This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

Network Reliability With Geographically Correlated Failures

Sebastian Neumayer and Eytan Modiano
Massachusetts Institute of Technology (MIT)
Email: {bastian,modiano}@mit.edu

***Abstract*—Fiber-optic networks are vulnerable to natural disasters, such as tornadoes or earthquakes, as well as to physical failures, such as an anchor cutting underwater fiber cables. Such real-world events occur in specific geographical locations and disrupt specific parts of the network. Therefore, the geography of the network determines the effect of physical events on the network's connectivity and capacity.**

In this paper, we develop tools to analyze network failures after a 'random' geographic disaster. The random location of the disaster allows us to model situations where the physical failures are not targeted attacks. In particular, we consider disasters that take the form of a 'random' line in a plane. Using results from geometric probability, we are able to calculate some network performance metrics to such a disaster in polynomial time. In particular, we can evaluate average two-terminal reliability in polynomial time under 'random' line-cuts. This is in contrast to the case of independent link failures for which there exists no known polynomial time algorithm to calculate this reliability metric. We also present some numerical results to show the significance of geometry on the survivability of the network and discuss network design in the context of random line-cuts. Our novel approach provides a promising new direction for modeling and designing networks to lessen the effects of geographical disasters or attacks.

I. Introduction

The world's communications infrastructure is primarily based on fiber-optic networks and as such has physical vulnerabilities. Anything from Electromagnetic Pulse (EMP) attacks [14], [27], [32] to dragging anchors [6], [31] can destroy critical network components. The geographical layout of the network affects the impact of such disasters since they occur in specific geographic locations. For example, an earthquake that cuts fiber-optic cables near a city which serves as a telecommunications hub is likely to cause significant disruption to communications services. In this paper, we attempt to evaluate the impact of geographically correlated failures on network connectivity. We first develop the necessary tools to evaluate network performance metrics under a geographic failure model. This, in turn, allows us to study network designs that can lessen the effects of regional disasters.

Real-world disasters affect particular geographical locations and thus will result in failures of neighboring network components. For example, the electronics (e.g., fiber-optic amplifier or repeaters) required to run optical links may fail due to a hurricane or an earthquake and lead to the failure of the associated optical fiber links. Therefore, it is necessary to consider the effect of disasters on the physical (fiber optic) network topology rather than on the logical (e.g., IP) network topology.

In our previous works we considered the problem of finding the worst-case location for a failure in a geographic network with respect to certain network connectivity measures [25], [26]. This models the scenario where the network is attacked with the intention to reduce its capacity or connectivity. On the other hand, in this paper we consider the impact of random disaster on network connectivity; the random location of the disasters can model failures resulting from natural disasters or collateral (non-targeted) damage in an attack. In the same vein, recently [33] studied the survivability of undersea cables under major disasters for a two node network topology.

Wide-spread failures have been studied extensively in the context of logical topology [8], [9], [11], [18]. Most of these papers model the topology of the Internet as a random graph [4] and use percolation theory to study the effects of random link failures on these graphs and whether the resulting network has a large connected component. However, most of these works focus on the logical topology rather than physical topology because they are motivated by failures of routers due to logical attacks (e.g., viruses and worms). Based on various measurements (e.g., [10]), it has been recently shown that the topology of the Internet is influenced by geographical factors such as population density [3], [15], [34]. Yet, these works do not consider the effect of failures that are geographically correlated.

Network survivability has also been studied extensively (e.g., [5], [12], [19], [35]), but most of the previous work in the study of physical topology and fiber networks (e.g., [22], [23]) focus on a small number of fiber failures (typically single isolated failures). In contrast, here we focus on events that induce a large number of failures in a certain geographical region (e.g., [6], [14], [27], [31]).

In this paper we model random disasters by lines which are 'randomly' located on the plane. Any links which are intersected by the randomly placed line are removed from the network. This can model damage to the communication infrastructure which is a result of a random natural disaster such as a tornado or collateral damage in an EMP attack. It can also model manufacturing errors in a VLSI ship or damage to a printed circuit board. We use geometric probability to assign a measure to lines in the plane that intersect some set of line segments. Using these basic tools which are introduced and explained in Section II, we are able to calculate network performance measures to a random line-cut in polynomial time. To the best of our knowledge this is the first paper to apply geometric probability techniques to network survivability ([17] applies similar techniques to detection in sensor networks). In particular, we calculate the average two-terminal reliability of a network in polynomial time with respect to a 'random' line-cut. This result is significant because there exists no known polynomial time algorithm to calculate two-terminal reliability under independent link failures; thus, it is clear that geographically correlated failures are fundamentally different from independent failures. We also present some numerical results that demonstrate the significance of geometry on the survivability of the network and discuss network design in the context of random line-cuts.

Our ultimate goal is to understand the effect of regional failures on the connectivity and reliability of a network, and to expose the design tradeoffs related to network survivability under a regional disaster. These tradeoffs may imply that there is a need to protect electronic components in certain regions (e.g., shielding against EMP attacks [14], [27]) or there may be a need to redesign parts of the network in order to make it less vulnerable to regional failures. This paper makes a small but fundamental step towards network design under geographically correlated failures.

The paper is organized as follows: In Sections II and III we introduce geometric probability and present an algorithm that allows us to evaluate joint link failure probabilities after a random line-cut. In Section IV we use these results to evaluate average two-terminal reliability to a random line-cut (among other measurements). In Section V we present some numerical results to show the significance of geometry on the survivability of the network and then in Section VI we present some network design problems in the context of random-line cuts. We conclude and discuss future research directions in Section VII.

II. Modeling Random Cuts in Geographic Networks

In this section we describe how to model random disasters using geometric probability. For simplicity, we focus only on disasters which remove links along a random line. For example, damage to communication infrastructure can be a result of a natural disaster such as a tornado or collateral damage in an EMP attack. As these disasters are not targeted, these events can be modeled as breaking fiber along random line. After introducing some basic definitions from geometric probability, we review classic results which allow us to find single and pairwise link failure probabilities. These results are requisite for Section III where we show how to find joint link failure probabilities to random line-cuts.

A. Geometric Probability

Geometric probability is the study of probabilities involved in geometric problems. In our case, we are interested in the probability that a 'randomly' placed line in a plane will intersect a certain set of links (e.g., links whose removal would disconnect the network). It should be noted that the problem we are interested in is very similar in nature to the Buffon's Needle problem [30].

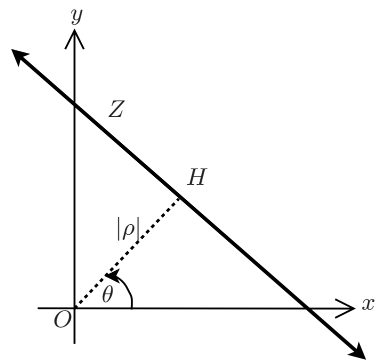

Fig. 1. Let Z be a line in the plane, O be the origin, and H be the closest point on Z to O (see Fig. 1). Note that every line in the plane can be parameterized by ρ and θ where θ is the smallest non-negative angle between the x-axis and the line's normal and $|\rho|$ is the Euclidean distance between H and O such that $\rho \in \mathbb{R}$, $\theta \in [0, \pi)$, and the equation of the line is $x \cos\theta + y \sin\theta - \rho = 0$.

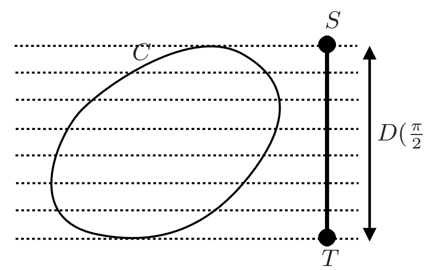

Fig. 2. Consider a set of parallel lines ($\theta = \frac{\pi}{2}$ but variable ρ) that intersect the convex set C. Projecting C onto a line perpendicular to the parallel lines results in line segment $\overline{ST}$. Let $D(\frac{\pi}{2})$ be the length of this line segment. By considering all angles θ, we have $m([C]) = \int_0^{\pi} D(\theta) d\theta$.

Before proceeding further, we will present some useful notation. Let C be a closed bounded convex set on the plane. Let L_C be the perimeter of C (where perimeter is the length of the boundary). Also, let $[C]$ denote the set of all lines in the plane which intersect C.

Geometric probability tells us how to assign a measure to sets of lines; let this measure be denoted by m. The rest of this section reviews results from geometric probability (see [20], [29]) that are necessary for the development of this work.

Let Z be a line in the plane, O be the origin, and H be the closest point on Z to O (see Fig. 1). Note that every line in the plane can be parameterized by ρ and θ where θ is the smallest non-negative angle between the x-axis and the line's normal and $|\rho|$ is the Euclidean distance between H and O such that $\rho \in \mathbb{R}$, $\theta \in [0, \pi)$, and the equation of the line is $x \cos\theta + y \sin\theta - \rho = 0$.

We first start by considering the set of parallel lines (fixed θ but variable ρ) that intersect the convex set C (see Fig. 2). Projecting C onto a line perpendicular to the parallel lines results in a line segment whose length is given by $D(\theta)$ (see Fig. 2). By considering all angles θ, we have $m([C]) = \int_0^{\pi} D(\theta) d\theta$.

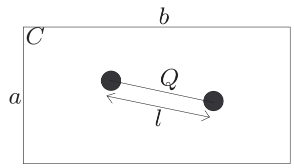

Fig. 3. Rectangle C with link Q inside.

For example, consider a circle of radius 1. Since the projection of this circle onto any line is a line segment of length 2, we know that $D(\theta) = 2$ for all θ. So $\int_0^\pi D(\theta)d\theta = 2\pi$. Next consider a horizontal line segment of length 1. By simple trigonometry, we know $D(\theta) = |\cos\theta|$ and thus $\int_0^\pi D(\theta)d\theta = \sin\theta|_0^{\frac{\pi}{2}} - \sin\theta|_{\frac{\pi}{2}}^{\pi} = 2$.

Notice that in the above examples $m([C])$ is equal to the perimeter of C, L_C. This result holds for all convex sets, as stated in the following lemma [20], [29].

Lemma 1 ([29]): Let C be a bounded closed convex set, then

$$m([C]) = \int_{[C]} d\rho d\theta = \int_0^\pi D(\theta)d\theta = L_C$$

B. Single Link Failures

Let $[Q]$ and $[C]$ be sets of lines in the plane such that $[Q] \subset [C]$. Given m, the probability a 'random' line is in the set $[Q]$ when it is known to be in the set $[C]$ is defined to be ratio of measures [29], $\frac{m([Q])}{m([C])}$. This definition appeals to intuition; $m([C])$ in some sense represents the 'weight' of lines in $[C]$ and $m([Q])$ represents the 'weight' of lines in $[Q]$. Therefore it makes sense that the probability a line in $[C]$ is also in $[Q]$ is $\frac{m([Q])}{m([C])}$.

We now present an example relating to network survivability. Consider a rectangle C with height a and width b and a line segment Q of length l inside C (see Fig. 3). Now we consider a random line-cut. We have:

$$\Pr(Q \text{ cut}|C \text{ cut}) = \frac{m([Q])}{m([C])} = \frac{L_Q}{L_C} = \frac{2l}{2(a+b)}$$

C. Pairwise Link Failures

We now present a classic definition and result in geometric probability which allows us to find pairwise link failure probabilities with respect to a random line-cut.

Definition 1 (Internal Cover): The internal cover of two bounded convex sets in the plane, A and B, denoted by $I(A, B)$ is given by the following. If $A \bigcap B = \emptyset$ then the internal cover is realized by a closed elastic string drawn about A and B and crossing over a point O placed between A and B [29] (see Fig. 4). If $A \bigcap B \neq \emptyset$, then the internal cover is realized by a string which is wrapped around the entire perimeter of both A and B. In this case, $L_{I(A,B)} = L_A + L_B$.

Let $conv(A)$ denote the convex hull of set A.

Lemma 2 ([29]): If A and B are bounded closed convex sets,

$$m([A] \bigcap [B]) = L_{I(A,B)} - L_{conv(A \cup B)}$$

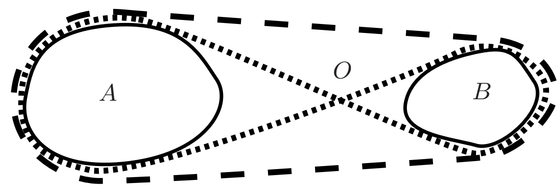

Fig. 4. The dotted curve shows the internal cover of A and B, denoted by $I(A, B)$. The dashed curve shows the boundary of the convex hull of $A \cup B$.

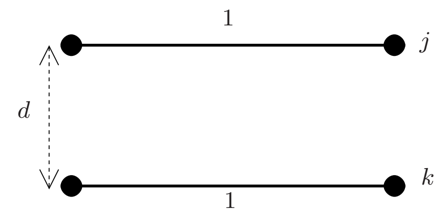

Fig. 5. $Pr($ j and k cut $|$ j cut $) = \sqrt{1+d^2} - d$

Given two links, j and k, by definition the probability a 'random' line is in the set $[j] \cap [k]$ when it known to be in the set $[j]$ (note that $[j] \cap [k] \subset [j]$) is the ratio of measures $\frac{m([j]\cap[k])}{m([j])}$. So, using Lemma 2, we find

$$\Pr(k \text{ cut}|j \text{ cut}) = \frac{m([k] \bigcap [j])}{m([j])} = \frac{L_{I(j,k)} - L_{conv(j \cup k)}}{L_j}$$

Examples demonstrating of the above result for pairwise link failures are given below.

Example 1: Two parallel links, j and k, of length 1 are separated by a distance d. The nodes form corners of a rectangle. See Fig. 5. Since the length of a diagonal is given by $\sqrt{1+d^2}$, we know $L_{I(j,k)} = 2 + 2\sqrt{1+d^2}$. Also, the perimeter of the rectangle is given by $2 + 2d$. Therefore, $Pr($ j and k cut $|$ j cut $) = \sqrt{1+d^2} - d$.

Example 2: Two links, j and k, of length 1 overlap as shown in Fig. 6 where the length of the overlap is ϵ. These links intersect, so $L_{I(j,k)} = 4$ by definition. Also, it is evident $L_{conv(j \cup k)} = 2(2 - \epsilon)$. Therefore, $Pr($ j and k cut $|$ j cut $) = \epsilon$.

Example 3: Two links, j and k, of length 1 are at an angle β to each other and share a common node (assume $\beta \in [0, \pi]$).

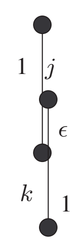

Fig. 6. $Pr($ j and k cut $|$ j cut $) = \epsilon$

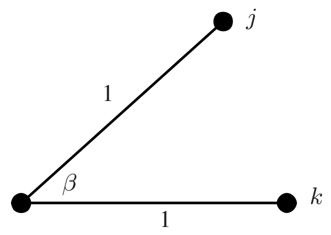

Fig. 7. $Pr(\ j$ and k cut $\mid j$ cut $) = 1 - \sin\frac{\beta}{2}$

See Fig. 7. These links intersect, so $L_{I(j,k)} = 4$. Also, the perimeter of the convex hull is given by $2+2\sin\frac{\beta}{2}$. Therefore, $Pr(\ j$ and k cut $\mid j$ cut $) = 1 - \sin\frac{\beta}{2}$. This result agrees with our intuition. If $\beta = 0$, then the links are on top of each other and the probability is one. If $\beta = \pi$, then only lines which intersect the shared node intersect both links and the probability is zero.

III. Geographically Correlated Link Failures

In this section we present an algorithm which calculates the measure of lines that intersect every line segment in a set of segments. This result will allow us to calculate the probability that a random line-cut intersects a certain set of links in a network (e.g. links whose removal would disconnect the network). We will then use this to efficiently calculate network performance measures with respect to random line-cuts. The details of this section may be skipped and the reader may proceed without loss of continuity to Section IV on evaluating network reliability.

Assume we are given a set of line segments, Q, on a plane such that the endpoints are in general form; that is, no three endpoints are collinear[1]. Let the i^{th} line segment be denoted by Q_i. Our goal is to find $m(\bigcap_{i=1}^{|Q|}[Q_i])$; that is, the measure of the set of lines that intersect all segments in Q.

Sylvester in [30] shows how to solve for $m(\bigcap_{i=1}^{|Q|}[Q_i])$. However, this approach takes exponential time in $|Q|$; this is because the perimeter of the convex hull of every subset of Q must be considered. Ambartzumian in [1] and [2] provides an algorithm to calculate $m(\bigcap_{i=1}^{|Q|}[Q_i])$ in polynomial time. The algorithm in this section presents an alternate way to solve for $m(\bigcap_{i=1}^{|Q|}[Q_i])$ in polynomial time by reducing the problem to finding pairwise link failures, as was shown in Section II-C.

In the following, for clarity of presentation we break down our procedure into steps for finding $m(\bigcap_{i=1}^{|Q|}[Q_i])$. The steps are supported with lemmas, however due to space constraints, all the proofs in this section are omitted (they can be found in [24]).

<u>*Step 1*</u>:

Let P be the set of endpoints of the line segments in Q. Let us impose an arbitrary ordering on P and denote the i^{th} point in P by P_i. Let $\overline{P_iP_j}$ be the line segment between P_i and P_j.

[1]This assumption is not restrictive as we can slightly perturb the location of the endpoints to satisfy this condition.

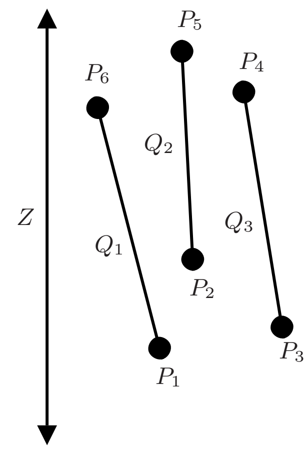

Fig. 8. Shown above are three line segments, Q_1, Q_2, and Q_3, and a line Z which does not intersect the convex hull of P. We want to find $m([Q_1]\bigcap[Q_2]\bigcap[Q_3])$ which is equivalent to $m([Z]\bigcap[Q_1]\bigcap[Q_2]\bigcap[Q_3])$.

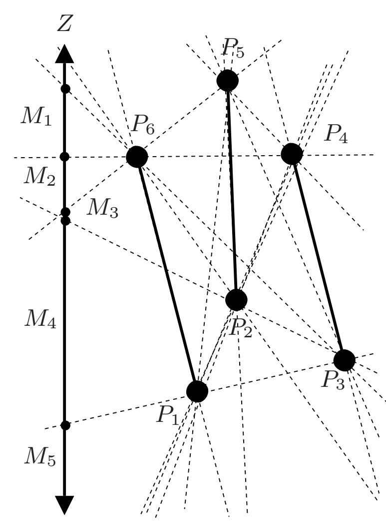

Fig. 9. Consider all lines that intersect two points in P. Let the intersection points of these lines and Z be denoted by α (shown as dots on Z above). Let the divisions of of Z into line segments and rays by points in α be denoted by M. Let us impose an ordering on M and denote M_i to be the i^{th} segment in M.

We start by arbitrarily placing a vertical line Z such that it does not intersect the convex hull of P (see Fig. 8). Note that $m(\bigcap_{i=1}^{|Q|}[Q_i]) = m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[Z]\big)$ because the set of all lines which do not intersect Z has measure zero.

<u>*Step 2*</u>:

Consider all lines that intersect two points in P. Let the intersection points of these lines and Z be denoted by α. Let the divisions of Z into line segments and rays by points in α be denoted by M (see Fig. 9). Let us impose an ordering on M and denote M_i to be the i^{th} segment in M.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

Step 3:

Now,

$$\begin{aligned} m(\bigcap_{i=1}^{|Q|}[Q_i]) &= m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[Z]\big) \\ &= m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap(\bigcup_{M_j\in M}[M_j])\big) \\ &= m\Big(\bigcup_{M_j\in M}((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[M_j])\Big) \end{aligned}$$

Since every $[M_j]$ is disjoint from $[M_k]$ when $j \neq k$ up to measure zero, we have:

$$m\Big(\bigcup_{M_j\in M}((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[M_j])\Big) = \sum_{M_j\in M} m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[M_j]\big)$$

Our problem is now reduced to computing $m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[M_j]\big)$ for every j. That is, the measure of the set of lines that intersect both M_j and each of the segments in Q. We will show that computing this is easy because it is equivalent to computing $m([\overline{P_kP_l}]\bigcap[M_j])$ for some k and l. That is, $m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[M_j]\big)$ is the same as the measure of the set of lines intersecting M_j and a line segment connecting two points in P.

Step 4:

In the following steps, we assume X is a point on Z such that $X \notin \alpha$.

Definition 2 ($T(X)$): $T(X)$ is an *ordered* set of all points in P such that when Z is rotated counter-clockwise about X, the order in which points in P are intersected is the ordering in $T(X)$.

For an illustration of this definition see Fig. 10.

Lemma 3: $T(X) = T(X')$ for every $X \in M_j$ and $X' \in M_j$.

Intuitively, this lemma states that the ordering of $T(X)$ is the same for all X in M_j.

Step 5:

Definition 3 (A_1^X): A_1^X is the last point in $T(X)$ such that there does not exist P_k and P_l ahead of A_1^X where $\overline{P_kP_l} \in Q$.

Definition 4 (A_2^X): A_2^X is the first point in $T(X)$ such that there exists a P_k before A_2^X where $\overline{P_kA_2^X} \in Q$.

See Fig. 10 for an example.

Lemma 4: If A_1^X comes before A_2^X, then $[X]\bigcap(\bigcap_{i=1}^{|Q|}[Q_i]) = [X]\bigcap[\overline{A_1^XA_2^X}]$, otherwise if A_2^X comes before A_1^X, then $[X]\bigcap(\bigcap_{i=1}^{|Q|}[Q_i]) = \emptyset$.

Intuitively, this lemma says the set all lines which intersect X and every line segment in Q is the same as the set of all lines which intersect X and some $\overline{P_kP_l}$. Take Fig. 10 as an example. The set of lines that intersect X and all three line segments is equivalent to the set of lines that intersect X and $\overline{P_2P_4}$.

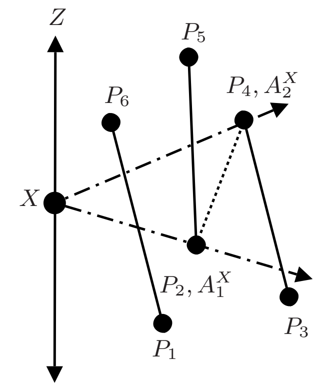

Fig. 10. An example to demonstrate the definitions of $T(X)$, A_1^X, and A_2^X. Here $T(X)$ is the ordered set $\{P_1, P_3, P_2, P_4, P_5, P_6\}$. A_1^X is the last point in $T(X)$ such that there does not exist a P_i and P_j ahead of A_1^X where $\overline{P_iP_j} \in Q$. Here $A_1^X = P_2$. A_2^X is the first point in $T(X)$ such that there exists a P_i before A_2^X where $\overline{P_iA_2^X} \in Q$. Here $A_2^X = P_4$. Note how the dotted line segment $\overline{A_1^XA_2^X}$ has the property $[X]\bigcap(\bigcap_{i=1}^{|Q|}[Q_i]) = [X]\bigcap[\overline{A_1^XA_2^X}]$.

Lemma 5: Assume $X \in M_j$. If A_1^X comes before A_2^X in $T(X)$, then $m\big([M_j]\bigcap(\bigcap_{i=1}^{|Q|}[Q_i])\big) = m([M_j]\bigcap[\overline{A_1^XA_2^X}])$, otherwise $m([M_j]\bigcap(\bigcap_{i=1}^{|Q|}[Q_i])) = 0$.

Intuitively, this lemma says the measure of all lines which intersect M_j and all line segments in Q is the same as the measure of all lines which intersect M_j and some line segment connecting two points in P. *This reduces the problem to finding the measure of the set of lines intersecting two line segments, a problem which we already know how to solve* (see Section II-C).

Proof: Direct result of Lemmas 3 and 4 and the fact $m([\alpha_i]) = 0 \;\; \forall \alpha_i \in \alpha$. ■

Summary:

In steps 1 and 2 we place a vertical line Z and partition it into a set of line segments M. In step 3 we show $m(\bigcap_{i=1}^{|Q|}[Q_i]) = \sum_{M_j\in M} m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[M_j]\big)$. Steps 4 and 5 when when combined with a lemma about rays (see appendix in [24]) show how to compute $m\big((\bigcap_{i=1}^{|Q|}[Q_i])\bigcap[M_j]\big)$ in constant time assuming we know A_1^X and A_2^X. For a given $X \in M_j$, $T(X)$ can be computed in polynomial time by sorting the angles between $\overline{XP_i}$ and Z for all i. A_1^X and A_2^X can then be found by enumerating through $T(X)$. Since $|M|$ is polynomial, this allows us to calculate $m(\bigcap_{i=1}^{|Q|}[Q_i])$ in polynomial time.

We believe the complexity of this algorithm can be reduced by going through all M_j 'in order,' thus eliminating the need to sort P for all M_j in M.

IV. Evaluating Network Reliability

In this section we introduce and show how to evaluate some performance metrics with respect to a random line-cut. After introducing our network model, we show that every performance metric can be evaluated in polynomial time. In particular, we can evaluate average two-terminal reliability in polynomial time under 'random' line-cuts. This is in contrast to the case of independent link failures for which there exists no known polynomial time algorithm to calculate this reliability metric.

A. Network Model

We start by describing our network model. Our geometric graph model contains a set of nodes N where each node is represented by a point on the plane. We assume the node locations are in general form; that is no three are collinear. A link between two nodes is represented by a line segment with endpoints at the respective node locations. In order to assign probabilities to random line events, we assume the set which contains all nodes and links ($conv(N)$) is a subset of some bounded closed convex set C with perimeter L_C. If a 'random' line that intersects C also intersects some links, those links are disrupted. Our goal is to evaluate the performance metrics described below in Definition 5 after a single random line-cut that intersects C.

B. Performance Metrics

We first introduce some network performance metrics and then describe how to evaluate each one after the removal of the intersected links. We will use tools from geometric probability to evaluate average values of these metrics with respect to a random line-cut.

Definition 5: [Performance Metrics]

- *ATR* - The all terminal reliability of the network. The all terminal reliability is defined here as 1 if the network is connected and 0 otherwise. In order to verify connectivity of the network a Breadth First Search algorithm can be used.
- *A2TR* - The average two terminal reliability of the network over all pairs of nodes. The two terminal reliability between two nodes is defined here as 1, if there is a path between them and 0, otherwise [28]. Effectively this metric is the probability a randomly chosen pair of nodes is connected. If the network is fully connected, the value of *A2TR* is 1. Otherwise, we have to sum over the number of node pairs in every connected component and divide it by the total number of node pairs in the network. That is, we sum the value of $k(k-1)$ over every connected component, where k is the number of nodes in each of the components, and then divide this sum by $N(N-1)$. This ratio gives the fraction of node pairs that are connected to each other. In order to verify connectivity or to count the number of nodes in each connected component a Breadth First Search algorithm can be used.

In this paper we only discuss the above two metrics. However, the following relevant metrics can also be evaluated with respect to random line-cuts using the results of this section.

- *TC* - The total capacity of the intersected links.
- *MFST* - The maximum flow between a given pair of nodes s and t.
- *AMF* - The average value of maximum flow between all pairs of nodes.

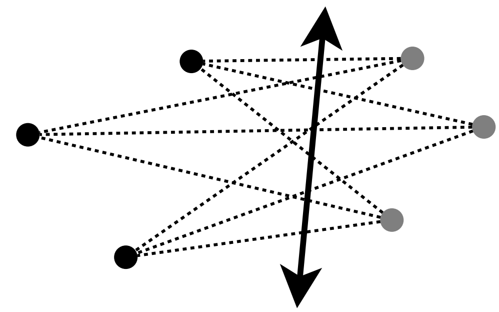

Fig. 11. Consider a line-partition of a set of nodes, N, into two non-empty subsets in the figure above. One subset has nodes colored grey and the other has nodes colored black. A line separates N into these subsets iff it intersects every dashed line segment connecting a grey node and black node.

It is apparent from the descriptions above that evaluating each metric after the removal of intersected links takes polynomial time in $|N|$.

C. Evaluation of the Metrics

We now show how to evaluate the metrics in Definition 5 with respect to a random-line cut. The basic idea is that every line which separates the nodes in the same way removes the same set of links. Using the techniques in Section III, we calculate the measure of the set of lines that separate the nodes in this way; this allows us to calculate the weighted average of a metric over all possible cuts. We start by introducing some useful terminology.

Definition 6 (Line-partition): A line-partition is a partition of a set of nodes into two subsets which are separated by a line. It is important to notice that not all partitions of N are line-partitions.

Let P be the set of all line-partitions created by a lines that intersect $conv(N)$. For each line-partition p in P, let $[p]$ be the set of all lines which form the line-partition p. For a particular p, let the set of all line segments connecting a node in one subset to a node in the other subset be given by Q_p.

Lemma 6: $m([p]) = m(\cap_{q\in Q_p}[q])$ for every $p \in P$

Proof: If a line intersects every line segment in Q_p, then it separates the endpoints of the line segments in Q_p into subsets that form p or it intersects a node. On the other hand, if a line forms a line partition p, then it separates nodes into two subsets and thus will intersect every line segment that has endpoints in both subsets (this is precisely Q_p). See Fig. 11. Thus $[p] = \cap_{q\in Q_p}[q]$ except for a set of lines which intersect nodes. Since the set of lines which intersect nodes has zero measure (points have zero perimeter), the result follows. ■

Now, let $[r]$ be the set of lines that intersect C but not $conv(N)$. That is, $[r] = [C]\backslash[conv(N)]$. Thus, $m([r]) = m([C]) - m([conv(N)])$ by countable additivity of measures.

Note that $(\cup_{p\in P}[p]) \cup [r] = [C]$ up to a set of measure zero. Now, since every line which forms the same line-partition removes the same links, evaluating the performance measures to a random line-cut becomes a weighted average over each partition. Let $Y(p)$ be the performance metric on the network

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

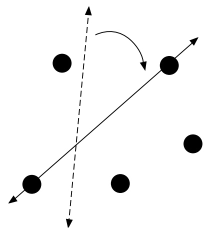

Fig. 12. Consider a line-cut resulting in a line-partition in which neither subset of nodes is empty. Now rotate this line-cut clockwise until nodes prevent any further clockwise movement (imagine that the line cannot pass through the nodes).

when links that intersect a line in [p] are removed. Since $\frac{m([p])}{L_C}$ is the probability a random line-cut will create a partition p, the performance metric to a random line-cut can be expressed as

$$\frac{m([r])}{L_C}Y(r) + \sum_{p \in P} \frac{m([p])}{L_C}Y(p) =$$
$$\frac{L_C - L_{conv(N)}}{L_C}Y(r) + \sum_{p \in P} \frac{m(\bigcap_{q \in Q_p}[q])}{L_C}Y(p) \quad (1)$$

Section III shows how to calculate $m(\bigcap_{q \in Q_p}[q])$ in polynomial time. The performance metrics in Definition 5 can be calculated in polynomial time as discussed above. In the following, we will show that $|P|$ is $O(|N|^2)$.

Lemma 7 ([13]): There are $O(|N|^2)$ line-partitions of a set of $|N|$ nodes.

Proof: [13] shows there are $\binom{|N|}{2} + 1$ line-partitions of a set of $|N|$ points, no three of which are collinear. ■

We will now provide some intuition behind the above result. Consider a line that forms a line-partition in which neither subset of nodes is empty (the line intersects $conv(N)$). Now rotate this line clockwise until nodes prevent any further clockwise movement (imagine that the line cannot pass through the nodes). There will be two points stopping the line from moving any further. Now, these two points specify this partition, and since there are $\binom{|N|}{2}$ ways to pick two nodes, there are $\binom{|N|}{2}$ partitions (see Fig. 12). The additional partition comes from the case when the line does not intersect $conv(N)$.

Theorem 1: Evaluating any performance metric in Definition 5 with respect to a random line cut takes polynomial time in $|N|$.

Proof: Since $|P|$ is polynomial in $|N|$ and evaluating $m(\bigcap_{q \in Q_p}[q])$ and $Y(p)$ takes polynomial time, Equation 1 can be evaluated for any performance metric in polynomial time. ■

This is particularly interesting for the case of $A2TR$ because there is no known polynomial algorithm to find $A2TR$ assuming independent link failures [7]!

V. Numerical Results to Random Line-cuts

We first present an example that demonstrates the significance of geometry on the survivability of the network. We then find $A2TR$ of a real-world network to a random line-cut.

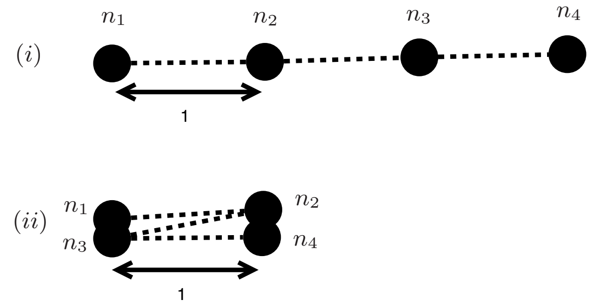

Fig. 13. Two networks of 4 nodes connected in serial by line segments of length 1. The network in case (i) resembles a line segment of length 3 and the network in case (ii) resembles a line segment of length 1.

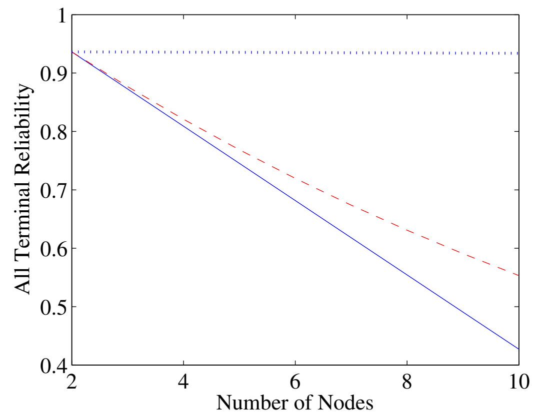

Fig. 14. Assuming $L_C = 10\pi$, the figure shows ATR versus number of nodes for different network configurations. The solid line represents the network which resembles a line segment of length $|N| - 1$, the dotted line represents the network which resembles a line segment of length 1, and the dashed line represents independent link failures.

A. An Example to Demonstrate the Importance of Geometry

In this example, every link has a length of one, so every link is intersected by a random line-cut with equal probability. We consider different geometries of the same network and evaluate ATR to random line-cuts. For comparison, we also evaluate ATR assuming independent links failures.

Consider a network of $|N|$ nodes connected in serial by line segments of length 1. We consider two different cases of geometries for this network. In case (i) the network resembles a line segment of length $|N| - 1$, and in case (ii) the network resembles a line segment of length 1 (see Fig. 13). Assuming $L_C = 10\pi$ and letting $|N|$ vary (assuming $|N| \leq 10$), we calculate ATR to random line-cuts in both cases using methods described in Section IV. Also, since any particular link of length 1 fails with probability $\frac{2}{10\pi}$ with respect to a random line-cut, we evaluate ATR when links fail independently with probability $\frac{2}{10\pi}$.

Fig. 14 shows the results. In case (i), ATR is approximately $1 - \frac{2(|N|-1)}{10\pi}$ since ATR is 1 if any link is intersected and 0 otherwise. In case (ii), ATR is approximately $1 - \frac{2}{10\pi}$ for all $|N|$ (again, since ATR is 1 if any link is intersected and

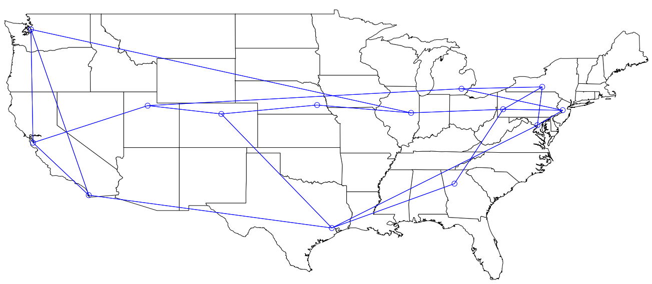

Fig. 15. NSFNET circa 1991 [21].

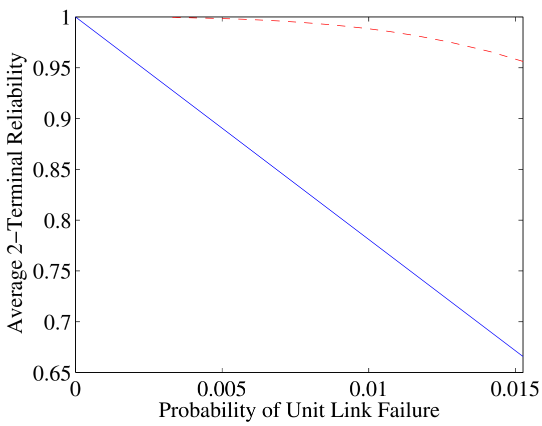

Fig. 16. The solid line shows $A2TR$ in NSFNET versus the probability a unit (latitude/longitude) of fiber is cut by a random line. The dashed line shows $A2TR$ in NSFNET assuming links fail *independently* such that links fail with the same probability as in the random line-cut case.

0 otherwise). When links fail independently with probability $\frac{2}{10\pi}$, ATR is given by $Pr(\text{no links fail}) = (1 - \frac{2}{10\pi})^{|\mathrm{N}|-1}$. Note this value lies between the results for the two geometric networks.

These results agree with intuition. In case (i) when the network resembles a line segment of length $|N| - 1$, the network is spread out and the probability of any link being intersected is larger when $|N|$ is larger. In case (ii) when the network resembles a line segment of length 1, the network has a perimeter of about 2 and thus the probability of any link being intersected is small. This example highlights the importance of node location on the survivability of the network (see Fig. 14).

B. A Real-World Example

In this example we consider NSFNET as found in 1991 [21] (see Fig. 15). This network has 14 nodes and connects major universities across the U.S.. We assume the network is contained within a convex set with perimeter L_C.

Using the results of Section IV, we calculate $A2TR$ of NSFNET to random-line cuts while L_C varies. Given the length of a particular link, the probability that link is cut is proportional to $\frac{2}{L_C}$. So we can plot $A2TR$ versus $\frac{2}{L_C}$. Note $\frac{2}{L_C}$ is the probability a unit of fiber is cut (since a unit of fiber has a perimeter of 2). See Fig. 16 for results. Note the linear form of the result in the figure; this agrees with Equation 1.

Next, we calculated $A2TR$ of NSFNET assuming independent link failures such that links fail with the same probability as in the random line-cut case. Thus the probability a link fails is still proportional to its length, however links fail independently. Since the total number of links is small, calculating $A2TR$ by enumerating all possible failures is still feasible (possible failures are exponential in number of links). See Fig. 16 for results. Note that $A2TR$ in the independent failure model is greater than in the case of random line-cuts. Perhaps this is because at least two links must fail independently to disconnect the network, however a line which intersects the network is guaranteed to disconnect it. Since most backbone networks are likely to be well connected, we expect a random line-cut to lead to lower $A2TR$ than independent link failures in the real-world setting.

VI. Network Design Under Random Line-cuts

In this section we present some network design problems in the context of random line-cuts. In all the proposed problems the location of every node is fixed; the problem is to find a set of links most robust to some metric under some constraints.

(i) Let N be a set of nodes fixed on the plane. As before, assume all links are represented as line segments between the points. A reasonable goal is to design a connected network with the least expected number of links cut by a random line. By linearity of expectation, the expected number of links cut is proportional to $\sum_{i \in Q} l_i$ where Q is the set of links and l_i is the length of link i. So, this problem reduces to minimizing the total length of links in the network while ensuring the graph is connected. This is equivalent to finding a Euclidean minimum spanning tree of N which can be done in polynomial time. Note however that the resulting network is not robust because a single link failure will disconnect it.

(ii) We next consider ring networks. Let N be a set of nodes fixed on the plane. A reasonable goal is to design a connected ring network with the least expected number of links cut by a random line. As before, the expected number of links cut is proportional to $\sum_{i \in Q} l_i$ where Q is the set of links and l_i is the length of link i. So this problem reduces to finding the minimum length cycle that visits every node exactly once and returns to the starting node. This is equivalent to the Euclidean traveling salesman problem which is hard to compute [16].

(iii) The final problem considers how to connect two nodes such that the path between them is robust to random-line disasters. More precisely, let N be a set of nodes fixed on the plane. Let $S \in N$ and $T \in N$ such that $S \neq T$. Let Q be a set of links. The problem is to find a path from S to T consisting of links from Q that has the minimum probability of being cut. This may correspond to finding the most robust path between two cities along preexisting conduits. Since the probability a random line will intersect the path is proportional to the perimeter of the convex hull of the path (a line intersects the path iff it intersects the convex hull of the path), we want to find a path such that the perimeter of the convex hull of the path is minimized. That is, we want to find a path with

minimum perimeter convex hull such that the path starts at S and ends at T, its edges belong to Q, and it contains no repeated vertices. The authors do not know a polynomial time algorithm to solve this problem (except in a trivialized setting).

VII. Conclusions

Motivated by applications in the area of network robustness and survivability, in this paper, we focused on the problem of geographically correlated network failures. Namely, we focused on randomly located geographical attacks on the network which can model the 'random' nature of a natural disaster or collateral damage. In particular, we focused on random line-cuts. Using tools from geometric probability we demonstrated how to compute failure probabilities and showed how to calculate $A2TR$ and other network performance metrics in polynomial time under this failure model. This is in contrast to the independent failure model because there does not exist a known polynomial time algorithm to calculate $A2TR$. We then presented some numerical results to demonstrate the significance of geometry on the survivability of the network and also discussed network design problems in the context of random line failures.

Our approach provides a fundamentally new way to look at network survivability that takes into account the geographical correlation between links. Some future research directions include the consideration of non-line cuts (e.g., circular cuts), multiple line-cuts (instead of a single line failure), and robust network design in the face of geographical failures.

Acknowledgments

This work was supported by NSF grants CNS-0626781 and CNS-0830961, DTRA grants HDTRA1-07-1-0004 and HDTRA-09-1-005, and the NDSEG Fellowship Program.

References

[1] R. Ambartzumian, "Combinatorial solution of the Buffon Sylvester problem," *Probability Theory and Related Fields*, vol. 29, no. 1, pp. 25–31, 1974.

[2] R. Ambartzumian and A. Baddeley, *Combinatorial Integral Geometry: With Applications to Mathematical Stereology*. John Wiley & Sons, 1982.

[3] S. Bar, M. Gonen, and A. Wool, "A geographic directed preferential Internet topology model," in *Proc. IEEE MASCOTS'05*, Sept. 2005.

[4] A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, October 1999.

[5] R. Bhandari, *Survivable networks: algorithms for diverse routing*. Kluwer, 1999.

[6] J. Borland, "Analyzing the Internet collapse," *MIT Technology Review*, Feb. 2008. [Online]. Available: http://www.technologyreview.com/Infotech/20152/?a=f

[7] T. Brecht and C. Colbourn, "Lower bounds on two-terminal network reliability." *Discrete Appl. Math.*, vol. 21, no. 3, pp. 185–198, 1988.

[8] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Resilience of the Internet to random breakdowns," *Phys. Rev. Lett.*, vol. 85, pp. 4626–4628, Nov. 2000.

[9] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Breakdown of the Internet under intentional attack," *Phys. Rev. Lett.*, vol. 86, no. 16, pp. 3682–3685, Apr 2001.

[10] P. Francis, S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, and L. Zhang, "IDMaps: A global internet host distance estimation service," *IEEE/ACM Trans. Netw.*, vol. 9, no. 5, pp. 525–540, October 2001.

[11] L. K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, and S. Havlin, "Stability and topology of scale-free networks under attack and defense strategies," *Phys. Rev. Lett.*, vol. 94, no. 18, 2005.

[12] O. Gerstel and R. Ramaswami, "Optical layer survivability: a services perspective," *IEEE Commun.*, vol. 38, no. 3, pp. 104–113, Mar. 2000.

[13] E. Harding, "The number of partitions of a set of n points in k dimensions induced by hyperplanes," *Proceedings of the Edinburgh Mathematical Society, Series II*, vol. 15, no. 285-289, p. 1967, 1966.

[14] 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.

[15] A. Lakhina, J. Byers, M. Crovella, and I. Matta, "On the geographic location of Internet resources," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 934–948, 2003.

[16] E. Lawler, A. Rinnooy-Kan, J. Lenstra, and D. Shmoys, *The traveling salesman problem: a guided tour of combinatorial optimization*. John Wiley & Sons Inc, 1985.

[17] R. P. Loukas Lazos and J. A. Ritcey, "Detection of mobile targets on the plane and in space using heterogeneous sensor networks," *Wireless Networks*.

[18] D. Magoni, "Tearing down the Internet," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 949–960, Aug. 2003.

[19] J. Manchester, D. Saha, and S. K. Tripathi (eds.), "Protection, restoration, and disoster recovery," *IEEE Network, Special issue*, vol. 18, no. 2, Mar.–Apr. 2004.

[20] A. Mathai, *An introduction to geometrical probability*. CRC Press, 1999.

[21] Merit Networks, Network Map. [Online]. Available: http://www.merit.edu/networkresearch/projecthistory/nsfnet/nsfnet%5Fmaps.php

[22] E. Modiano and A. Narula-Tam, "Survivable lightpath routing: a new approach to the design of WDM-based networks," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 4, pp. 800–809, May 2002.

[23] A. Narula-Tam, E. Modiano, and A. Brzezinski, "Physical topology design for survivable routing of logical rings in WDM-based networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 8, pp. 1525–1538, Oct. 2004.

[24] S. Neumayer and E. Modiano, "Network reliability with geographically correlated failures," Massachusetts Institute of Technology, EE, Tech. Rep. 2009-07-31, Jul. 2009. [Online]. Available: http://www.mit.edu/%7Ebastian/INFOCOM10-Tech-Report-2009-07-31.pdf

[25] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the impact of geographically correlated network failures," in *IEEE Military Communications Conference, 2008. MILCOM 2008*, 2008, pp. 1–6.

[26] ——, "Assessing the vulnerability of the fiber infrastructure to disasters," in *Proc. IEEE INFOCOM'09*, Apr. 2009.

[27] W. Radasky, "High-altitude electromagnetic pulse (HEMP): A threat to our way of life," *IEEE-USA Today's Engineer*, Sep. 2007. [Online]. Available: http://www.todaysengineer.org/2007/Sep/HEMP.asp

[28] S. Rai and D. P. Agrawal, *Distributed Computing Network Reliability*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1990.

[29] L. Santalo, *Integral Geometry and Geometric Probability*. Cambridge University Press, 2004.

[30] J. Sylvester, "On a funicular solution of Buffon's problem of the needle in its most general form," *Acta Mathematica*, vol. 14, no. 1, pp. 185–205, 1890.

[31] R. Wilhelm and C. Buckridge, "Mediterranean fibre cable cut – a RIPE NCC analysis," Apr. 2008. [Online]. Available: http://www.ripe.net/projects/reports/2008cable-cut/index.html

[32] C. Wilson, "High altitude electromagnetic pulse (HEMP) and high power microwave (HPM) devices: Threat assessments," CRS Report for Congress, Aug. 2004. [Online]. Available: http://www.fas.org/man/crs/RL32544.pdf

[33] W. Wu, B. Moran, J. H. Manton, and M. Zukerman, "Topology design of undersea cables considering survivability under major disasters," in *AINA Workshops*, 2009, pp. 1154–1159.

[34] S. Yook, H. Jeong, and A. Barabasi, "Modeling the Internet's large-scale topology," *Proc. Natl. Acad. Sci.*, vol. 99, no. 21, pp. 13 382–13 386, 2002.

[35] D. Zhou and S. Subramaniam, "Survivability in optical networks," *IEEE Network*, vol. 14, no. 6, pp. 16–23, Nov.-Dec. 2000.