optimal lightpath routings, beyond the MCLC maximization criterion. In Section III, we develop new optimization criteria that characterize in greater detail optimal lightpath routings in the low failure probability regime.

Furthermore, the routing in Fig. 1(b) can be obtained by



(b) Lightpath routing

Fig. 2. Example showing that the reliability can be further improved via logical topology augmentation: in (a), dashed lines are added lightpaths.

In addition to the lightpath rerouting approach, the new optimization criteria can also be used to further enhance the reliability in a different manner. In particular, we consider logical topology augmentation. For instance, suppose that two (diagonal) logical links are added to the logical topology in the example of Fig. 1 (see Fig. 2(a)). Fig. 2(b) is an example of routing the two new lightpaths. The new network has far better reliability than the old one in the low failure probability regime since the MCLC value has been raised from 2 to 3. This example shows that augmenting the logical topology can significantly improve the reliability. In Section IV-B, using the new optimization criteria, we study how to choose the new logical link that achieves maximum reliability improvement.

III. Properties of Optimal Lightpath Routings

We first study the properties of optimal lightpath routings. These properties will give insight on how routings should be designed for better reliability. Since the failure probability p is typically small in many practical scenarios, we mainly focus on the low failure probability regime. The properties of optimal lightpath routings for large p will be briefly discussed in Section VI.

A. Uniformly and Locally Optimal Routings

We start with a discussion of routings that are most reliable for all failure probabilities. The observations in this section will motivate a local (in p) optimization approach to the design of lightpath routing, which is relatively easy compared with an optimization over all the values of p. We begin with the following definition:

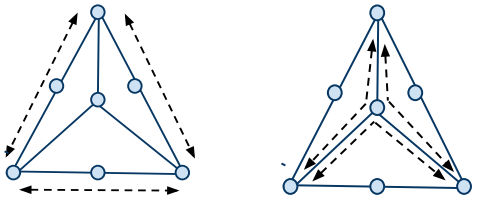
Definition 1: For given logical and physical topologies, a lightpath routing is said to be *uniformly optimal* if its reliability is greater than or equal to that of any other lightpath routing for every value of p.

Therefore, a uniformly optimal lightpath routing yields the best reliability for all $p \in [0, 1]$. Based on the failure polynomial of a lightpath routing, one can immediately develop a sufficient condition for a uniformly optimal lightpath routing:

Observation 1: Given a lightpath routing R, let N_i^R be the number of cross-layer cuts with size i. Then R is a uniformly optimal lightpath routing if, for any other lightpath routing $R^{'}$, $N_i^R \leq N_i^{R^{'}}$ for all $i \in \{0, \ldots, m\}$, where m is the number of physical links.

While it is desirable to design a uniformly optimal routing, such a routing does not always exist. Intuitively, for small p, only a small number of links are likely to fail simultaneously, and hence for better reliability it is important to remain connected after a small number of failures. In contrast, for large p, it is likely that a large number of links fail simultaneously, and thus it is important to withstand a large number of failures. These two objectives conflict because the former prefers disjoint lightpath routing whereas the latter prefers shortest lightpath routing.

For example, Fig. 3 shows two different lightpath routings. In Fig. 3(a), the logical links are routed over physically disjoint paths, and its reliability is given by $3(1-p)^4 - 2(1-p)^6$.



(a) Optimal Routing in Low Regime

(b) Optimal Routing in High Regime

Fig. 3. Example showing that optimal routings depend on the value of p. Physical topology is solid line, logical topology is the triangle formed by the 3 corner nodes and 3 edges, and lightpath routing is dashed line.

In contrast, in Fig. 3(b), every pair of logical links share a physical link, and its reliability is $(1-p)^3$. While disjoint path routing is considered to be more reliable, it is easy to see that in this example the disjoint routing has better reliability only for small values of p whereas for large p (e.g., $p > 0.7$) the non-disjoint routing is more reliable.

Since uniformly optimal lightpath routings are not always attainable, we are motivated to focus on *locally* optimal routings, where the probability regime of optimality is restricted to a subrange within $[0, 1]$. A locally optimal lightpath routing is defined as follows:

Definition 2: For given logical and physical topologies, a lightpath routing is said to be *locally optimal* if there exists $0 \leq a < b \leq 1$, such that its reliability is greater than or equal to that of any other lightpath routing for every value of $p \in [a, b]$. In addition, the interval $[a, b]$ is called the *optimality regime* for the lightpath routing.

Note that a uniformly optimal lightpath routing is also locally optimal with optimality regime $[0, 1]$. Theorem 1 below is a crucial result to this study; namely, it reveals a connection between local optimality and uniform optimality.

Theorem 1: Consider a pair of logical and physical topologies (G_L, G_P) for which there exists a uniformly optimal routing. Then, any locally optimal lightpath routing for (G_L, G_P) is also uniformly optimal.

Proof: Denote by $F^*(p)$ the failure polynomial of a uniformly optimal lightpath routing. By definition, $F^*(p)$ is no greater than any other failure polynomial for $p \in [0, 1]$. Consider a locally optimal lightpath routing L with optimality regime $[p_1, p_2]$, and let $F^L(p)$ be its failure polynomial.

The polynomial equation $F^L(p) - F^*(p) = 0$ has degree at most m and thus has at most m roots unless the polynomial $F^L(p) - F^*(p)$ is trivially zero. However, by the definitions of local optimality and uniform optimality, the equation has an infinite number of solutions over the interval $[p_1, p_2]$. Consequently, $F^L(p)$ is identical to $F^*(p)$, which implies that lightpath routing L is also uniformly optimal. ■

Motivated by this result, we study locally optimal lightpath routings. In particular, we develop the conditions for a lightpath routing to be optimal for the low failure probability regime (small p).

B. Low Failure Probability Regime

It is easy to see that in the failure polynomial, the terms corresponding to small cross-layer cuts dominate when p is small. Hence, for reliability maximization in the low failure probability regime, it is desirable to minimize the number of small cross-layer cuts. We use this intuition to derive the properties of optimal routings for small p. We begin with the following definition:

Definition 3: Consider two lightpath routings 1 and 2. Routing 1 is said to be *more reliable* than routing 2 *in the low failure probability regime* if there exists a positive number p_0 such that the reliability of routing 1 is higher than that of routing 2 for $0 < p < p_0$. A lightpath routing is said to be *locally optimal in the low failure probability regime* if it is more (or equally) reliable than any other routing in the low failure probability regime.

In the following, we characterize the impact of small cuts on the reliability. Let d_j be the size of the MCLC under routing $j(= 1, 2)$. Let N_i and M_i be the numbers of cross-layer cuts of size i under routings 1 and 2 respectively. We call the vector $N = [N_i, \forall i]$ the *cut vector*. The following is an example of cut vectors N and M with $d_1 = 4$ and $d_2 = 3$:

i	0	1	2	3	4	5	$\cdots$	m
N_i	0	0	0	0	20	26	$\cdots$	1
M_i	0	0	0	9	19	30	$\cdots$	1.

Using cut vectors of lightpath routings, we define *lexicographical ordering* as follows:

Definition 4: Routing 1 is lexicographically smaller than routing 2 if $N_d < M_d$ where d is the smallest i at which N_i and M_i differ.

Note that a lightpath routing with a larger MCLC size is lexicographically smaller by Definition 4. In the above example, we have $d = 3$ and $N_d < M_d$, hence routing 1 is lexicographically smaller. Therefore, if a lightpath routing is lexicographically smaller than another, it has fewer small cross-layer cuts and thus yields better reliability for small p.

Theorem 2: Given two lightpath routings 1 and 2 with cut vectors $[N_i | i = 0, \ldots, m]$ and $[M_i | i = 0, \ldots, m]$ respectively, where m is the number of physical links, if routing 1 is lexicographically smaller than routing 2, then routing 1 is more reliable than routing 2 in the low failure probability regime. In particular, let $d = \min_{0 \leq i \leq m} \{i : M_i \neq N_i\}$ be the index where the elements in the cut vectors first differ. Then, lightpath routing 1 is more reliable than routing 2 for $p < p_0 = \frac{(d+1)(M_d - N_d)}{2m\binom{m}{d}}$.

Proof: This is implied by Theorem 3, which will be discussed in Section III-C. ■

Clearly, Theorem 2 leads to a local optimality condition; that is, if a lightpath routing minimizes the cut vector lexicographically, then it is locally optimal in the low failure probability regime. An interesting case is when routing 1 has larger MCLC than routing 2 (as in the above example). In this case, routing 1 is lexicographically smaller than routing 2, and Theorem 2 implies the following corollary.

Corollary 1: If $d_1 > d_2$, then routing 1 is more reliable than routing 2 in the low failure probability regime.

Consequently, a lightpath routing with the maximum size MCLC yields the best reliability for small p. We note that the same result was shown in [23]. Similarly, routing 1 is also lexicographically smaller than routing 2 when they have the

same size of MCLC but routing 1 has fewer MCLCs. This leads to the following result:

Corollary 2: If $d_1 = d_2$ and $N_{d_1} < M_{d_2}$, then routing 1 is more reliable than routing 2 in the low probability regime.

The expression for p_0 given in Theorem 2 also provides some insight into how the difference of the cut vectors affects the guaranteed regime. For example, if d is small and $M_d - N_d$ is large, the guaranteed regime is larger. In other words, if one lightpath routing has fewer small cross-layer cuts than the other, it will achieve higher reliability for a larger range of p in the low probability regime.

Therefore, for reliability maximization in the low failure probability regime, it is desirable to maximize the size of the MCLC while minimizing the number of such MCLCs. This condition will be used to develop the algorithms in Section IV.

C. Extension of Optimal Probability Regimes

The expressions in Theorem 2 only consider the first element in the two cut vectors that are different. As a result, the guaranteed regime is rather conservative. For instance, the expression fails to capture the uniform optimality for a lightpath routing that satisfies the condition in Theorem 1. In this section, we will develop a more general expression for the regime bound that includes other elements in the cut vectors.

Consider two lightpath routings 1 and 2. Let $F_j(p)$ be the failure polynomial of routing j $(= 1, 2)$, and N_i's and M_i's be the coefficients in $F_1(p)$ and $F_2(p)$ respectively. Define the following vector of partial sums:

$$\overrightarrow{\mathbf{N}} = \left[\sum_{i=0}^{k} N_i | k = 0, ..., m \right]$$

The vector $\overrightarrow{\mathbf{M}}$ is defined similarly. Note that the i-th element $\overrightarrow{N}_i$ of vector $\overrightarrow{\mathbf{N}}$ is the total number of cross-layer cuts of size at most i. We first extend the definition of *lexicographical ordering* as follows:

Definition 5: Lightpath routing 1 is said to be k-lexicographically smaller than lightpath routing 2 if

$$k = \max \left\{ j : \overrightarrow{N}_i \le \overrightarrow{M}_i, \quad \forall i < d + j \right\} \text{ and } k \ge 1,$$

where d is the position of the first element where the two cut vectors differ.

Therefore, a lightpath routing is lexicographically smaller (in the original sense) if and only if it is k-lexicographically smaller for some $k \ge 1$. The k-lexicographical ordering thus compares two lightpath routings based on structures beyond the smallest cuts, making it possible to establish a larger optimality regime. Roughly speaking, the value of k reflects the degree of dominance of a lightpath routing in the low probability regime: a k-lexicographically smaller lightpath routing means that it has fewer "small" cuts, where the definition for "small" is broader if k is larger.

It is obvious that when $p \le 0.5$, the failure probability of a cross-layer cut is a non-increasing function of the cut size, because $p^i(1-p)^{m-i} \ge p^{i+1}(1-p)^{m-(i+1)}$ for $p \le 0.5$. Suppose that routing 1 has smaller total number of cuts of size up to i than routing 2, i.e., $\overrightarrow{N}_i \le \overrightarrow{M}_i$. To compare cross-layer cuts of size at most $i+1$, suppose further that the relative increment $N_{i+1} - M_{i+1}$ in the number of larger cuts does not exceed the surplus $\overrightarrow{M}_i - \overrightarrow{N}_i$ from smaller cuts, i.e., $\overrightarrow{N}_{i+1} \le \overrightarrow{M}_{i+1}$. Then, with respect to cut size at most $i+1$, routing 1 will have smaller failure probability than routing 2, provided that the same was true for cut size up to i. This observation leads to the following theorem on the relationship between lexicographical ordering and probability regime.

Theorem 3: Given two vectors $\mathbf{N}=[N_i | i = 0, \ldots, m]$ and $\mathbf{M}=[M_i | i = 0, \ldots, m]$, let $F_1(p) = \sum_{i=0}^{n} N_i p^i (1-p)^{m-i}$ and $F_2(p) = \sum_{i=0}^{n} M_i p^i (1-p)^{m-i}$. For any j, let $\overrightarrow{\Delta}_j = \sum_{i=0}^{j} (M_i - N_i)$ and $\overrightarrow{\delta}_j = \max_{j+1 \le i \le m} \left\{ \frac{N_i - M_i}{\binom{m}{i}} \right\}$. If the vector $\mathbf{N}$ is k-lexicographically smaller than $\mathbf{M}$, then

$$F_1(p) \le F_2(p) \text{ for } p \le p_0^l = \min \left\{ 0.5, \max_{d \le j \le d+k-1} B_j \right\}$$

where $d = \min \{ i : N_i < M_i \}$ and

$$B_j = \begin{cases} 0.5, & \text{if } j = m \\ \frac{1}{\frac{m}{j+1} + \overrightarrow{\delta}_j \binom{m}{j+1} / \overrightarrow{\Delta}_j}, & \text{otherwise.} \end{cases}$$

Proof: See Appendix A. ∎

Therefore, the probability regime bound p_0^l in Theorem 3 is a non-decreasing function of k, which means that a lightpath routing with smaller number of cuts over a larger size range will be more reliable over a larger probability regime. This is consistent with the conclusion in Section III-B, that the lightpath routing design should minimize the lexicographical ordering of the cut vector.

Theorem 2 follows from Theorem 3. For a lexicographically smaller lightpath routing, the term B_d is given by:

$$\begin{aligned} \frac{1}{\frac{m}{d+1} + \overrightarrow{\delta}_d \binom{m}{d+1} / \overrightarrow{\Delta}_d} &= \frac{1}{\frac{m}{d+1} + \overrightarrow{\delta}_d \binom{m}{d+1} / (M_d - N_d)} \\ &\ge \frac{(d+1)(M_d - N_d)}{m(M_d - N_d) + (d+1)\binom{m}{d+1}}, \\ &\ge \frac{(d+1)(M_d - N_d)}{m\binom{m}{d} + (m-d)\binom{m}{d}} \\ &\ge \frac{(d+1)(M_d - N_d)}{2m\binom{m}{d}}, \end{aligned}$$

where the first inequality is due to $\overrightarrow{\delta}_d \le 1$.

An interesting special case is when $d + k - 1 = m$, that is, $\overrightarrow{M}_j \ge \overrightarrow{N}_j$ for all $j = 0, \ldots, m$. In that case, the term $B_{d+k-1} = B_m = 0.5$, implying that the optimality regime is $[0, 0.5]$. We summarize this as the following corollary:

Corollary 3: If $\overrightarrow{N}_j \le \overrightarrow{M}_j$ for all $j = 0, \ldots, m$, then lightpath routing 1 is at least as reliable as lightpath routing 2 for $p \le 0.5$, i.e., $F_1(p) \le F_2(p)$ for $p \le 0.5$.

Note that the condition in Corollary 3 requires every partial sum in the vector $\mathbf{M}$ to be at least the corresponding partial sum in the vector $\mathbf{N}$, which is a stronger condition than the lexicographic comparison in Theorem 2. This stronger condition allows the better optimality regime to be established in Corollary 3.

IV. Maximizing Reliability by Improving Lightpath Routing and Logical Connectivity

In this section, we explore ways to improve the reliability of a layered network. Typically, the physical topology is static and difficult to change. Therefore, the reliability of a layered network can be improved by one of two ways: (i) improving the lightpath routing, or (ii) improving the logical topology.

We have shown in Section III that when physical link failures are rare, the lightpath routing that minimizes the lexicographical ordering will maximize the reliability. This new observation gives us an exact optimization criterion for designing reliable layered networks.

As discussed in Section I, the traditional approach of jointly routing all logical links is often too complex, which makes it infeasible for larger networks. This motivates the incremental approach introduced in this section, where the layered network is improved one logical link at a time. This significantly reduces the problem space and allows us to use a more sophisticated objective function based on the optimziation criterion we studied in Section III.

Within this context, we study two optimization problems that are fundamental to improving the lightpath routing and logical connectivity:

1) **Lightpath Rerouting**: Given the physical, logical topologies and a lightpath routing, find a logical link to reroute, such that the resulting reliability is maximized.
2) **Logical Topology Augmentation**: Given the physical, logical topologies and a lightpath routing, find a pair of logical nodes, as well as a physical path between the nodes, such that the addition of the corresponding logical link will provide maximum reliability improvement.

The above two problems are basic building blocks for designing reliable layered networks. For example, given an existing layered network, we can iteratively reroute existing lightpaths in the network until no further improvement is possible (e.g. Figure 4). **Hence, given the physical and logical topologies, the iterative rerouting algorithm can be described as follows:**

1) **Generate an arbitrary initial lightpath routing.**
2) **Reroute a logical link using ILP/approximation algorithm introduced in Section IV.**
3) **Repeat Step 2 until no further improvement can be made by rerouting a single lightpath.**

Similarly, if it is feasible to add new logical links, we can iteratively augment the logical topology to further improve the reliability; and studying the Logical Topology Augmentation problem allows us to select such new logical links effectively. These iterative rerouting and augmentation algorithms will be used for performance evaluation in Section V.

In this section, we present algorithms for the rerouting and augmentation problems. In the next section, we will evaluate the effectiveness of rerouting and augmentation on improving cross-layer reliability.

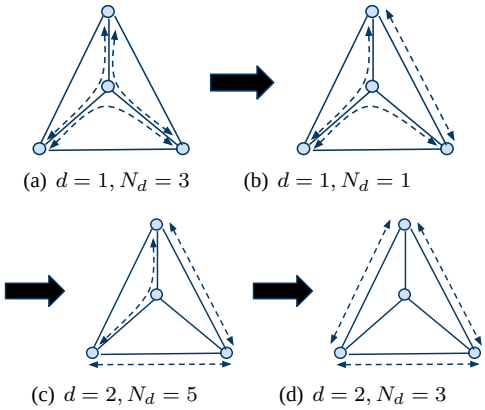


Fig. 4. Improving reliability via lightpath rerouting. The physical topology is in solid lines, and the lightpath routing of the logical topology is in dashed lines. The MCLC value and the number of MCLCs in the lightpath routings are denoted by d and N_d.

A. Lightpath Rerouting

Given a layered network and its lightpath routing, the objective of the *Lightpath Rerouting* problem is to find the best way to reroute a lightpath, so that the reliability improvement is maximized. Recall that with low link failure probability, the reliability of a network is maximized when the lexicographical ordering of its cut vector is minimized. Therefore, the most effective reroute should maximize the MCLC of the resulting lightpath routing, and also minimize the number of MCLCs.

In the following sections, we first analyze the effect of rerouting a lightpath and characterize conditions where such a rerouting is beneficial. This provides the groundwork for our rerouting algorithms. Based on these observations, we develop an ILP to find the optimal lightpath to reroute. Next, we propose an approximation algorithm that computes a near-optimal solution in much shorter time. This gives us a scalable algorithm that can be used for designing large layered networks.

1) Effects of Rerouting a Lightpath: Let d be the size of the MCLC under the initial routing. When the physical route of a logical link changes, some cross-layer cuts will be converted into non-cuts, and some non-cuts will be converted into cross-layer cuts. In the low failure probability regime, the reliability will be improved by the rerouting if the following is true:

1) The conversion of cross-layer cuts with size d to non-cuts outnumbers the conversion in the opposite direction.
2) The MCLC value does not decrease.

Therefore, we can formulate the lightpath rerouting as an optimization problem to maximize the reduction in the number of MCLCs, subject to the constraint that no non-cuts of size smaller than d is converted to cross-layer cuts. The exact conditions for the conversion between cuts and non-cuts are described as follows, which will be used as the basis of the ILP formulation as well as the approximation algorithm.

Given the physical topology $G_P = (V_P, E_P)$ and the logical topology $G_L = (V_L, E_L)$, we model a lightpath routing as a set of binary constants $\left\{f_{ij}^{st}\right\}$, where $f_{ij}^{st} = 1$ if and only if logical link (s,t) uses physical link (i,j) in the lightpath routing. For a given set of physical links S, we define the *logical residual graph* for S, denoted as G_L^S, to be

$\left\{(s,t) \in E_L : \sum_{(i,j)\in S} f_{ij}^{st} = 0\right\}$. In other words, the residual graph consists of logical links that use none of the physical links in S. By definition, the set S is a cross-layer cut if and only if its logical residual graph is disconnected. Given a cross-layer cut S, it is called a *k-way cross-layer cut* if its logical residual graph has k connected components. In addition, given a cross-layer non-cut T for a lightpath routing, we call a logical link (s,t) *critical* to T if (s,t) is a cut edge of the residual graph G_L^T, that is, it is an edge in G_L^T whose removal will disconnect the residual graph.

The following theorems describe the conditions for a lightpath rerouting that results in conversions between cross-layer cuts and non-cuts. The proofs can be found in [33].

Theorem 4: Let S be a cross-layer cut for a lightpath routing. Rerouting logical link (s,t) from physical path P_1 to P_2 turns S into a non-cut if and only if the following conditions are true:

1) S is a 2-way cross-layer cut.
2) s and t are disconnected in the residual graph for S.
3) P_2 does not use any physical links in S.

Proof: Suppose all the above conditions are true. Since the new route P_2 does not use any physical links in S, the logical link (s,t) will be in the logical residual graph for S under the new lightpath routing. Other logical links that are in the original residual graph will remain, because none of their physical routes have changed. Therefore, the residual graph will become connected now that (s,t) is added to it, which implies S becomes a non-cut. It can be easily verified that the residual graph will remain disconnected if any of the above conditions do not hold. ■

Theorem 5: Let T be a cross-layer non-cut for a lightpath routing. Rerouting logical link (s,t) from physical path P_1 to P_2 turns T into a cross-layer cut if and only if the following conditions are true:

1) (s,t) is critical to T.
2) P_2 uses some physical link in T.

Proof: Suppose both conditions are true. Since P_2 uses some physical fiber in T, the logical link will be removed from the residual graph for T under the new lightpath routing. Since (s,t) is critical to the non-cut T, its removal will disconnect the residual graph, which means that T will become a cross-layer cut. It can be easily verified that the residual graph will remain connected if any of the two conditions do not hold. ■

Therefore, the optimal rerouting should maximize the number of cross-layer cuts satisfying Theorem 4 and minimize the number of non-cuts satisfying Theorem 5. However, it is also important to ensure that none of the non-cuts with size smaller than d is converted to cross-layer cuts by the rerouting, since otherwise the MCLC value will decrease. The following theorem states that only non-cuts with size at least $d-1$ can be converted into a cross-layer cut by rerouting a single lightpath.

Theorem 6: Let d be the Min Cross Layer Cut value of a lightpath routing and let $\mathcal{NC}$ be the set of cross-layer non-cuts that can be converted into cross-layer cuts by rerouting a single logical link. Then $|T| \geq d-1$ for all $T \in \mathcal{NC}$.

Proof: Suppose $\mathcal{NC}$ contains a convertible non-cut S with size less than $d-1$. Since S is convertible, there exists a logical link (s,t) that is critical to S. Now let l be a fiber used by (s,t), then the fiber set $S \cup \{l\}$ would disconnect the logical residual graph and is therefore a cross-layer cut. However, such a set contains at most $d-1$ fibers, contradicting that d is the Min Cross Layer Cut. ■

Therefore, when rerouting a lightpath, we need to make sure that none of the non-cuts with size $d-1$ get converted into cuts in order to prevent the MCLC value from decreasing. Based on these observations, we next develop an ILP for the lightpath rerouting problem.

2) ILP for Lightpath Rerouting: For the given lightpath routing, let d be the MCLC value, and let $\mathcal{C}_d, \mathcal{NC}_d$ and $\mathcal{NC}_{d-1}$ be the sets of 2-way cross-layer cuts with size d, non-cuts with size d, and non-cuts with size $d-1$ respectively. The lightpath rerouting problem can be formulated as an ILP that finds the logical link, and its new physical route, that maximizes the net reduction in MCLCs. **In other words, the optimal reroute should result in the minimum number of cross-layer cuts with size d, without creating any cross-layers cuts with size $d-1$.**

The ILP can be considered as a path selection problem on an auxiliary graph $G_P^{'} = (V_P^{'}, E_P^{'})$, where $V_P^{'} = V_P \cup \{u, v\}$, with u and v being the additional source and sink nodes in the auxiliary graph; and $E_P^{'} = E_P \cup \{(u,x), (x,v) : x \in V_P\}$. Figure 5 illustrates the construction of the auxiliary graph.

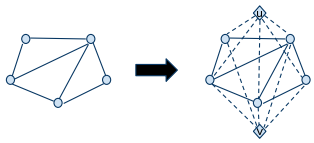


Fig. 5. Construction of the auxiliary graph for the ILP. u and v are the additional source and sink nodes, and the dashed lines are the additional links in the auxiliary graph.

We first define the following variables and parameters:

1) Variables:
   - $\{g_{st} : (s,t) \in E_L\}$: 1 if logical link (s,t) is rerouted, and 0 otherwise.
   - $\left\{f_{ij} : (i,j) \in E_P^{'}\right\}$: Flow variables describing a path in $G_P^{'}$ from node u to node v.
   - $\{y^c : c \in \mathcal{C}_d\}$: 1 if the cross-layer cut c is converted into a non-cut by the lightpath rerouting, and 0 otherwise.
   - $\{z^c : c \in \mathcal{NC}_d\}$: 1 if the non-cut c is converted into a cross-layer cut by the lightpath rerouting, and 0 otherwise.
2) Parameters:
   - $\{h_{st}^c : c \in \mathcal{C}_d, (s,t) \in E_L\}$: 1 if logical nodes s and t are disconnected by the 2-way cut c, and 0 otherwise.
   - $\{q_{st}^c : c \in \mathcal{NC}_d \cup \mathcal{NC}_{d-1}, (s,t) \in E_L\}$: 1 if logical link (s,t) is critical to the non-cut c, and 0 otherwise.
   - $\left\{l_{ij}^c : \forall c \in \mathcal{C}_d \cup \mathcal{NC}_d \cup \mathcal{NC}_{d-1}, (i,j) \in E_P\right\}$: 1 if physical link (i,j) is in set c, and 0 otherwise.

The lightpath rerouting can be formulated as follows:

$$\text{REROUTE}: \quad \text{Maximize} \sum_{c\in\mathcal{C}_d} y^c - \sum_{c\in\mathcal{NC}_d} z^c, \quad \text{subject to:}$$

$$g_{st} \le (f_{us} + f_{tv})/2, \quad \forall (s,t) \in E_L \quad (2)$$

$$\sum_{(s,t)\in E_L} g_{st} = 1 \quad (3)$$

$$l_{ij}^c f_{ij} + \sum_{(s,t)\in E_L} q_{st}^c g_{st} \le 1, \quad \forall c \in \mathcal{NC}_{d-1}, (i,j) \in E_P \quad (4)$$

$$l_{ij}^c f_{ij} + \sum_{(s,t)\in E_L} q_{st}^c g_{st} \le z^c + 1, \forall c \in \mathcal{NC}_d, (i,j) \in E_P \quad (5)$$

$$y^c \le \sum_{(s,t)\in E_L} h_{st}^c g_{st}, \quad \forall c \in \mathcal{C}_d \quad (6)$$

$$y^c \le 1 - l_{ij}^c f_{ij}, \quad \forall (i,j) \in E_P, \forall c \in \mathcal{C}_d \quad (7)$$

$$\{(i,j) : f_{ij} = 1\} \text{ forms an } (u,v)\text{-path in } G_P^{'} \quad (8)$$

$$f_{ij}, g_{st} \in \{0,1\}, 0 \le y^c, z^c \le 1$$

The formulation can be interpreted as a path selection problem on the auxiliary graph $G_P^{'}$. Constraint (8), which requires that the variables f_{ij} describe a path from u to v, can be expressed by the standard flow conservation constraints. As a result, in a feasible solution to the formulation, the variables f_{ij} represent a path $u \to s \rightsquigarrow t \to v$, which corresponds to the new physical route for the logical link (s,t) after the rerouting.

Constraint (2) ensures that g_{st} can be set to 1 only if f_{ij} represents the path $u \to s \rightsquigarrow t \to v$, and Constraint (3) makes sure that the chosen (s,t) is indeed a logical link in E_L. Therefore, exactly one logical link (s,t) can have $g_{st} = 1$, and a feasible solution to this ILP corresponds to a rerouting of the logical link.

In Constraint (4), the two terms correspond to the conditions in Theorem 6. The constraint makes sure that at most one of the conditions is satisfied, thereby disallowing the non-cuts of size $d-1$ to be converted into a cross-layer cut. Similarly, Constraint (5) makes sure $z^c = 1$ for any non-cut $c \in \mathcal{NC}_d$ that is converted into a cut by the rerouting.

Finally, Constraints (6) and (7) describe conditions 2) and 3) of Theorem 4 respectively. Therefore, y^c can be 1 only if both conditions are satisfied. Since c also satisfies condition 1) by definition of $\mathcal{C}_d$, this implies that cross-layer cut c is converted into a non-cut when $y^c = 1$.

Since the objective is to maximize y^c and minimize z^c, in an optimal solution $y^c = 1$ if and only if cross-layer cut c is converted into a non-cut, and $z^c = 1$ if and only if non-cut c is converted into a cross-layer cut. As a result, the objective function reflects the net reduction in the number of MCLCs.

Finally, note that the variables y^c and z^c will take on binary values in an optimal solution even if they are not constrained to be integral. This observation helps to reduce the number of binary variables in the formulation.

The ILP REROUTE **approximates the lexicographical ordering minimization by minimizing the number of MCLCs in the network. It can be extended to consider cross-layer cuts of size larger than d, thus achieving a better approximation. In this case, the set of cross layer cuts and non-cuts $\mathcal{C}_d$ and $\mathcal{NC}_d$ will be replaced by sets that include the cut and non-cuts up to size $k > d$, denoted as $\mathcal{C}_{\le k}$ and $\mathcal{NC}_{\le k}$. The objective function will be changed to**

$$Maximize \quad \sum_{c\in\mathcal{C}_{\le k}} y^c w^c - \sum_{c\in\mathcal{NC}_{\le k}} z^c w^c, \quad (9)$$

where w^c is a weight constant assigned to each cut c so that a smaller cut will have weight that dominates cuts of larger size. In particular, if k is set to $|E_P|$, the extended ILP will return the optimal solution that minimizes the lexicographical ordering. However, such a formulation will contain an exponential number of variables y^c and z^c, and is generally not feasible for practical use. Therefore, in the rest of the paper, we will focus on the problem of minimizing the number of MCLCs, though the techniques discussed in this paper are also be applicable to the more general setting.

3) Approximation Algorithm for Lightpath Rerouting: For larger networks, however, solving the rerouting ILP may still be infeasible. Therefore, in this section, we present an approximation algorithm for the rerouting problem that provides near-optimal solutions within a much shorter time.

We focus on the following question: Given the lightpath routing, and a logical link (s,t), what is the best way to reroute (s,t) assuming the routes for all other logical links are fixed? A solution to this problem will allow us to solve the lightpath rerouting problem, since we can run the algorithm once for each logical link, and return the best solution.

Similar to the previous section, let $\mathcal{C}_d, \mathcal{NC}_d$ and $\mathcal{NC}_{d-1}$ be the set of cross-layer cuts of size d, non-cuts of size d and non-cuts of size $d-1$ respectively. Now suppose Q is a new physical route for logical link (s,t). Let $\mathcal{NC}_d^{st}$ and $\mathcal{NC}_{d-1}^{st}$ be the subsets of $\mathcal{NC}_d$ and $\mathcal{NC}_{d-1}$ that satisfy condition (1) of Theorem 5. These two sets represent the non-cuts that can potentially be converted into a cut by rerouting (s,t). It immediately follows that any (s,t) path that uses a physical link in $\cup_{T\in\mathcal{NC}_{d-1}^{st}} T$ will create a cross-layer cut with size $d-1$, which should be forbidden for the new physical route. In addition, for any physical link (i,j), the set $\mathcal{L}_{ij}^{\mathcal{NC}} = \left\{T \in \mathcal{NC}_d^{st} : (i,j) \in T\right\}$ represents the non-cuts with size d that will be converted into cross-layer cuts if the new route Q contains the physical link (i,j).

Similarly, let $\mathcal{C}_d^{st} \subseteq \mathcal{C}_d$ be the set of cross-layer cuts that do not satisfy conditions (1) or (2) of Theorem 4. This represents the set that will continue to be cross-layer cuts regardless of the new physical route Q for (s,t). In addition, for each $(i,j) \in E_P$, the cross-layer cuts in the set $\mathcal{L}_{ij}^{\mathcal{C}} = \{S \in \mathcal{C}_d - \mathcal{C}_d^{st} : (i,j) \in S\}$ will also continue to be cross-layer cuts if the new route Q contains the physical link (i,j), as they do not satisfy condition 3) of Theorem 4.

Now, for each physical link (i,j), let $\mathcal{L}_{ij} = \mathcal{L}_{ij}^{\mathcal{C}} \cup \mathcal{L}_{ij}^{\mathcal{NC}}$. If a physical link (i,j) is used by the logical link (s,t) in the new route Q, it will cause the set $\mathcal{L}_{ij} \cup \mathcal{C}_d^{st}$ to become cross-layer cuts. Since every set of physical links in $\mathcal{C}_d^{st}$ will be cross-layer cuts regardless of the physical route taken by (s,t), the lightpath rerouting problem for logical link (s,t) can be formulated as choosing the (s,t)-path Q in $G_P^{'} = (V_P, E_P - \cup_{T\in\mathcal{NC}_{d-1}^{st}} T)$ that minimizes $|\mathcal{L}(Q)| = |\cup_{(i,j)\in Q}$

$\mathcal{L}_{ij}|$. Although this is an instance of the NP-Hard *Minimum Color Path* [34] problem, a simple d-approximation algorithm exists, as described below:

Algorithm 1 REROUTE_SP(s,t)

1: Construct a weighted graph on $G_P^{'} = (V_P, E_P - \cup_{T \in \mathcal{NC}_{d-1}^{st}} T)$, where each edge (i,j) is assigned with weight $w(i,j) = |\mathcal{L}_{ij}|$.
2: Run Dijkstra's algorithm to find the shortest (s,t)-path in the weighted graph.

Theorem 7: Let Q^* be the optimal physical route for (s,t) that results in the minimum number of MCLCs, and let Q^{SP} be the new route for (s,t) returned by REROUTE_SP. For any (s,t)-path Q, let $N_d(Q)$ be the number of cross-layer cuts with size d after rerouting (s,t) with Q, where d is the size of the MCLC. Then $N_d(Q^{\mathsf{SP}}) \leq d \cdot N_d(Q^*)$.

Proof: See Appendix B. ∎

Therefore, the number of cross-layer cuts of size d given by REROUTE_SP is at most d times the optimal reroute. Note that if the optimal new route for (s,t) eliminates every MCLC of size d, i.e., $N_d(Q^*) = 0$, the approximation algorithm will find a new route that achieves that as well. We state this observation as the following corollary.

Corollary 4: REROUTE_SP(s,t) will return a new route for (s,t) that increases the size of MCLC of the layered network, if such a new route exists.

We can extend algorithm REROUTE_SP, which is based on the Dijkstra's shortest path algorithm, by using the k-shortest path algorithm [35] to successively compute the next shortest path in $G_P^{'}$ and keep track of the path Q with the minimum value of $|\mathcal{L}(Q)|$. The value k reflects a tradeoff between running time and quality of the solution. As we will see in Section V, by picking a good value of k, we can obtain a lightpath routing within a much shorter time than solving the ILP without sacrificing much in solution quality.

A Note on Complexity: **The sets $\mathcal{C}_d$ and $\mathcal{NC}_d$ can be constructed by enumerating all the $\binom{|E_P|}{d}$ subsets of physical links and each of them can be classified as a cut or non-cut in $O(|E|_L)$ time by running a breath-first search on the logical topology. Similarly, for each subset $S \in \mathcal{C}_d \cup \mathcal{NC}_d$, we can decide whether each of its member (i,j) is in $\mathcal{L}_{ij}$ and $\mathcal{NC}_{d-1}^{st}$ by breath-first search. Therefore, the time to compute all $\mathcal{L}_{ij}$ is $O(\binom{|E_P|}{d}(|E_L| + d)) = O(|E_P|^d|E_L|)$. Overall, the time complexity to construct the graph $G_P^{'}$ is $O(|E_P|^d|E_L|)$. The k-shortest path algorithm on $G_P^{'}$ can be run in $O(k|V_P|(|E_P| + |V_P|\log|V_P|))$ time [35]. Therefore, the overall time complexity of REROUTE_SP(s,t) is $O(|E_P|^d|E_L| + k|V_P|(|E_P| + |V_P|\log|V_P|))$.**

B. Logical Topology Augmentation

The *Logical Topology Augmentation* problem involves finding the best way to augment the logical topology with a single logical link, in order to maximize the reliability improvement. Even though the augmentation problem has been extensively studied for single-layer networks, [36]–[40], this has not been studied before in the context of multi-layer networks. In addition to the *placement* aspect of finding the end points for the new link as for the single-layer networks, there is also the *routing* aspect for the layered networks. This adds a new dimension of complexity to the augmentation problem.

As it turns out, the insights from our study of the lightpath rerouting problem are largely applicable to the logical topology augmentation problem. In the following sections, we will first discuss the similarity between the augmentation problem and the lightpath rerouting problem, and then develop a similar ILP formulation and approximation algorithm.

1) Effects of a Single-Link Augmentation: Similar to the rerouting problem, the new logical link chosen by the augmentation algorithm should maximize the reduction in the number of MCLCs. However, unlike rerouting, adding a new link never converts a non-cut into a cross-layer cut. Therefore, in augmentation we only need to consider the effect of the new logical link on the existing cross-layer cuts.

Suppose that an initial lightpath routing is given for the physical topology $G_P = (V_P, E_P)$ and the logical topology $G_L = (V_L, E_L)$. Let d be the size of the MCLC under the initial routing. Let G_L^S be the logical residual graph for any cross-layer cut S, that is, the logical subgraph in which the logical links do not use any physical links in S. The following theorem characterizes the effect of a single-link augmentation: The proof can be found in [33].

Theorem 8: Let S be a cross-layer cut for a lightpath routing. Augmenting the network with a new logical link (s,t) over physical route P converts a cross-layer cut S into a non-cut if and only if:

1) S is a 2-way cross-layer cut.
2) s and t are disconnected in the residual graph for S.
3) P does not use any physical links in S.

Proof: The proof is the same as Theorem 4. The new logical link will make the residual graph connected if and only if the above conditions are true. ∎

Note that the characterizations for augmentation (Theorem 8) and rerouting (Theorems 4 and 5) differ only in that the conditions in Theorem 5 are no longer applicable to augmentation, because augmentation never converts any non-cut into a cross-layer cut. Therefore, we can revise the ILP REROUTE accordingly to formulate an ILP for the augmentation problem.

2) ILP for Logical Topology Augmentation: We will revise the ILP REROUTE presented in Section IV to develop the ILP for the augmentation problem. In REROUTE, the variables $\{z^c = 1 : c \in \mathcal{NC}_d\}$ correspond to the set of non-cuts that will be converted into cuts by the rerouting, and Constraints (4) and (5) describe the conditions for such conversion. As previously discussed, such conversion is not applicable in augmentation and therefore these variables and constraints can be removed from the ILP. In addition, unlike rerouting where we choose from the set of existing logical links, in augmentation we can pick any two logical nodes for the new logical link. Therefore, we will replace the variable set $\{g_{st} : (s,t) \in E_L\}$ in REROUTE by $\{g_{st} : (s,t) \in V_L \times V_L\}$ and remove Constraint (3). This gives us the following ILP

for augmentation:

$$\text{AUGMENT}: \quad \text{Maximize} \sum_{c\in\mathcal{C}_d} y^c, \quad \text{subject to:}$$

$$g_{st} \le (f_{us} + f_{tv})/2, \quad \forall (s,t) \in V_L \times V_L \tag{10}$$

$$y^c \le \sum_{(s,t)\in V_L\times V_L} h^c_{st} g_{st}, \quad \forall c \in \mathcal{C}_d \tag{11}$$

$$y^c \le 1 - l^c_{ij} f_{ij}, \quad \forall (i,j) \in E_P, \forall c \in \mathcal{C}_d \tag{12}$$

$$\{(i,j) : f_{ij} = 1\} \text{ forms an } (u,v)\text{-path in } G'_P \tag{13}$$

$$f_{ij}, g_{st} \in \{0,1\}, 0 \le y^c \le 1$$

Similar to the interpretation of REROUTE, in a feasible solution to AUGMENT, the variables f_{ij} represent a path $u \to s \rightsquigarrow t \to v$, as described by Constraint (13). This corresponds to the new logical link to be added, along with its physical route. Constraint (10) ensures that $g_{st} = 1$ if and only if (s,t) is the new logical link selected. Constraints (11) and (12) describe the conditions in Theorem 8. In particular, the variable y^c describes whether the cross-layer cut c is converted into non-cut by the augmentation. Therefore, the ILP maximizes the number of such conversions, which translates to maximizing the improvement in reliability.

3) Approximation Algorithm For Logical Topology Augmentation: One can also design an approximation algorithm similar to REROUTE_SP introduced in Section IV-A3 for the logical topology augmentation problem. We will again focus on the following question: Given a layered network, and a new logical link (s,t), find the physical route for (s,t) such that the resulting number of cross-layer cuts of size d is minimized. We can then apply the algorithm for this problem for every possible pair of logical nodes s and t, to find out the new logical link that would result in the maximum reliability improvement.

Let d be the size of the MCLC of the layered network and $\mathcal{C}^{st}_d$ be the set of 2-way cross-layer cuts of size d that separate the logical nodes s and t. Then by Theorem 8, the set $\mathcal{L}_{ij} = \{S \in \mathcal{C}^{st}_d : (i,j) \in S\}$ represents the sets in $\mathcal{C}^{st}_d$ that will remain to be cross-layer cuts if the physical link (i,j) is used by the (s,t) path Q. We can then develop an approximation algorithm for the augmentation problem similar to REROUTE_SP:

Algorithm 2 AUGMENT_SP(s,t)

1: Construct a weighted graph on $G_P = (V_P, E_P)$, where each edge (i,j) is assigned with weight $w(i,j) = |\mathcal{L}_{ij}|$.
2: Run Dijkstra's algorithm to find the shortest (s,t)-path in the weighted graph.

Since each cross-layer cut S in $\mathcal{C}^{st}_d$ has size d, there are exactly d physical links (i,j) such that $S \in \mathcal{L}_{ij}$. As a result, AUGMENT_SP is a d-approximation algorithm, with the same proof as Theorem 7.

V. Simulation Results

The single-link rerouting and augmentation methods developed in the previous section can be used as a building block for improving the reliability of an existing layered network. For example, by iteratively rerouting the logical links for a given lightpath routing until no further improvement is possible, we can obtain to a locally optimal solution. In this section, we study the effectiveness of such approach via extensive simulation studies.

A. Iterative Rerouting for Survivable Lightpath Routing

We first apply iterative rerouting to solve the Survivable Lightpath Routing problem, whose objective is to obtain a lightpath routing that maximizes the reliability for given physical and logical topologies. The Survivably Lightpath Routing has been previously studied in the literature, where the best known algorithmn [32] is based on an ILP fomulation that maximizes the MCLC of the network. In contrast, the objective for lightpath rerouting algorithm is based on the lexicographical ordering of the cut vector, which captures more precisely the survivability characteristics of the network. As we will learn from the result, using this improved objective significantly improves the quality of the solution.

In this study, we use the NSFNET (Figure 6), extended with new links to raise its connectivity to 4, as the physical topology. For logical topologies, we generate 350 random graphs with connectivity 4, ranging from 6 to 12 nodes; and 13 to 38 links. For each algorithm under evaluation, we compute a lightpath routing for each pair of physcial and logical topologies. The average reliability among the 350 lightpath routings will be presented as the performance metric.

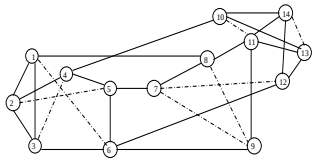


Fig. 6. The extended NSFNET. The dashed lines are the new links.

We will first study the effect of the different initial lightpath routings on the reliability of the final solution. Next, we'll compare the performance of the rerouting algorithm variants based on ILP and the approximation algorithm. Throughout these studies, we also compare the solutions generated by these algorithms with the solution generated by the best known lightpath routing algorithm in the literature, MCF_{LF} [32] (denoted as MCF in this paper for simplicity), as well as the simple shortest path algorithm SP.

1) Performance of ILP-Based Rerouting: We first evaluate the reliability performance of the ILP-based lightpath rerouting approach introduced in Section IV-A2, with initial lightpath routings generated by two different algorithms:

- RR_{SP}: The initial lightpath routing is generated by the Shortest-Path algorithm SP, which routes each lightpath with minimum number of physical hops.
- RR_{MCF}: The initial lightpath routing is generated by the algorithm MCF introduced in [32].

Compared with SP, MCF provides initial lightpath routings with much higher reliability at the expense of longer running

time. Given the initial lightpath routing, the rerouting algorithm repeatedly solves the rerouting ILP in Section IV-A2 to improve the reliability, until it reaches a local optimum.

Figure 7 illustrates the average unreliability of the different algorithms. Even with initial lightpath routings generated by the best known lightpath routing algorithm MCF, the rerouting algorithm $\mathsf{RR}_{\mathsf{MCF}}$ is able to further reduce the unreliability of the lightpath routings. In fact, while only 50% of the lightpath routings generated by MCF has MCLC 4, which is the connectivity of the logical topologies and is therefore the highest MCLC value achievable, the rerouting algorithm $\mathsf{RR}_{\mathsf{MCF}}$ is able to archieve this maximum MCLC value 98% of the time. This means that the lightpath rerouting approach is able to produce lightpath routings that are much more reliable than existing algorithms.

In addition, even though the initial lightpath routings generated by SP and MCF differ significantly in reliability, the iterative rerouting eliminates most of the difference. This suggests that the rerouting approach is robust to initial routings, and we can use a simple algorithm, such as Shortest-Path, to generate the initial lightpath routing and rely on iterative rerouting to obtain a reliable solution.

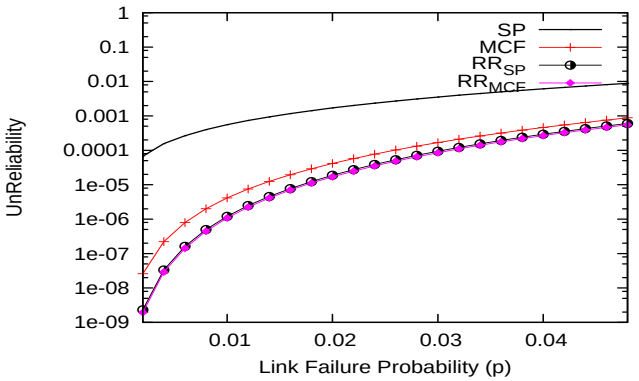


Fig. 7. UnReliability performance by different algorithms.

Table I shows the average physical path length by the lightpaths generated by the different algorithms. The higher reliability of the rerouting algorithms comes with a cost of longer paths, as the algorithms often select the longer physical routes in order to achieve higher reliability. This reflects the tradeoff between the reliability and bandwidth resource used by the lightpath routings.

Number of Logical Nodes	Average Path Length			
	SP	MCF	$\mathsf{RR}_{\mathsf{SP}}$	$\mathsf{RR}_{\mathsf{MCF}}$
6	1.93	2.32	2.60	2.57
7	1.90	2.28	2.58	2.57
8	1.89	2.30	2.68	2.62
9	1.88	2.29	2.57	2.60
10	1.91	2.34	2.68	2.64
11	1.90	2.32	2.60	2.64
12	1.86	2.22	2.51	2.51

TABLE I
AVERAGE PATH LENGTH OF THE SHORTEST-PATH ALGORITHM SP, LIGHTPATH ROUTING ALGORITHM MCF, AS WELL AS THE REROUTING ALGORITHMS USING SHORTEST-PATH ($\mathsf{RR}_{\mathsf{SP}}$) AND MCF ($\mathsf{RR}_{\mathsf{MCF}}$) TO GENERATE THE INITIAL LIGHTPATH ROUTINGS.

Table II shows the average running times of the rerouting algorithms, (not including the time to generate the initial routings), as well as the average number of rerouting iterations. Compared with the lightpath routing algorithm MCF, the rerouting algorithms are able to terminate faster with a better solution. This is because this iterative rerouting approach effectively decomposes the joint lightpath routing problem into simpler single-link rerouting steps, where the ILP in each step is much smaller than the lightpath routing formulation in MCF.

Between the two rerouting variants, $\mathsf{RR}_{\mathsf{SP}}$ requires more iterations than $\mathsf{RR}_{\mathsf{MCF}}$ to reach the local optimum, because it starts with a much less reliable initial lightpath routing. However, the difference in total running time is less significant. This is because the size of the rerouting ILP formulation is larger when the MCLC of the lightpath routing is large, and thus takes longer to solve. In most cases, $\mathsf{RR}_{\mathsf{SP}}$ starts with an initial lightpath routings with a lower MCLC value. As a result, most of the additional rerouting steps consist of solving the smaller ILPs to bring up the MCLC value. Therefore, these additional steps take much shorter time.

Number of Logical Nodes	Running Time (seconds)			Number of Iterations	
	MCF	$\mathsf{RR}_{\mathsf{SP}}$	$\mathsf{RR}_{\mathsf{MCF}}$	$\mathsf{RR}_{\mathsf{SP}}$	$\mathsf{RR}_{\mathsf{MCF}}$
6	1652	265	164	7.0	3.0
7	1655	314	257	8.9	4.2
8	1732	500	365	10.3	5.0
9	1838	745	525	11.6	6.2
10	2032	1238	824	14.1	7.3
11	2219	1389	1280	14.0	8.0
12	2716	1268	1104	14.1	8.2

TABLE II
RUNNING TIMES OF THE LIGHTPATH ROUTING ALGORITHM MCF, AS WELL AS THE REROUTING ALGORITHMS USING SHORTEST-PATH ($\mathsf{RR}_{\mathsf{SP}}$) AND MCF ($\mathsf{RR}_{\mathsf{MCF}}$) TO GENERATE THE INITIAL LIGHTPATH ROUTINGS; AND THE NUMBER OF ITERATIONS OF THE REROUTING ALGORITHMS.

2) Performance of Approximation Algorithm: Next, we compare the performance of the approximation algorithm introduced in Section IV-A3 with the ILP counterpart. As discussed, the approximation algorithm is based on the k-shortest-path algorithm, where the parameter k reflects a tradeoff between running time and reliability performance. We evaluate this algorithm, $\mathsf{Shortest}_k$, with k =1, 10 and 100.

We use the lightpath routings generated by the Shortest Path algorithm as the initial routings. Figure 8 shows the average unreliability of the lightpath routings produced by the algorithms. While $\mathsf{Shortest}_1$ brings in the majority of the improvement, increasing the value of k is able to further reduce the unreliability. In particular, when $k = 100$, the approximation algorithm performs almost as well as solving the rerouting ILP.

Table III compares the running time of the algorithms. As shown in the table, the approximation algorithms are significantly faster than the ILP-based algorithm. This suggests that the approximation algorithm is promising rerouting alternative to the ILP for improving the reliability of large networks, without the need to solve complex mathematical programs.

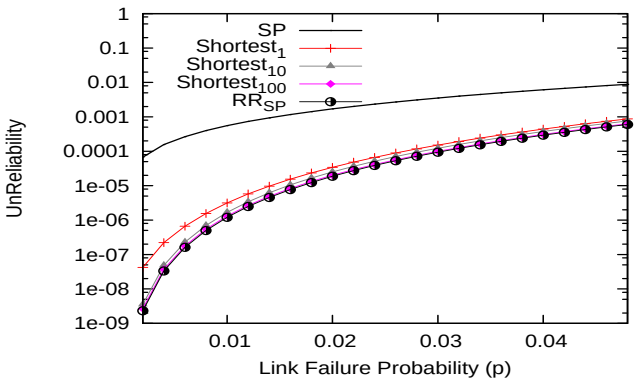


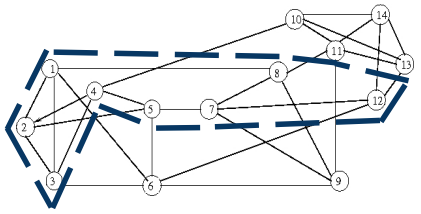
Fig. 8. Lightpath rerouting: performance of approximation algorithm.

Number of Logical Nodes	Running Time (seconds)			
	RR_{SP}	$Shortest_1$	$Shortest_{10}$	$Shortest_{100}$
6	265	12	14	24
7	314	20	26	43
8	500	32	43	79
9	745	45	55	123
10	1238	68	91	199
11	1389	83	104	254
12	1268	113	135	344

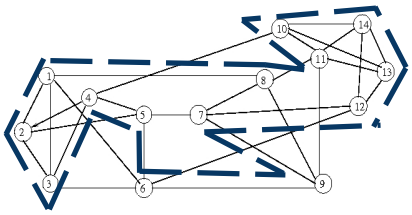
TABLE III
RUNNING TIMES OF THE REROUTING ALGORITHMS BASED ON ILP (RR_{SP}) AND k-SHORTEST PATHS.

B. Effects of Logical Topology Augmentation

Next, we study the effect of augmenting the logical topology on the network reliability. We study a 10-node and a 14-node logical ring on the augmented NFSNET, as shown in Figure 9, and incrementally augment the rings to study the reliability improvement from the addition of new logical links.



(a) 10 Node Logical Ring

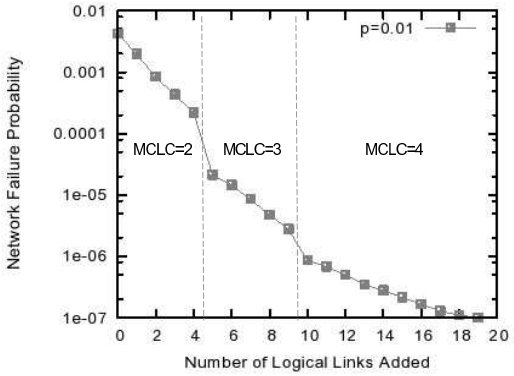


(b) 14 Node Logical Ring

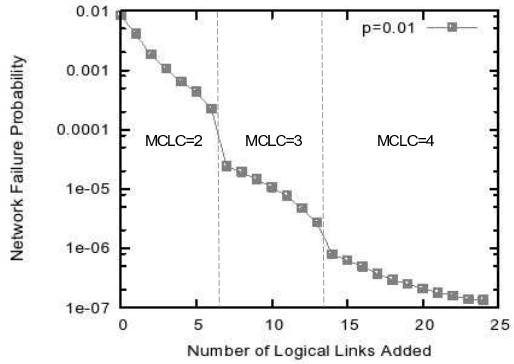
Fig. 9. Logical rings on extended NSFNET.

The cross-layer reliability of the networks after each augmentation step is shown in Figure 10. With link failure probability $p = 0.01$, the unreliability declines as we add more logical links to the rings. The key observation from these figures is that the improvement in reliability is most prominent when the augmentation increases the MCLC of the network. This suggests that networks with a small number of MCLCs have a greater potential to significantly improve the reliability by augmentation, as it is more likely to improve their MCLC values by a small number of new logical links.

In the case where the additional link does not cause an MCLC increase, the marginal reliability improvement decreases with the current MCLC value. This means that augmentation is more effective when the MCLC value is lower.



(a) 10 Node Logical Ring



(b) 14 Node Logical Ring

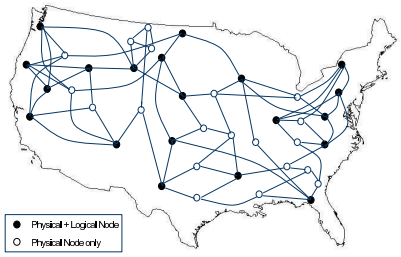
Fig. 10. Impact on reliability by augmenting logical rings.

C. Case Study: A Real-World IP-over-WDM Topology

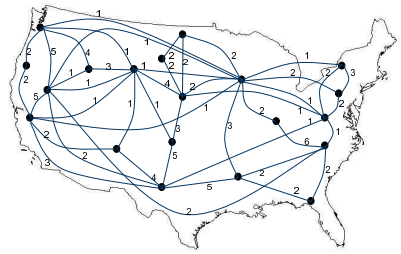
Finally, we evaluate the performance of the rerouting and augmentation algorithms on a large layered network based on a real-world IP-over-WDM network. The physical and logical topologies, shown in Figure 11, are constructed based on the network maps available from Qwest Communications [41]. Both the physical and logical topologies are extended with new links so that the graphs have connectivity 4. The physical topology has 39 nodes and 72 links, and the logical topology has 20 nodes and 101 links.

The study on larger networks allows us to reevaluate the performance of the lightpath algorithms, both in terms of scalability and solution quality. In this study, we run the following lightpath routing algorithms and compare their solutions:

1) MCF: The multi-commodity flow algorithm introduced in [32]. As in Section V-A, the algorithm is evaluated as the performance baseline.
2) REROUTE: The iterative lightpath rerouting algorithm, based on the k-shortest path algorithm presented in Section IV-A3, where k is set to 5000 in our experiment.



(a) WDM (physical) network.



(b) IP/MPLS (logical) network. The numbers indicate the number of parallel logical links between the logical nodes.

Fig. 11. Physical and logical topologies.

3) AUGMENT$_n$: The logical topology augmentation algorithm, based on the k-shortest path algorithm presented in Section IV-B3, where k is set to 5000 in our experiment. The augmentation algorithm is run successively to add n new edges on the lightpath routing given by REROUTE, where $n = 1, \ldots, 9$.

The MCLC values and the number of MCLCs of the lightpath routings generated by each algorithm are shown in Table IV. These numbers are compared against the lower bound, which is computed by counting the number of minimum sized physical fiber sets whose removal will *physically* disconnect some logical nodes. These sets of physical links are cross-layer cuts regardless of the lightpath routing, and therefore will provide a lower bound on the number of MCLCs.

It was observed in [32] that the survivability performance of the multi-commodity flow formulation MCF declines as the network size increases. In this case, the solution produced by the algorithm only has MCLC value 2. On the other hand, the rerouting algorithm REROUTE continues to produce a lightpath routing with the maximum possible MCLC value 4. Augmenting the logical topology can further improve the reliability of the layered network by reducing the number of MCLCs, though the incremental effect declines as more logical links are added to the network. The number of MCLCs hits the lower bound when the logical topology is augmented with 9 additional logical links.

Figure 12 compares the algorithms in terms of the cross-layer reliability in the low failure probability regime. As suggested by Table IV, the iterative algorithms achieve significantly higher reliability than the existing algorithm MCF (by about 3 orders of magnitude). In particular, the majority of the improvement is achieved by the lightpath rerouting algorithm REROUTE. This is because the lightpath rerouting method alone is able to achieve the maximum MCLC value. As we observed in Section V-B, adding logical links is more effective only if the new links can raise the MCLC of the network. In other words, even without adding new logical links, we can obtain a near optimal solution by improving the existing lightpath routing via the iterative rerouting method.

Algorithm	MCLC	Number of MCLCs
MCF	2	5
REROUTE	4	216
AUGMENT$_1$	4	84
AUGMENT$_3$	4	34
AUGMENT$_5$	4	25
AUGMENT$_9$	4	20
Lower Bound	**4**	**20**

TABLE IV
MCLC VALUES AND MCLC COUNTS OF DIFFERENT LIGHTPATH ROUTINGS. THE LIGHTPATH ROUTING ON A LOGICAL TOPOLOGY AUGMENTED WITH n NEW LOGICAL LINKS IS DENOTED BY AUGMENT$_n$.

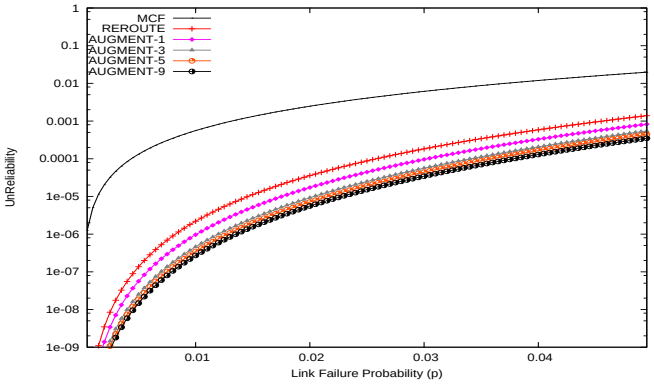


Fig. 12. Unreliability of different lightpath routings.

VI. HIGH FAILURE PROBABILITY REGIME

As discussed in Section I, natural disasters or physical attacks can lead to widespread network link failures. While such events may be extremely rare, certain networks that are of critical importance to national security and our day to day lives may need to be designed so that they can withstand such rare events. Moreover, certain "specialized" networks, such as those onboard an aircraft or a ship may need to be designed to withstand very high link failure probabilities that result from a catastrophic failure event (e.g., well over 50% link failures) [42]. In this section, we briefly discuss network design in this high failure probability scenario.

In Section III-B, we showed that when p is small, it is important to minimize the number of small cuts. Analogously, for large p, large cuts are dominant, and hence, minimizing the number of large cuts would result in maximum reliability. In other words, the cut vector should be minimized for large cuts for better reliability in the high failure probability regime. Similar to the case of low probability regime, we define the following vector of partial sums:

$$\overleftarrow{\mathbf{N}} = \left[\sum_{i=m-k}^{m} N_i | k = 0, ..., m \right].$$

The vector $\overleftarrow{\mathbf{M}}$ is defined similarly. Note that the i-th element $\overleftarrow{N}_i$ is the total number of cross-layer cuts of size at least $m-i$. We will use these vectors to develop the conditions that incrementally include larger cuts and characterize the probability regime where one lightpath routing is more reliable than any other for large p.

First, the k-colexicographical ordering (an analogy to k-lexicographical ordering in Section III-C) is defined as follows:

Definition 6: Lightpath routing 1 is said to be k-colexicographically smaller than lightpath routing 2 if

$$k = \max\left\{j : \overleftarrow{N}_i \leq \overleftarrow{M}_i, \quad \forall i > c-j\right\} \text{ and } k \geq 1,$$

where c is the position of last element where the two cut vectors differ.

In contrast to the k-lexicographical ordering, this colexicographical ordering starts from the largest cuts, and incrementally includes the smaller cuts. The following result is similar to Theorem 3:

Theorem 9: Given two vectors $\mathbf{N}=[N_i|i = 0,\ldots,m]$ and $\mathbf{M}=[M_i|i = 0,\ldots,m]$. For any j, let $\overleftarrow{\Delta}_j = \sum_{i=m-j}^{m}(M_i - N_i)$ and $\overleftarrow{\delta}_j = \max_{0\leq i\leq m-j-1}\left\{\frac{N_i-M_i}{\binom{m}{i}}\right\}$. If $\mathbf{N}$ is k-colexicographically smaller than $\mathbf{M}$, then

$$F_1(p) \leq F_2(p) \text{ for } p \geq p_0^h = 1 - \max\left\{0.5, \min_{c\leq j\leq c+k-1} C_j\right\},$$

where $c = \min\{i : N_{m-i} < M_{m-i}\}$ and

$$C_j = \begin{cases} 0.5, & \text{if } j = m \\ 1 - \frac{1}{\frac{m}{j+1} + \overleftarrow{\delta}_j\binom{m}{j+1}/\overleftarrow{\Delta}_j}, & \text{otherwise.} \end{cases}$$

Proof: See Appendix C. ∎

The following corollary is analogous to Corollary 3 for the high failure regime:

Corollary 5: If $\overleftarrow{N}_j \leq \overleftarrow{M}_j$ for all $j = 0,\ldots,m$, then routing 1 is at least as reliable as routing 2 for $p \geq 0.5$, i.e., $F_1(p) \leq F_2(p)$ for $p \geq 0.5$.

Combining Corollaries 3 and 5 gives a condition for uniformly optimal lightpath routing:

Corollary 6: If $\overrightarrow{N}_j \leq \overrightarrow{M}_j$ and $\overleftarrow{N}_j \leq \overleftarrow{M}_j$ for all $j = 0,\ldots,m$, then lightpath routing 1 is uniformly optimal.

Theorems 3 and 9 provide a single optimality regime expression for lightpath routings that exhibit different degrees of dominance. Note that the conditions of (co)lexicographical ordering in Corollary 6 are satisfied by the uniform optimality condition $N_i \leq M_i, \forall i$ given in Theorem 1. Therefore, this unified theorem allows for a broader class of uniformly optimal lightpath routings.

More importantly, Theorem 9 can be used to derive practical conditions for optimal lightpath routings in the high failure probability regime. We begin with the following definitions:

Definition 7: Consider two lightpath routings 1 and 2. Routing 1 is said to be *more reliable* than routing 2 *in the high failure probability regime* if there exists a number $p_0 < 1$ such that the reliability of routing 1 is higher than that of routing 2 for $p_0 < p < 1$.

Definition 8: A *cross-layer spanning tree* is a *minimal* set of fibers whose survival keeps the logical network connected. Hence, if T is a cross-layer spanning tree, then the survival of just $T \setminus \{(i,j)\}$ renders the logical network disconnected for any fiber $(i,j) \in T$.

Note that the cross-layer spanning tree is a generalization of the single-layer spanning tree. However, unlike a single-layer graph where all spanning trees have the same size, in a layered graph, spanning trees can have different sizes. Thus, we define a *Min Cross Layer Spanning Tree (MCLST)* as a spanning tree with minimum number of physical links.

In the high failure probability regime, it is likely that there are a large number of failures. Hence, the MCLST is an important parameter in the high failure probability regime because logical networks with small MCLST may remain connected even if only a small number of physical links survive. This intuition together with Theorem 9 leads to practical conditions for optimal routings in the high failure probability regime.

Note that in the failure polynomial, $N_i \leq \binom{m}{i}$. Let $m-c$ be the size of MCLST. Then, c is the largest i such that $N_i < \binom{m}{i}$, and we have $N_i = \binom{m}{i}, \forall i > c$, meaning that more than c failures would always disconnect the logical network. Let $m-c_j$ be the size of MCLST for routing j. It is obvious that if $c_1 > c_2$ or $c_1 = c_2$ & $N_{c_1} < M_{c_2}$, then routing 1 is k-colexicographically smaller than routing 2. This observation leads to the corollaries similar to the low regime case:

Corollary 7: If $c_1 > c_2$ (i.e., if routing 1 has smaller MCLST than routing 2), then routing 1 is more reliable than routing 2 in the high failure probability regime.

Corollary 8: If $c_1 = c_2$ and $N_{c_1} < M_{c_2}$ (i.e., routings 1 and 2 have the same size of MCLST, but routing 1 has more MCLSTs), then routing 1 is more reliable than routing 2 in the high failure probability regime.

Therefore, for reliability maximization in the high failure probability regime, it is desirable to find a lightpath routing that minimizes the size of MCLST and maximizes the number of MCLSTs. This observation is similar to the single-layer setting where maximizing the number of spanning trees maximizes the reliability for large p [8]. The major difference in the multi-layer case is that, since spanning trees may have different sizes, minimizing the *size* of the Min Cross-Layer Spanning Tree becomes the primary objective. Moreover, computing the size of the MCLST is NP-hard [23], and therefore, designing a lightpath routing that minimizes the MCLST is likely to be a difficult problem. We developed an ILP-based algorithm that finds a lightpath routing with minimum-size MCLST, and its details can be found in Appendix D.

VII. Conclusion

We studied the reliability maximization problem in layered networks with random link failures. We introduced the notion of lexicographical ordering for lightpath routings, and fully identified optimization criteria for maximum reliability in the low failure probability regime. In particular, we showed that a lightpath routing with the maximum size of Min Cross Layer Cut (MCLC) and the minimum number of MCLCs is most reliable in the low failure probability regime. Based on this insight, we developed a novel lightpath rerouting

approach to design reliable layered networks for the low failure probability regime. By incrementally improving the lightpath routing, this rerouting approach is able to achieve a locally optimal solution. Our simulation results show that the rerouting algorithms developed in this paper are able to produce much more reliable lightpath routings than existing algorithms (by about 3 orders of magnitude in real IP-over-WDM Network), and are more scalable to large networks. Using the optimization criteria, we also developed logical topology augmentation algorithms that can further improve the reliability of a given layered network.

We also showed that the high failure probability regime requires different optimization criteria that a routing with the minimum size of Min Cross Layer Spanning Tree (MCLST) and the maximum number of MCLSTs maximizes reliability. Our results in the high failure probability regime lay the foundation for the design of networks facing increased concern about large scale failures due to natural disasters or attacks.

APPENDIX A
PROOF OF THEOREM 3

We first prove the following lemma.

Lemma 1: If vector **N** is k-lexicographically smaller than vector **M**, then for all $j \leq d+k-1$, where $d = \min\{d : N_d < M_d\}$, and for $0 \leq p \leq 0.5$,

$$\sum_{i=0}^{j}(M_i - N_i)p^i(1-p)^{m-i} \geq \overrightarrow{\Delta}_j p^j(1-p)^{m-j}. \quad (14)$$

Proof: We prove, by induction on j, that (14) holds for all $j \leq d+k-1$. Since $\overrightarrow{\Delta}_0 = M_0 - N_0$, we have for $j=0$,

$$\sum_{i=0}^{j}(M_i - N_i)p^i(1-p)^{m-i} = \overrightarrow{\Delta}_0(1-p)^m$$

Therefore, (14) holds for $j=0$. Now suppose (14) holds for all $i \leq j$ for some $j < d+k-1$. Then, we have:

$$\begin{aligned}
&\sum_{i=0}^{j+1}(M_i - N_i)p^i(1-p)^{m-i}\\
=&\sum_{i=0}^{j}(M_i - N_i)p^i(1-p)^{m-i} + (M_{j+1} - N_{j+1})p^{j+1}(1-p)^{m-(j+1)}\\
\geq&\overrightarrow{\Delta}_j p^j(1-p)^{m-j} + (M_{j+1} - N_{j+1})p^{j+1}(1-p)^{m-(j+1)}\\
\geq&\overrightarrow{\Delta}_j p^{j+1}(1-p)^{m-(j+1)} + (M_{j+1} - N_{j+1})p^{j+1}(1-p)^{m-(j+1)}\\
=&\overrightarrow{\Delta}_{j+1} p^{j+1}(1-p)^{m-(j+1)},
\end{aligned}$$

where the first inequality is due to the induction hypothesis, and the second inequality is because $p/(1-p) \leq 1$. Therefore, by induction, (14) is true for all $j \leq k$. ∎

Lemma 2: Given a fixed k, if $\overrightarrow{\Delta}_i \geq 0$ for all $i \leq d+k-1$, then for any $d \leq j \leq d+k-1$:

$$F_1(p) \leq F_2(p),$$

for $0 \leq p \leq \min\{0.5, B_j\}$, where:

$$B_j = \begin{cases} 0.5, & \text{if } j = m \\ \frac{1}{\frac{m}{j+1} + \overrightarrow{\delta}_j\binom{m}{j+1}/\overrightarrow{\Delta}_j}, & \text{otherwise.} \end{cases}$$

Proof: First, note that by definition of $\overrightarrow{\delta}_j$, for any $i \geq j$:

$$\overrightarrow{\delta}_j\binom{m}{i} \geq N_i - M_i. \quad (15)$$

If $k = m-d+1$, then Lemma 1 implies that, for $p \leq 0.5$:

$$\sum_{i=0}^{m}(M_i - N_i)p^i(1-p)^{m-i} \geq \overrightarrow{\Delta}_m p^m \geq 0.$$

Therefore, the lemma is true for $k = m-d+1$. Now suppose $k < m-d+1$. If $\overrightarrow{\delta}_j \leq 0$ for some $j \leq k$, this implies for any $d+k \leq l \leq m$:

$$\begin{aligned}
\overrightarrow{\Delta}_l &= \overrightarrow{\Delta}_{d+k-1} + \sum_{i=d+k}^{l}(M_i - N_i)\\
&\geq \overrightarrow{\Delta}_{d+k-1} - \sum_{i=d+k}^{l}\overrightarrow{\delta}_j\binom{m}{i} \geq 0,
\end{aligned}$$

where the first inequality is due to (15). The second inequality is due to the fact that $\overrightarrow{\delta}_j \leq 0$, and that $\overrightarrow{\Delta}_{d+k-1} \geq 0$, since **N** is k-lexicographically smaller than **M**. Therefore, in this case, the vector **N** is also $(m-d+1)$-lexicographically smaller than **M**, and the lemma is true as proved above. Therefore, in the rest of the proof, we assume that $\overrightarrow{\delta}_j > 0$.

Since $p \leq 0.5$ and $\overrightarrow{\Delta}_i \geq 0$ for all $i \leq d+k-1$, by Lemma 1 we have, for all $j \leq d+k-1$:

$$\sum_{i=0}^{j}(M_i - N_i)p^i(1-p)^{m-i} \geq \overrightarrow{\Delta}_j p^j(1-p)^{m-j}. \quad (16)$$

Next, we will use the following result to bound the tail probability of the Binomial distribution:

Lemma 3 ([?]): For $r > mp$,

$$\sum_{i=r}^{m}\binom{m}{i}p^i(1-p)^{m-i} \leq \binom{m}{r}p^r(1-p)^{m-r}\cdot\frac{r(1-p)}{r-mp}.$$

Therefore, since $p \leq \frac{1}{\frac{m}{j+1} + \overrightarrow{\delta}_j\binom{m}{j+1}/\overrightarrow{\Delta}_j} < \frac{j+1}{m}$, by Lemma 3, we have:

$$\begin{aligned}
&\sum_{i=j+1}^{m}\binom{m}{i}p^i(1-p)^{m-i} \quad (17)\\
&\leq \binom{m}{j+1}p^{j+1}(1-p)^{m-(j+1)}\cdot\frac{(j+1)(1-p)}{j+1-mp}\\
&= \binom{m}{j+1}p^j(1-p)^{m-j}\cdot\frac{(j+1)p}{j+1-mp}. \quad (18)
\end{aligned}$$

In addition, since $p \leq \frac{1}{\frac{m}{j+1} + \overrightarrow{\delta}_j\binom{m}{j+1}/\overrightarrow{\Delta}_j}$, we have:

$$\begin{aligned}
\frac{(j+1)p}{j+1-mp} &= \frac{1}{\frac{1}{p} - \frac{m}{j+1}}\\
&\leq \frac{1}{\frac{\overrightarrow{\delta}_j\binom{m}{j+1}}{\overrightarrow{\Delta}_j} + \frac{m}{j+1} - \frac{m}{j+1}} = \frac{\overrightarrow{\Delta}_j}{\overrightarrow{\delta}_j\binom{m}{j+1}}. \quad (19)
\end{aligned}$$

It follows that:

$$\sum_{i=0}^{m}(M_i - N_i)p^i(1-p)^{m-i}$$

$$
\begin{aligned}
&= \sum_{i=0}^{j}(M_i - N_i)p^i(1-p)^{m-i} + \sum_{i=j+1}^{m}(M_i - N_i)p^i(1-p)^{m-i} \\
&\geq \sum_{i=0}^{j}(M_i - N_i)p^i(1-p)^{m-i} - \sum_{i=j+1}^{m}\overrightarrow{\delta}_j\binom{m}{i}p^i(1-p)^{m-i} \\
&\geq \overrightarrow{\Delta}_j p^j(1-p)^{m-j} - \overrightarrow{\delta}_j\binom{m}{j+1}p^j(1-p)^{m-j}\cdot\frac{(j+1)p}{j+1-mp} \\
&= p^j(1-p)^{m-j}\overrightarrow{\delta}_j\left(\frac{\overrightarrow{\Delta}_j}{\overrightarrow{\delta}_j} - \binom{m}{j+1}\cdot\frac{(j+1)p}{j+1-mp}\right) \\
&\geq p^j(1-p)^{m-j}\overrightarrow{\delta}_j\left(\frac{\overrightarrow{\Delta}_j}{\overrightarrow{\delta}_j} - \binom{m}{j+1}\cdot\frac{\overrightarrow{\Delta}_j}{\overrightarrow{\delta}_j\binom{m}{j+1}}\right) = 0.
\end{aligned}
$$

The first inequality is due to (15), the second inequality is due to (16) and (18), and the last inequality is due to (19). ■

As a result of Lemma 2, we can pick the $d \leq j \leq d+k-1$ such that B_j is maximized to obtain the largest upper bound for p, and Theorem 3 follows.

Appendix B
Proof of Theorem 7

Given any (s,t) path Q, define $\mathcal{L}(Q) = \cup_{(i,j)\in Q} L_{ij}$, it follows that $N_d(Q) = |\mathcal{L}(Q)| + |\mathcal{C}_d^{st}| = |\mathcal{L}(Q)| + K$, where $K = |\mathcal{C}_d^{st}|$ is a constant. In addition, let $w(Q)$ be the total weight sum of the path Q in the weighted graph constructed by $\mathsf{REROUTE\_SP}(s,t)$.

Since each set of physical links $S \in \mathcal{L}(Q)$ has size d, we have $|\{(i,j) : S \in \mathcal{L}_{ij}\}| \leq d$, which implies:

$$
\begin{aligned}
w(Q) &= \sum_{(i,j)\in Q}|\mathcal{L}_{ij}| \\
&\leq \sum_{S\in\mathcal{L}(Q)}|\{(i,j) : S \in \mathcal{L}_{ij}\}| \\
&\leq d\cdot|\mathcal{L}(Q)| = d\cdot(N_d(Q) - K) \qquad (20)
\end{aligned}
$$

Now, since Q^{SP} is the minimum weight (s,t) path in the graph, it follows that:

$$
\begin{aligned}
N_d(Q^{\mathsf{SP}}) &= |\mathcal{L}(Q^{\mathsf{SP}})| + K \\
&\leq w(Q^{\mathsf{SP}}) + K \\
&\leq w(Q^*) + K \\
&\leq d\cdot(N_d(Q^*) - K) + K, \quad \text{by Equation (20)} \\
&\leq d\cdot N_d(Q^*).
\end{aligned}
$$

Appendix C
Proof of Theorem 9

Let $N_i^{'} = N_{m-i}$ and $M_i^{'} = M_{m-i}$, for $i = 0,\ldots,m$; and let $\overrightarrow{N_k^{'}} = \sum_{i=0}^{k} N_i^{'}$ and $\overrightarrow{M_k^{'}} = \sum_{i=0}^{k} M_i^{'}$. It follows that the vector $\overrightarrow{N^{'}} := \left[\overrightarrow{N_i^{'}} | i = 0,\ldots,m\right]$ is k-lexicographically smaller than the vector $\overrightarrow{M^{'}} := \left[\overrightarrow{M_i^{'}} | i = 0,\ldots,m\right]$. Let $q = 1-p$. Then, by Theorem 3,

$$
\sum_{i=0}^{m}(M_i - N_i)p^i(1-p)^{m-i} = \sum_{i=0}^{m}(M_i^{'} - N_i^{'})q^i(1-q)^{m-i} \geq 0,
$$

for $q \leq \min\left\{0.5, \max_{d\leq j\leq d+k-1} B_j\right\}$, where:

$$
B_j = \begin{cases} 0.5, & \text{if } j = m \\ \frac{1}{\frac{m}{j+1} + \overrightarrow{\delta}_j^{'}\binom{m}{j+1}/\overrightarrow{\Delta}_j^{'}}, & \text{otherwise} \end{cases}
$$

In the above expression, we have:

$$
\overrightarrow{\Delta}_j^{'} = \sum_{i=0}^{j} M_i^{'} - N_i^{'} = \overleftarrow{\Delta}_j, \quad \text{and}
$$

$$
\overrightarrow{\delta}_j^{'} = \max_{j+1\leq i\leq m}\left\{\frac{N_i^{'} - M_i^{'}}{\binom{m}{i}}\right\} = \overleftarrow{\delta}_j.
$$

Note that $B_j = 1 - C_j$ for $d \leq j \leq d+k-1$. Therefore, routing 1 is at least as reliable as routing 2 for

$$
\begin{aligned}
p = 1-q &\geq 1 - \min\left\{0.5, \max_{d\leq j\leq d+k-1} B_j\right\} \\
&= \max\left\{0.5, \min_{d\leq j\leq d+k-1} C_j\right\}.
\end{aligned}
$$

This completes the proof.

Appendix D
Lightpath Routing ILP to Minimize Minimum Cross Layer Spanning Tree (MCLST) Size

As discussed in Section VI, lightpath routings with smaller MCLST size will be more reliable in the high failure probability regime. In this section, we present an ILP for the lightpath routing formulation that minimizes the MCLST.that are optimized for the high failure probability regime. We first define the following variables:

- $\left\{f_{ij}^{st} | (s,t) \in E_L, (i,j) \in E_P\right\}$: Flow variables representing the lightpath routing.
- $\{y_{ij} | (i,j) \in E_P\}$: 1 if fiber (i,j) survives, 0 otherwise.
- $\{z^{st} | (s,t) \in E_L\}$: 1 if lightpath (s,t) survives, 0 otherwise.
- $\{x^{st} | (s,t) \in E_L\}$: Flow variables on the logical topology.

$$
\mathsf{MCLST}: \quad \text{Minimize} \sum_{(i,j)\in E_P} y_{ij}, \quad \text{subject to:}
$$

$$
\sum_{t\in V_L} x^{st} - \sum_{t\in V_L} x^{ts} = \begin{cases} |V_L| - 1, & \text{if } s = 0 \\ -1, & \text{if } s \in V_L - \{0\} \end{cases} \qquad (21)
$$

$$
(V_L - 1)\cdot z^{st} \geq x^{st}, \quad \forall (s,t) \in E_L \qquad (22)
$$

$$
y_{ij} \geq z^{st} + f_{ij}^{st} - 1 \quad \forall (s,t) \in E_L, \forall (i,j) \in E_P \qquad (23)
$$

$$
\{(i,j) : f_{ij}^{st} = 1\} \text{ forms an } (s,t)\text{-path in } G_P, \forall (s,t) \in E_L
$$

$$
0 \leq y_{ij} \leq 1; \quad 0 \leq x^{st}; \quad z_{ij}, f_{ij}^{st} \in \{0,1\}
$$

The variables x^{st} represent a flow on the logical topology where 1 unit of flow is sent from logical node 0 to every other logical node, as described by Constraint (21). Constraint (22) requires these flows to be carried only on the surviving logical links, which implies that the surviving links form a connected logical subgraph. Constraint (23) ensures the survival of physical links that are used by any surviving logical links. Since the objective function minimizes $\sum_{(i,j)\in E_P} y_{ij}$, the

optimal solution will represent a minimum set of physical links whose survival will allow the logical link to be connected.

Therefore, the set of physical links (i, j) with $y_{ij} = 1$ forms a cross-layer spanning tree. As a result, the optimal solution to the above ILP yields a lightpath routing that minimizes the size of the MCLST.

REFERENCES

[1] J. Zhang and B. Mukherjee, "A review of fault management in WDM mesh networks: Basic concepts and research challenges," *IEEE Network*, pp. 41–48, Mar 2004.

[2] J. Zhang, K. Zhu, H. Zang, and B. Mukherjee, "Service provisioning to provide per-connection-based availability guarantee in WDM mesh networks," in *OFC*, Atlanta, GA, Mar 2003.

[3] M. Tornatore, G. Maier, and A. Pattavina, "Availability design of optical transport networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 8, pp. 1520–1532, August 2005.

[4] J. Segovia, E. Calle, P. Vila, J. Marzo, and J. Tapolcai, "Topology-focused availability analysis of basic protection schemes in optical transport networks," *J. of Opt. Net.*, vol. 7, pp. 351–364, Mar. 2008.

[5] F. Boesch, X. Li, and C. Suffel, "On the existence of uniformly optimally reliable networks," *Networks*, vol. 21, pp. 181–194, 1991.

[6] G. Wang, "A proof of Boesch's conjecture," *Networks*, vol. 24, pp. 277–284, 1994.

[7] W. Myrvold, K. Cheung, L. Page, and J. Perry, "Uniformly-most reliable networks do not always exist," *Networks*, vol. 21, pp. 417–419, 1991.

[8] D. Bauer, F. Boesch, C. Suffel, and R. Van Slyke, "On the validity of a reduction of reliable network design to a graph extremal problem," *IEEE Trans. on Circ. and Sys.*, vol. 34, no. 12, pp. 1579–1581, December 1987.

[9] W. Myrvold, "Reliable network synthesis: Some recent developments," in *International Conference on Graph Theory, Combinatorics, Algorithms, and Applications, volume II*, 1996, pp. 650–660.

[10] Y. Ath and M. Sobel, "Some conjectured uniformly optimal reliable networks," *Probab. Eng. Inf. Sci.*, vol. 14, no. 3, pp. 375–383, 2000.

[11] ——, "Counterexamples to conjectures for uniformly optimally reliable graphs," *Probab. Eng. Inf. Sci.*, vol. 14, no. 2, pp. 173–177, 2000.

[12] F. Boesch, "On unreliability polynomials and graph connectivity in reliable network synthesis," *Journal of Graph Theory*, vol. 10, no. 3, pp. 339–352, 1986.

[13] M. Kurant, H. X. Nguyen, and P. Thiran, "Survey on dependable IP over fiber networks," *LNCS*, vol. 4028, pp. 55–81, November 2006.

[14] J. Armitage, O. Crochat, and J.-Y. L. Boudec, "Design of a survivable WDM photonic network," in *IEEE INFOCOM*, April 1997.

[15] O. Crochat and J.-Y. L. Boudec, "Design protection for WDM optical networks," *IEEE JSAC*, vol. 16, no. 7, Sep. 1998.

[16] Q. Deng, G. Sasaki, and C.-F. Su, "Survivable IP over WDM: a mathematical programming problem formulation," in *Proc 40th Allerton Conference on Communication, Control and Computing*, Oct. 2002.

[17] M. Kurant and P. Thiran, "Survivable routing of mesh topologies in IP-over-WDM networks by recursive graph contraction," *IEEE JSAC*, vol. 25, no. 5, pp. 922–933, June 2007.

[18] H. Lee, H. Choi, S. Subramaniam, and H.-A. Choi, "Survivable embedding of logical topologies in WDM ring networks," *Inf. Sci. Inf. Comput. Sci.*, vol. 149, no. 1-3, pp. 151–160, 2003.

[19] E. Modiano and A. Narula-Tam, "Survivable lightpath routing: A new approach to the design of WDM-based networks," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 4, May 2002.

[20] A. Sen, B. Hao, and B. H. Shen, "Survivable routing in WDM networks," in *Proceedings of the Seventh International Symposium on Computers and Communications*, Washington, DC, USA, 2002.

[21] ——, "Minimum cost ring survivability in WDM networks," in *Workshop on High Performance Switching and Routing*, 2003, pp. 183–188.

[22] A. Todimala and B. Ramamurthy, "Survivable virtual topology routing under Shared Risk Link Groups in WDM networks," in *BROADNETS*, Washington, DC, USA, 2004.

[23] K. Lee, H.-W. Lee, and E. Modiano, "Reliability in layered networks with random link failures," *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1835 –1848, dec. 2011.

[24] M. To and P. Neusy, "Unavailability analysis of long-haul networks," *IEEE JSAC*, vol. 12, no. 1, pp. 100–109, 1994.

[25] W. Wu, B. Moran, J. Manton, and M. Zukerman, "Topology design of undersea cables considering survivability under major disasters," in *WAINA*, May 2009.

[26] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the vulnerability of the fiber infrastructure to disasters," *Networking, IEEE/ACM Transactions on*, vol. 19, no. 6, pp. 1610 –1623, dec. 2011.

[27] J. S. Foster Jr. et al., "Report of the commission to assess the threat to the united states from electromagnetic pulse (emp) attack, vol. i: Executive report," Apr. 2004.

[28] C. Wilson. (2004) High altitude electromagnetic pulse (hemp) and high power microwave (hpm) devices: Threat assessments. [Online]. Available: http://www.fas.org/man/crs/RL32544.pdf

[29] W. Radasky. (2007) High-altitude electromagnetic pulse (hemp): A threat to our way of life. [Online]. Available: http://www.todaysengineer.org/2007/Sep/HEMP.asp

[30] "Detailed technical report on emp and severe solar flare threats to the u.s. power grid," Oct. 2010.

[31] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. Porcellinis, and R. Setola, "Modelling interdependent infrastructures using interacting dynamical models," *International Journal of Critical Infrastructures*, vol. 4, no. 1, pp. 63–79, 2008.

[32] K. Lee, E. Modiano, and H.-W. Lee, "Cross-layer survivability in WDM-based networks," *IEEE/ACM Transactions on Networking*, vol. 19, no. 4, pp. 1000 –1013, aug. 2011.

[33] K. Lee, "Survivability in layered networks," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2011.

[34] S. Yuan, S. Varma, and J. P. Jue, "Minimum-color path problems for reliability in mesh networks," in *INFOCOM*, 2005, pp. 2658–2669.

[35] J. Y. Yen, "Finding the k shortest loopless paths in a network," *Management Science*, vol. 17, no. 11, pp. 712–716, 1971.

[36] A. Frank, "Connectivity augmentation of networks: structures and algorithms," *Mathematical Programming*, vol. 84, pp. 439–441, 1999.

[37] G.-R. Cai and Y.-G. Sun, "The minimum augmentation of any graph to a k-edge-connected graph," *Networks*, vol. 19, no. 1, pp. 151–172, January 1989.

[38] T.-S. Hsu, T. Suel, R. Tamassia, and H. Y. D. The, "Graph augmentation and related problems: Theory and practice," 1993.

[39] T. Watanabe, Y. Higashi, and A. Nakamura, "An approach to robust network construction from graph augmentation problems," vol. 4, May 1990, pp. 2861–2864.

[40] B. Jackson and T. Jordán, "Connectivity augmentation of graphs," *Electronic Notes in Discrete Mathematics*, vol. 5, pp. 185–188, 2000, 6th International Conference on Graph Theory.

[41] "Qwest Communications," https://www.qwest.com/business/resource-center/network-ma

[42] G. Weichenberg, V. Chan, and M. Medard, "High-reliability topological architectures for networks under stress," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1830 – 1845, nov. 2004.

PLACE PHOTO HERE

Hyang-Won Lee (S'02-M'08) received the BS, MS, and PhD degrees in all electrical engineering and computer science from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 2001, 2003, and 2007, respectively. He was a postdoctoral research associate at the Massachusetts Institute of Technology during 2007-2011, and a research assistant professor at KAIST between 2011 and 2012. He is currently an assistant professor in the Department of Internet and Multimedia Engineering, Konkuk University, Seoul, Republic of Korea. His research interests are in the areas of congestion control, stochastic network control, robust optimization and reliable network design.

PLACE PHOTO HERE

Kayi Lee (S'99-M'05) received the B.S. degree in Computer Science and Mathematics from the Massachusetts Institute of Technology (MIT), Cambridge, in 1999 and his M.Eng., and PhD degrees in Computer Science from MIT in 2000 and 2011 respectively. He is currently a Software Engineer at Google Inc., Cambridge, MA. His research interests include routing and network survivability.

PLACE PHOTO HERE

Eytan Modiano (S'90-M'93-SM'00-F'12) received the BS degree in electrical engineering and computer science from the University of Connecticut at Storrs in 1986 and the MS and PhD degrees, both in electrical engineering, from the University of Maryland, College Park, in 1989 and 1992, respectively. He was a US Naval Research Laboratory fellow between 1987 and 1992 and a US National Research Council postdoctoral fellow during 1992-1993. Between 1993 and 1999, he was with the Massachusetts Institute of Technology (MIT) Lincoln Laboratory where he was the project leader for MIT Lincoln Laboratorys Next Generation Internet (NGI) project. Since 1999, he has been on the faculty at MIT, where he is presently a professor. His research is on communication networks and protocols with emphasis on satellite, wireless, and optical networks. He is currently an associate editor for the IEEE/ACM Transactions on Networking and for The International Journal of Satellite Communications and Networks. He had served as associate editor for the IEEE Transactions on Information Theory (communication networks area) and as a guest editor for the IEEE JSAC special issue on WDM network architectures, the Computer Networks journal special issue on broadband internet access, the Journal of Communications and Networks special issue on wireless ad-hoc networks, and for the IEEE Journal of Lightwave Technology special issue on optical networks. He was the technical program cochair for Wiopt 2006, IEEE INFOCOM 2007, and ACM MobiHoc 2007. He is an associate fellow of the AIAA and a fellow of the IEEE.