Computer Networks 94 (2016) 14–28

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Network reliability under geographically correlated line and disk failure models☆

Sebastian Neumayer*, Eytan Modiano

Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA

ARTICLE INFO

Article history:
Received 24 February 2015
Revised 20 August 2015
Accepted 26 November 2015
Available online 4 December 2015

Keywords:
Network survivability
Geographically correlated failures
Fiber-optic
Electromagnetic Pulse (EMP)

ABSTRACT

Optical fiber networks consist of fibers that are laid out along physical terrestrial paths. As such, they are vulnerable to geographical physical failures, such as earthquakes and Electromagnetic Pulse (EMP) attacks. Moreover, such disasters can lead to multiple, geographically correlated, failures on the fiber network. Thus, the geographical layout of the fiber infrastructure has a critical impact on the robustness of the network in the face of such geographical physical failures.

In this paper, we develop tools to analyze network connectivity 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 'randomly' located disk or line in a plane. Using results from geometric probability, we are able to calculate certain network performance metrics to such a disaster in polynomial time. In particular, we can evaluate average two-terminal reliability in polynomial time under both 'random' failure models. 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. Finally, we present numerical results that make clear geographically correlated failures are fundamentally different from independent failures. Our novel approach provides a promising new direction for modeling and designing networks to lessen the effects of geographically correlated failures.

1. Introduction

Optical fiber networks consist of fibers that are laid out along physical terrestrial paths. As such, they are vulnerable to geographical physical failures, such as hurricanes and Electromagnetic Pulse (EMP) attacks [22,36,42]. Moreover, such disasters can lead to multiple, geographically correlated, failures on the fiber network. Thus, the geographical layout of the fiber infrastructure has a critical impact on the robustness of the network in the face of such geographical physical failures. In this paper, we evaluate the impact of geographically correlated failures on network connectivity with respect to a randomly located line or disk failure.

Previous works have considered the problem of finding the worst-case location for a geographic failure, represented as a disk or line segment, in a geographic network with respect to certain network connectivity measures [1,3,34]. 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 a randomly located disaster on network connectivity. The random location of the disaster can model failure resulting from a natural

☆ This work was supported by NSF grants CNS-0626781, CNS-1017800, and CNS-0830961, DTRA grants HDTRA1-07-1-0004 and HDTRA-09-1-005, and the NDSEG Fellowship Program. Preliminary and partial versions of this paper appeared in Proc. IEEE INFOCOM'10, Mar. 2010 [31] and Proc. IEEE GLOBECOM'11, Dec. 2011 [32].

* Corresponding author. Tel.: +18152586322.
E-mail addresses: bastian@mit.edu (S. Neumayer), modiano@mit.edu (E. Modiano).

http://dx.doi.org/10.1016/j.comnet.2015.11.025
1389-1286/

disaster such as a hurricane or collateral (non-targeted) damage in an EMP attack.

In the context of the logical topology, many works have studied the effect of large-scale failures [12,13,18,25]. Some of these works consider random link failures by modeling the Internet as a random graph [9] and applying percolation theory. Additionally many papers have studied network reliability (e.g., [10,19,26,45]), however most works that focus on physical topologies consider only a small number of fiber failures (e.g., [29,30]). It has been shown through measurements (e.g., [17]) that factors such as population density influence the topology of the Internet [8,23,44]. However, these papers do not account for geographically correlated failures. In contrast, in this paper we consider disaster events which cause many failures in a certain geographical area (e.g., [6,11,22,36]). Only recently has the impact of geographically correlated failures started to receive attention (e.g. [39], [34], [1]).

Our method is to use geometric probability to assign a measure to sets of lines and disks in the plane that intersect some set of line segments (e.g. network links whose removal would disconnect the network). Using these basic tools which are introduced and explained in Section 2, we are able to calculate network performance metrics to these random cuts in polynomial time.

To the best of our knowledge our work is among the first to apply geometric probability techniques to network survivability. In [43] the survivability of undersea cables with respect to a randomly located disk is studied, however only a two node topology was considered. Also, [24] applied geometric probability techniques to detection in sensor networks, and [41] considers a probabilistic region failure model whose failure region consists of several annuli and has probabilistic effect. For a more comprehensive overview of geographically correlated failure models see [20].

Although the two failure models presented in this paper are idealized and do not capture the non-uniform effects and varying sizes of certain real-world failures, these models can be used as a 'first order' approximation for randomly located disasters.

A notable contribution of this paper is the development of an algorithm to calculate the average two-terminal reliability of a network in polynomial time with respect to non-targeted line or disk failure. This result is significant because calculating this metric assuming independent link failures is known to be NP-hard [7]. Another contribution of this paper is numerical results for some network performance metrics using our algorithms which show the importance of node placement on the survivability of the network. We also present results based on real-world networks that show independent and geographically correlated failures are fundamentally different.

This paper is organized as follows: In Sections 2 and 3 we introduce geometric probability and present algorithms that allow us to evaluate joint link failure probabilities after a random line or disk cut. In Section 4 we use these results to demonstrate how to evaluate average two-terminal reliability under our failure models (among other metrics). In Section 5 we present some numerical results to show the significance of geometry on the survivability of the network and that make clear that geographically correlated failures are fundamentally different from independent failures. We conclude and discuss future research directions in Section 6.

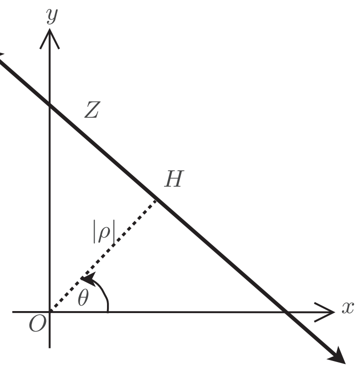

Fig. 1. Let Z be a line in the plane, O be the origin, and H be the closest point on Z to O (see Fig. 1). Note that every line in the plane can be parameterized by ρ and θ where θ is the smallest non-negative angle between the x-axis and the line's normal and $|\rho|$ is the Euclidean distance between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

2. Modeling randomly located disasters in geographic networks

In this section we describe how to model disasters as randomly located lines or disks using geometric probability. We initially focus only on disasters which remove links along a random line. This may model damage to communication infrastructure in a localized area that is the result of a natural disaster (e.g. an earthquake). We then shift focus to randomly located disks of a particular radius. The circular form of the attack may better model the effect of large storms that affect large areas or collateral damage in an EMP or bomb attack. After introducing some basic definitions and tools from geometric probability, we review classical results which allow us to find single link failure probabilities. These results are requisite for Section 3 where we show how to find relevant joint link failure probabilities to a random line or disk failure.

2.1. Random line disaster

Geometric probability is the study of probabilities involved in geometric problems. In our case, we are interested in the probability that a 'randomly' (infinitely long) 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 [40].

Before proceeding further, we will present some useful notation. Let C be a closed bounded convex set on the plane. Denote the perimeter (where perimeter is the length of the boundary) of a set of points in the plane C by L_C and its area by R_C. 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 [27,38]) 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

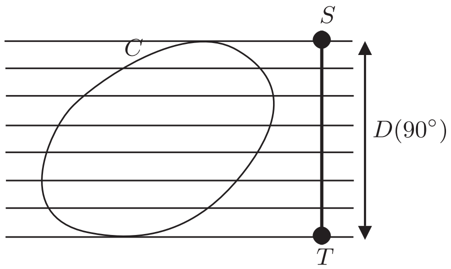

Fig. 2. Consider a set of parallel lines ($\theta = 90^\circ$ but variable ρ) that intersect C. Projecting C onto a line perpendicular to the parallel lines results in line segment $\overline{ST}$. Let $D(90^\circ)$ be the length of this line segment. It seems reasonable for a measure to map the set of parallel lines in the plane with $\theta = 90^\circ$ that intersect C to $D(90^\circ)$. Note that $D(\theta)$ remains unchanged if C is translated.

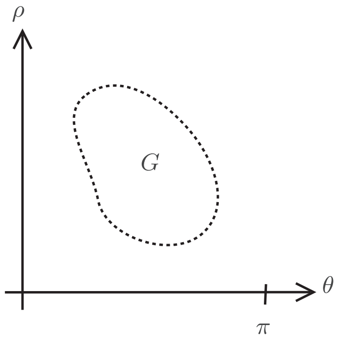

Fig. 3. Since every line in the plane can be parameterized by ρ and θ, we can represent a line in G as a point in the (ρ, θ) plane. Integrating over the set of all these points allows us to assign a measure to the set G.

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 $x \cos\theta + y \sin\theta - \rho = 0$.

Let C be a bounded closed convex set in the plane. We first start by considering a set of parallel lines (fixed θ but variable ρ) that intersect C (see Fig. 2). Projecting C onto a line perpendicular to the parallel lines results in a line segment (see Fig. 2). Let $D(\theta)$ be the length of this line segment.

By considering $D(\theta)$ over all angles, we have $m([C])$ as $\int_0^\pi D(\theta)d\theta$. Note that $D(\theta)$ is invariant under the translation of C and since $D(\theta) = D(\theta + \pi)$ we know $\int_0^\pi D(\theta)d\theta$ is invariant under the translation and rotation of C. We now present the definition of the measure.

Definition 1 (Measure of a set of lines)**.** The measure of a set of lines G is defined as the integral

$$m(G) = \int_G d\rho d\theta$$

Note we use G to denote both a set of lines and its equivalent set of points in the (ρ, θ) plane. In some sense, this integral is the area of G in the (ρ, θ) plane. For a visualization of the measure, see Fig. 3. Since every line in the plane can be parameterized by ρ and θ, we can represent each line in G as a point (ρ, θ) in $\mathbb{R} \times [0, \pi)$.

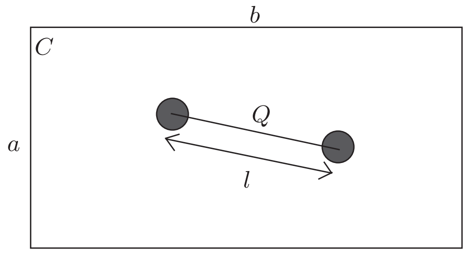

Fig. 4. Rectangle C with link Q inside.

We will now present two examples of evaluating $\int_0^\pi D(\theta)d\theta$. 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 $\int_0^\pi D(\theta)d\theta = \sin\theta|_0^{\frac{\pi}{2}} - \sin\theta|_{\frac{\pi}{2}}^{\pi} = 2$.

We now present an important result from geometric probability.

Lemma 1. *Let C be a bounded closed convex set and $D(\theta)$ be defined the same as above. Now,*

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

Note this is consistent with the above examples (recall L_C is the perimeter of C). See [38, p. 30][27, p. 72] for a justification of the above lemma.

2.1.1. 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 [38, p. 30], $\frac{m([Q])}{m([C])}$. 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 the above 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. 4). Now we consider a random line cut. We have:

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

2.1.2. 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. These results are important because they are used in Section 3 to develop a method for finding the probability a particular set of links fail.

Definition 2 (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 [38, p. 32-34]

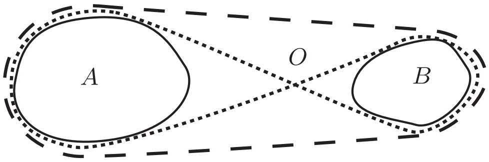

Fig. 5. The dotted curve shows the internal cover of A and B, denoted by $I(A, B)$. The dashed curve shows the boundary of the convex hull of $A \cup B$.

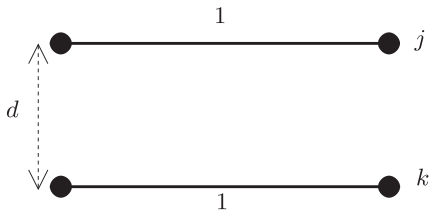

Fig. 6. $Pr($ j and k cut $|$ j cut $) = \sqrt{1+d^2} - d$.

(see Fig. 5). 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, by convention, $L_{I(A,B)} = L_A + L_B$.

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

Lemma 2 ([38]). *If A and B are bounded closed convex sets,*

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

For a proof, see [38, p. 32–34]. This lemma gives us the measure of the set of lines that intersect two convex shapes (e.g. line segments) in terms of perimeters of related shapes.

In the following, we use the above lemma to find the probability a link is cut (i.e. intersected by a random line) given another particular link is also cut. Given two links, j and k, by definition the probability a 'random' line is in the set $[j] \cap [k]$ when it is 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. 6. 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. 7 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. See Fig. 8. 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 | j cut $)$

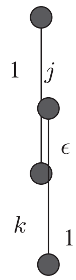

Fig. 7. $Pr($ j and k cut $|$ j cut $) = \epsilon$.

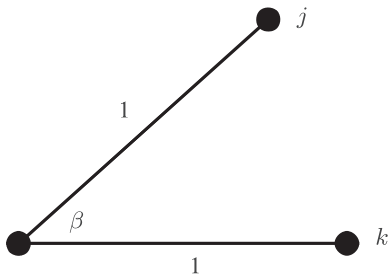

Fig. 8. $Pr($ j and k cut $|$ j cut $) = 1 - \sin\frac{\beta}{2}$.

$= 1 - \sin\frac{\beta}{2}$. 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.

2.2. Random circular disaster

We now shift focus from the random line failure model to a randomly located disk failure model. We are interested in the probability that a 'randomly' placed disk (of a particular radius) in the plane will intersect a certain set of links. We model a failure event in the network as a *single* randomly located disk of a radius r_b.

Before proceeding further, we will present some useful notation. Given a set in the plane, let $\langle \cdot \rangle$ denote the set of all disks in the plane of radius r_b that intersect it.

Geometric probability tells us how to assign a measure to sets of disks; let this measure be denoted by μ. The rest of this section reviews results from geometric probability (see [27,38]) that are necessary for the development of this work.

Note that every disk in the plane of radius r_b can be parameterized by the location of its center. Denote the center of disk D as $[x_D, y_D]$ (see Fig. 9). Let D_O be the disk of radius r_b centered at the origin.

We now present the definition of the measure μ.

Definition 3 (Measure of a set of disks)**.** The measure μ of a set of disks G is defined as the integral

$$\mu(G) = \int_G dxdy$$

Note we use G to denote both a set of disks and the set of centers of these disks. This integral is the area of the set of centers of the disks in G and will be denoted by $area(G)$. This definition appeals to intuition; in the same way the 'size' of a

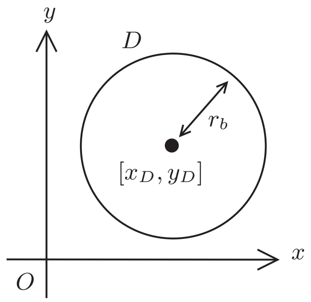

Fig. 9. Every disk in the plane of radius r_b can be parameterized by the location of its center. Denote the center of disk D as $[x_D, y_D]$.

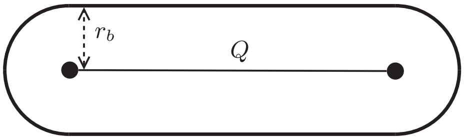

Fig. 10. The shape above represents $Q \oplus D_O$. This shape is known as a hippodrome and it represents the set of all points whose shortest distance to Q is less than or equal to r_b [16]. Denote the hippodrome corresponding to a link Q and radius r_b by $H(Q, r_b)$. Note that a disk D of radius r_b intersects Q iff $[x_D, y_D] \in H(Q, r_b)$.

set of points in the plane is its area, the 'size' of a set of disks is the area of the disks centers.

Definition 4 (Minkowski sum)**.** The Minkowski sum of two sets in the plane A and B in Euclidean space, denoted by $A \oplus B$, is given by

$$A \oplus B = \{a + b | a \in A, b \in B\}$$

Intuitively, every point in the Minkowski sum $C \oplus D_O$ (where C is a convex set) represents a center of a disk of radius r_b that intersects C. We will now discuss an important example. Let Q be a line segment link; consider $Q \oplus D_O$ (see Fig. 10). This shape is known as a hippodrome and it represents the set of all points whose distance to Q is less than or equal to r_b [16]. Denote the hippodrome corresponding to a link Q and radius r_b by $H(Q, r_b)$. Note that a disk D of radius r_b intersects Q iff $[x_D, y_D] \in H(Q, r_b)$.

Lemma 3. *[38, pp. 80–123] Let C be a bounded closed convex set of points in the plane, then*

$$\mu(\langle C \rangle) = \int_{\langle C \rangle} dxdy = area(C \oplus D_O) = R_C + L_C r_b + \pi {r_b}^2$$

(again, where R_C is the area of C). Intuitively, every point in the Minkowski sum $C \oplus D_O$ represents a center of a unique disk of radius r_b that intersects C. Integrating over the set of centers of these disks yields the measure of $\langle C \rangle$. For example, consider a line segment link Q of length d. Now the measure of the set of disks of radius r_b that intersect Q is $\mu(\langle Q \rangle) = area(Q \oplus D_O) = area(H(Q, r_b)) = 2\,dr_b + \pi {r_b}^2$

Let $\langle Q \rangle$ and $\langle C \rangle$ be sets of disks of radius r_b in the plane such that $\langle Q \rangle \subset \langle C \rangle$. Similar to the random line case, given μ, the probability that a 'random' disk is in the set $\langle Q \rangle$ given it is in the set $\langle C \rangle$ is given by the ratio $\frac{\mu(\langle Q \rangle)}{\mu(\langle C \rangle)}$. Note that C contains the centers of all possible disk failures and is required for normalization purposes.

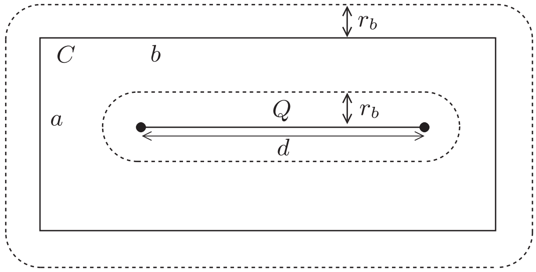

Fig. 11. The rectangle C and line segment link Q are shown in solid. The center of every disaster that intersects Q is given by the dotted hippodrome $H(Q, r_b)$. The center of every disaster that intersects C is given by $C \oplus D_O$ which is shown as the larger dotted shape.

Table 1
Table of geometric probability notation.

L_C	Perimeter of convex set C
R_C	Area of convex set C
$[C]$	The set of lines that intersect convex set C
$\langle C \rangle$	The set of disks of a particular radius whose center lies in convex set C
$conv(A)$	The convex hull of set A
$I(A, B)$	The internal cover of convex sets A and B
$A \oplus B$	The Minkowski sum of sets A and B
$m()$	A measure on sets of lines. Given a convex set C, this measure has the property $m([C]) = L_C$
$\mu()$	A measure on sets of disks. Given a convex set C, this measure has the property $\mu(\langle C \rangle) = R_C + L_C r_b + \pi {r_b}^2$ where r_b is the radius of disks considered

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 d inside C (see Fig. 11). Now we consider a random disk-cut. We have:

$$\Pr(Q \text{ cut} | C \text{ cut}) = \frac{\mu(\langle Q \rangle)}{\mu(\langle C \rangle)} = \frac{area(Q \oplus D_O)}{area(C \oplus D_O)} = \frac{2\,dr_b + \pi {r_b}^2}{ab + 2(a+b)r_b + \pi {r_b}^2} \tag{1}$$

In Table 1 we summarize the descriptions of geometric probability notation introduced so far.

3. Geographically correlated link failures

In this section we present algorithms that calculate the measure of disks or lines intersecting only a particular set of links. This result will allow us to calculate the probability that a randomly located disaster intersects a certain set of links in a network (e.g. links whose removal would disconnect the network). We will then use these algorithms to efficiently calculate network performance metrics with respect to random failures. The details of this section may be skipped and the reader may proceed without loss of continuity to Section 4 on evaluating network reliability.

3.1. Random line disaster

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

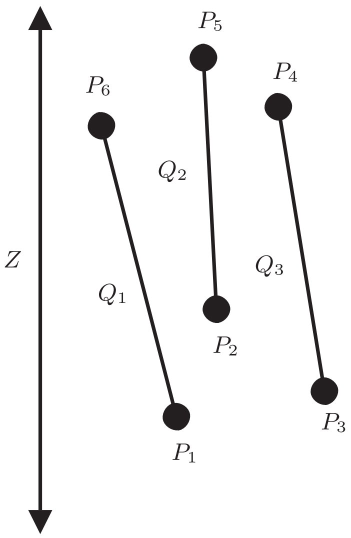

Fig. 12. Shown above are three line segments, Q_1, Q_2, and Q_3, and a line Z which does not intersect the convex hull of P. We want to find $m([Q_1]\bigcap[Q_2]\bigcap[Q_3])$ which is equivalent to $m(Z\bigcap[Q_1]\bigcap[Q_2]\bigcap[Q_3])$.

endpoints are collinear[1]. Let the ith 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 [40] shows how to solve for $m(\bigcap_{i=1}^{|Q|}[Q_i])$. However, their 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 [4] and [5] 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 done in Section 2.1.2.

In the following, for clarity of presentation we break down our procedure into steps for finding $m(\bigcap_{i=1}^{|Q|}[Q_i])$.

Step 1:

In this step we will introduce a line on the plane and partition it into a set of line segments and rays (a ray is a subset of a line consisting of a point and all points on that line which extend infinitely in one direction away from that point). We then find an alternate expression for $m(\bigcap_{i=1}^{|Q|}[Q_i])$ which include these line segments and rays.

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

We start by arbitrarily placing a vertical line Z such that it does not intersect the convex hull of P (see Fig. 12). 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.

Now, 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. 13). Let us impose an ordering on M and denote M_i to be the ith segment in M.

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

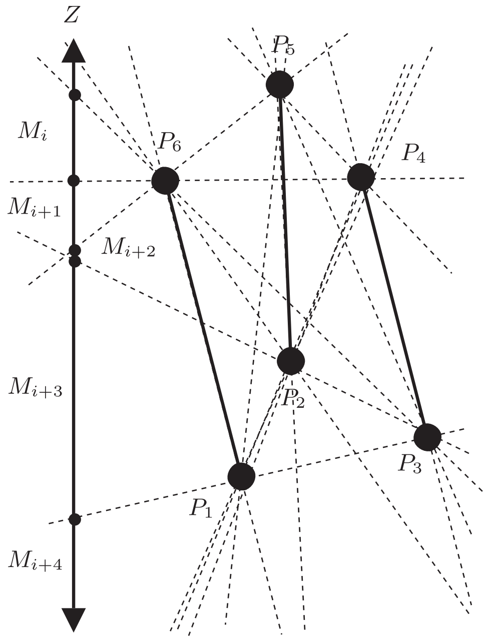

Fig. 13. Consider all lines that intersect two points in P. Let the intersection points of these lines and Z be denoted by α (shown as dots on Z above). Let the divisions of Z into line segments and rays by points in α be denoted by M. Let us impose an ordering on M and denote M_i to be the ith segment in M.

Now,

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

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

$$\begin{aligned} & m\left(\bigcup_{M_j\in M}\left(\left(\bigcap_{i=1}^{|Q|}[Q_i]\right)\bigcap[M_j]\right)\right) \\ & \quad = \sum_{M_j\in M} m\left(\left(\bigcap_{i=1}^{|Q|}[Q_i]\right)\bigcap[M_j]\right) \end{aligned}$$

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.

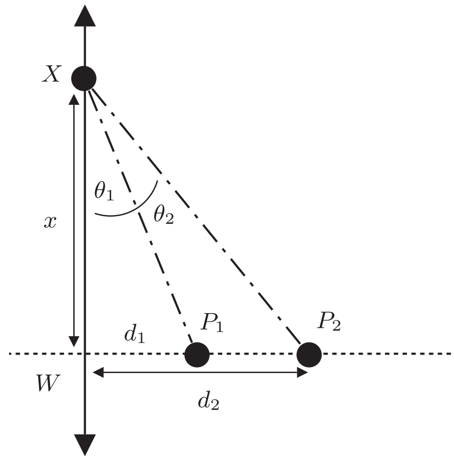

Fig. 14. This is the setup for the proof of Lemma 4. Here X is above W, so $\theta_1 = \arctan(\frac{d_1}{x})$ and $\theta_2 = \arctan(\frac{d_2}{x})$.

Step 2:
In this step we show for all points on an M_j, the ordering of points P (in a rotational sense) is the same.

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

Definition 5.. ***(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. 15.

Lemma 4. *$T(X) = T(X')$ 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.

Proof. We want to show the ordering of $T(X)$ is constant for all X in any M_j. This is equivalent to showing the pairwise ordering in $T(X)$ is constant for all X in any M_j.

We will use geometry techniques to prove the pairwise ordering in $T(X)$ is constant. In order to do this we will present some notation. Consider two different points in P, P_1 and P_2. Assume the line that contains P_1 and P_2 intersects Z (the proof is trivial otherwise) and denote this intersection point by W. Let the distance between W and P_1 be given by d_1 and the distance between W and P_2 be given by d_2. Without loss of generality, assume $d_1 < d_2$. Let X be a point on Z such that $X \neq W$. Let x denote the distance between X and W. Let θ_1 be the angle Z must rotate counter-clockwise about X to intersect P_1 and θ_2 be the angle to intersect P_2. Realizing arctan is a strictly monotonically increasing function, from geometry we have:

$$\theta_1 = \arctan\left(\frac{d_1}{x}\right) < \arctan\left(\frac{d_2}{x}\right) = \theta_2 \;\forall\, X \text{ above W}$$

$$\theta_1 = \pi - \arctan\left(\frac{d_1}{x}\right) > \pi - \arctan\left(\frac{d_2}{x}\right) = \theta_2 \;\forall\, X \text{ below W}$$

See Fig. 14 for clarification.

Now let $X' \in M_j$ for some j such that $X' \neq W$. Now, the equations above imply: P_1 comes before P_2 in both $T(X)$ and $T(X')$ or P_1 comes after P_2 in both $T(X)$ and $T(X')$. Because the above holds for any two different P_1 and P_2, this completely specifies the ordering of $T(X)$ and $T(X')$ and also implies $T(X) = T(X')$. □

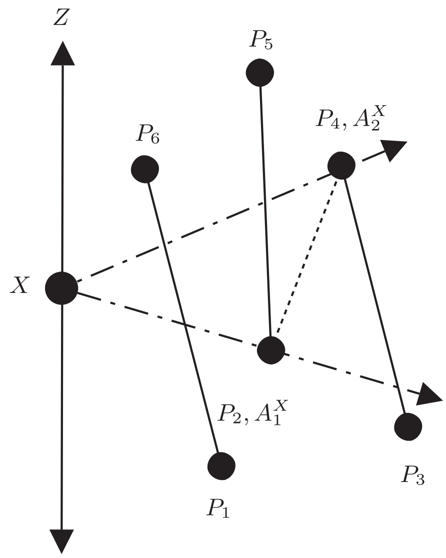

Fig. 15. An example to demonstrate the definitions of $T(X)$, A_1^X, and A_2^X in steps 2 and 3. Here $T(X)$ is the ordered set $\{P_1, P_3, P_2, P_4, P_5, P_6\}$. A_1^X is the last point in $T(X)$ such that there does not exist a P_i and P_j ahead of A_1^X where $\overline{P_iP_j} \in Q$. Here $A_1^X = P_2$. 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 $[X] \bigcap (\bigcap_{i=1}^{|Q|}[Q_i]) = [X] \bigcap [\overline{A_1^XA_2^X}]$.

Step 3:
In this step we show 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 Lemma 2).

Definition 6.. **(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 7.. **(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. 15 for an example.

Lemma 5. *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 that 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. 15 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 $\overline{P_2P_4}$.

Proof. We first use the definitions of A_1^X and A_2^X to find the angles of lines which intersect X and every $Q_i \in Q$. Then conditioned on the ordering of A_1^X and A_2^X in $T(X)$, we use this set of angles to prove the lemma.

We first introduce some useful terminology. Let V_θ denote the line which is a counter-clockwise rotation of Z about X by θ degrees. Let V_{θ_1} denote the line which intersects both A_1^X and X and let V_{θ_2} denote the line which intersects both A_2^X and X. The definition of A_1^X implies θ_1 is the smallest θ such that for every $Q_i \in Q$ there exists a V_θ with $\theta \le \theta_1$ such that Q_i is intersected. Intuitively, θ_1 is the smallest angle θ such that V_θ can intersect every $Q_i \in Q$. The definition of A_2^X implies θ_2 is the largest θ such that for every $Q_i \in Q$ there exists a V_θ with $\theta \ge \theta_2$ such that Q_i is intersected. Intuitively, θ_2 is the largest angle θ such that V_θ can intersect all $Q_i \in Q$. Since θ_1 is the smallest angle θ such that V_θ can intersect every $Q_i \in Q$ and θ_2 is the largest angle θ such that V_θ can intersect all $Q_i \in Q$, this implies V_θ intersects every $Q_i \in Q$ iff $\theta_1 \le \theta \le \theta_2$.

If we assume A_1^X comes before A_2^X in $T(X)$, this implies $\theta_1 \le \theta_2$. Note from geometry, we know a line subsects angle $\angle A_1^X X A_2^X$ iff this line intersects X and $\overline{A_1^X A_2^X}$; thus $\theta_1 \le \theta \le \theta_2$ iff V_θ intersects $\overline{A_1^X A_2^X}$. Since V_θ intersects every $Q_i \in Q$ iff $\theta_1 \le \theta \le \theta_2$, this implies $[X] \bigcap (\bigcap_{i=1}^{|Q|}[Q_i]) = [X] \bigcap [\overline{A_1^X A_2^X}]$.

If we assume A_2^X comes before A_1^X in $T(X)$, this implies $\theta_2 \le \theta_1$. Since V_θ intersects every $Q_i \in Q$ iff $\theta_1 \le \theta \le \theta_2$, this implies if $\theta_2 < \theta_1$ no line intersects X and every $Q_i \in Q$. Also, if $\theta_2 = \theta_1$ only the line V_{θ_1} intersects X and every $Q_i \in Q$. In this case, since V_{θ_1} intersects A_1^X and A_2^X, V_{θ_1} is equivalent to $B_{A_1^X, A_2^X}$. $\square$

Lemma 6. *Assume $X \in M_j$. If A_1^X comes before A_2^X in $T(X)$, then $m\big([M_j] \bigcap (\bigcap_{i=1}^{|Q|}[Q_i])\big) = m([M_j] \bigcap [\overline{A_1^X A_2^X}])$, otherwise $m([M_j] \bigcap (\bigcap_{i=1}^{|Q|}[Q_i])) = 0$.*

Proof. Direct result of Lemmas 4 and 5 and the fact $m([\alpha_i]) = 0 \ \ \forall \alpha_i \in \alpha$. $\square$

Summary:
In step 1 we place a vertical line Z and partition it into a set of line segments and rays M. We then show $m(\bigcap_{i=1}^{|Q|}[Q_i]) = \sum_{M_j \in M} m\big((\bigcap_{i=1}^{|Q|}[Q_i]) \bigcap [M_j]\big)$. Lemma 6 along with a lemma about rays (see the technical report [33]) shows 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.

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.

3.2. Random circular disaster

Let $\mathcal{L}$ be the set of all line segment links in the network and C be a convex polygon that contains $\mathcal{L}$. Consider some set of links $K \subset \mathcal{L}$. We wish to find the measure of all disks of radius r_b that intersect C and every link in K but intersect no links in $\mathcal{L} - K$. See Fig. 16 for an example. This measure is given by $\mu\big(\langle C\rangle \cap (\cap_{k \in K}\langle k\rangle) - \cup_{q \in (\mathcal{L}-K)}\langle q\rangle\big)$. It is clear that a disk D belongs to this measured set iff (i) $[x_D, y_D] \in C \oplus D_O$, (ii) $[x_D, y_D] \in H(k, r_b) \forall k \in K$, and (iii) $[x_D, y_D] \notin H(q, r_b) \ \forall q \in (\mathcal{L} - K)$. So, this measure can also be written as $area\big((C \oplus D_O) \cap (\cap_{k \in K} H(k, r_b)) - \cup_{q \in (\mathcal{L}-K)} H(q, r_b)\big)$. For ease of presentation we abuse notation and denote this measure by $area(K)$.

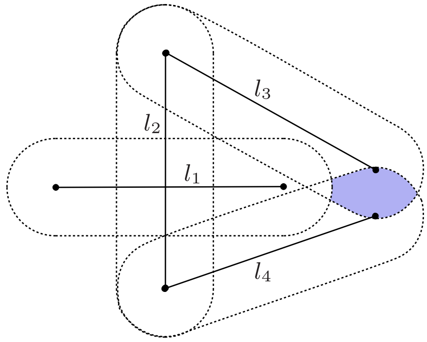

Fig. 16. The measure of disks that intersect C, l_3, and l_4 but neither l_1 or l_2 is given by the area of the shaded region above (for ease of presentation we do not picture C). This measure can be written as $\mu(\langle C\rangle \cap \langle l_3\rangle \cap \langle l_4\rangle - \langle l_1\rangle \cup \langle l_2\rangle)$ or alternatively as $area((C \oplus D_O) \cap H(l_3, r_b) \cap H(l_4, r_b) - H(l_1, r_b) \cup H(l_2, r_b))$.

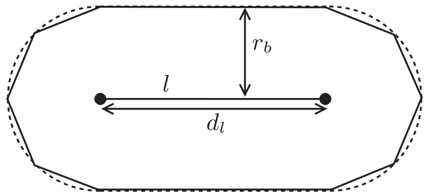

Fig. 17. We approximate $H(l, r_b)$, shown as a dashed hippodrome above, by the inscribing polygon $\hat{H}_n(l, r_b)$ such that $H(l, r_b)$ shares the line segment portion of its boundary with $\hat{H}_n(l, r_b)$ and each end of $\hat{H}_n(l, r_b)$ forms half of a regular $2n$-sided polygon. The solid polygon above is $\hat{H}_4(l, r_b)$.

Definition 8.. ***(area(K))***. Let area(K) be given by the measure of all disks of radius r_b that intersect C and every link in K but intersect no links in $\mathcal{L} - K$.

We note that finding $area(K)$ exactly is difficult because it requires finding the area of intersections and unions of hippodromes. In the following we describe a method for approximating $area(K)$ which is based on approximating hippodromes by polygons for which there are known methods to calculate intersections, unions, and area. We approximate $H(l, r_b)$ by the inscribing polygon $\hat{H}_n(l, r_b)$ such that $H(l, r_b)$ shares the line segment portion of its boundary with $\hat{H}_n(l, r_b)$ and each end of $\hat{H}_n(l, r_b)$ forms half of a regular $2n$-sided polygon (see Fig. 17). Let $\widehat{area}_n(K)$ be defined the same as $area(K)$ except that every hippodrome is replaced by its polygon approximation. Using techniques for finding the intersection, union, and area of polygons [35], we can find $\widehat{area}_n(K)$ in polynomial time.

Lemma 7. $\lim_{n\to\infty} \widehat{area}_n(K) = area(K) \ \forall \ K \subset \mathcal{L}$

The above lemma shows that $\widehat{area}_n(K)$ is a good approximation for $area(K)$ for large enough n. A proof may be found in the technical report [33].

4. Evaluating network reliability metrics

In this section we introduce and show how to evaluate some performance metrics with respect to the random line and disk failure models. After introducing our network model, we show that our performance metrics of interest can be evaluated in polynomial time. In particular, we can evaluate average two-terminal reliability in polynomial time. This result is significant because calculating this metric under independent link failures is known to be NP-hard [7].

4.1. 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. Let L be the set of all links in the network; we assume every link has a capacity associated with it. 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' disaster that intersects C also intersects some links, those links are disrupted. Our goal is to evaluate the performance metrics described below in Definition 9 after a single random line or disk failure that intersects C.

4.2. 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 the tools developed in Section 3 to evaluate average values of these metrics with respect to a random failure.

Definition 9 (Performance metrics)**.**

- *ATR* - The all terminal reliability of the network. The all terminal reliability is defined to be 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.
- *ATTR* - The average two terminal reliability of the network over all pairs of nodes. The two terminal reliability between two nodes is defined to be 1, if there is a path between them and 0, otherwise [37]. Effectively this metric is the probability a randomly chosen pair of nodes is connected. If the network is fully connected, the value of *ATTR* 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 failures 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.

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

4.3. Evaluation of the metrics under the random line failure model

We now show how to evaluate the metrics in Definition 9 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 3, 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 10 (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 lines that intersects $conv(N)$. 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 (including those which represent links in the network).

Lemma 8. *$m([p]) = m(\cap_{q\in Q_p}[q])$ 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. 18. 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 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

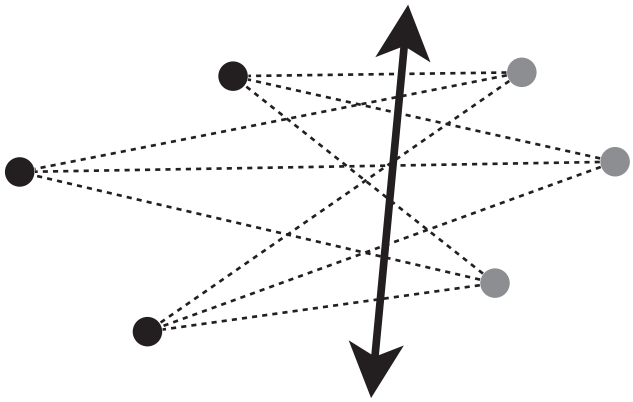

Fig. 18. Consider a line-partition of a set of nodes, N, into two non-empty subsets in the figure above. One subset has nodes colored grey and the other has nodes colored black. A line separates N into these subsets iff it intersects every dashed line segment connecting a grey node and black node.

expressed as

$$\frac{m([r])}{L_C}Y(r) + \sum_{p\in P}\frac{m([p])}{L_C}Y(p)$$
$$= \frac{L_C - L_{conv(N)}}{L_C}Y(r) + \sum_{p\in P}\frac{m(\bigcap_{q\in Q_p}[q])}{L_C}Y(p) \quad (2)$$

Section 3 shows how to calculate $m(\bigcap_{q\in Q_p}[q])$ in polynomial time. The performance metrics in Definition 9 can be calculated in polynomial time as discussed above. In the following, we will show that $|P|$ is $O(|N|^2)$.

Lemma 9 ([21]). *There are $O(|N|^2)$ line-partitions of a set of $|N|$ nodes.*

Proof. Harding [21] 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. 19). The additional partition comes from the case when the line does not intersect $conv(N)$.

Theorem 1. *Evaluating any performance metric in Definition 9 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, Eq. 2 can be evaluated for any performance metric in polynomial time. □

This is particularly interesting for the case of *ATTR* because calculating this metric assuming independent link failures is known to be NP-hard [7]. This is a consequence of the fact not all partitions of N are line-partitions (i.e. there are many more possible partitions under the independent link failure model).

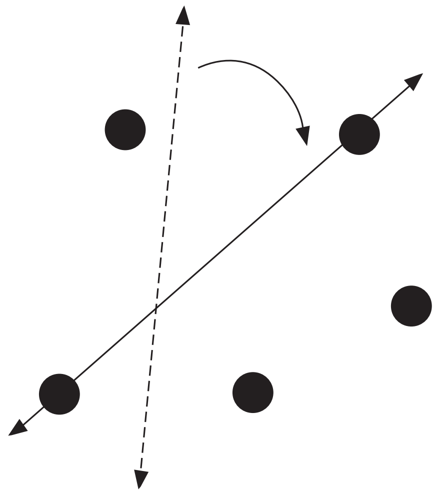

Fig. 19. Consider a line-cut resulting in a line-partition in which neither subset of nodes is empty. Now rotate this line-cut clockwise until nodes prevent any further clockwise movement (imagine that the line cannot pass through the nodes).

4.4. Evaluation of the metrics under the random circular disk failure model

Next we show how to evaluate the metrics above with respect to a random-disk cut of radius r_b. The basic idea is that the center of all disks of radius r_b that intersect a particular set of links (and no other links) is some set in the plane. By showing that the number of these sets we need to consider grows polynomially in N and by evaluating the area of each set, we can evaluate a 'weighted average' of a metric over all possible cuts.

Let U be the set of all subsets of L that can be intersected by exactly one disk of radius r_b. Evaluating performance metrics to a random disk-cut is a weighted average over every $K \in U$. Let $Y(K)$ be a reliability metric evaluated after the removal of every link in K. Since $\frac{area(K)}{area(C\oplus D_O)}$ is the probability a random disk of radius r_b that intersects C also intersects every link in K and no links in $(L-K)$, the performance metric to a

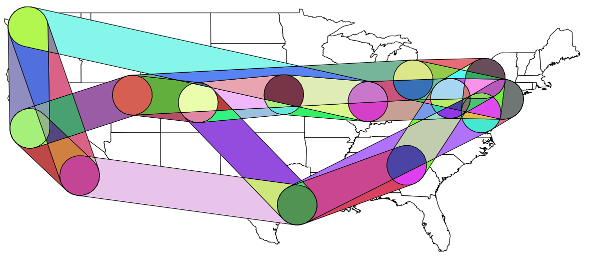

Fig. 20. This figure shows the hippodromes and related intersections with respect to NSFNET [28] and a circular cut of approximately 120 miles.

random disk-cut can be expressed as:

$$\sum_{K \in U} \frac{area(K)}{area(C \oplus D_O)} Y(K) \tag{3}$$

Section 3 shows how to approximate $area(K)$ in polynomial time. $Y(K)$ for the performance metrics can also be calculated in polynomial time. In the following, we apply the theory of arrangements to show that the size of U grows polynomially with respect to N.

Let ∂ denote the boundary of a set. Consider the set of curves $R = \partial C \cup \{\partial H(l, r_b) | l \in L\}$. These curves partition C, the set containing the network, into regions called faces. For example, in Fig. 20 which corresponds to the NSFNET with $r_b = 2$, every colored area represents a different face. By enumerating these faces, we can enumerate every element in U (since every disk centered in a particular face intersects the same links). Arrangements, a computational geometry tool, allow us to enumerate the faces of a set of curves in the plane in polynomial time [15], [2]. However, the theory requires that every pair of curves intersect in a finite number of locations [15] which does not hold in our setting. Nonetheless, the theory can be applied with a minor perturbation to the geometry.

Since enumerating U, evaluating $Y(K)$, and approximating $area(K)$ all take polynomial time, the network performance metrics can be approximated in polynomial time under a random disk failure.

5. Numerical results

In this section we evaluate some network performance metrics using the results of the previous section. We first introduce an example that demonstrates the significance of geometry on the survivability of the network. We then present results based on real-world networks that show that independent and geographically correlated failures are fundamentally different. Lastly, we present results that show the effect of the size of disk and line *segment* failures.

5.1. 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 failure 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 link failures.

Consider a network of $|N|$ nodes connected in series by line segments of length 1. We consider two different cases of

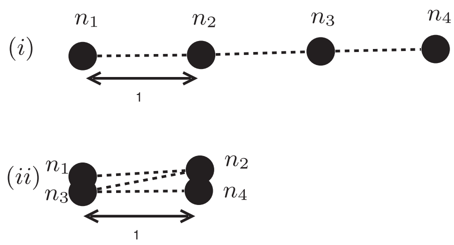

Fig. 21. Two networks of 4 nodes connected in serial by line segments of length 1. The network in case (i) resembles a line segment of length 3 and the network in case (ii) resembles a line segment of length 1.

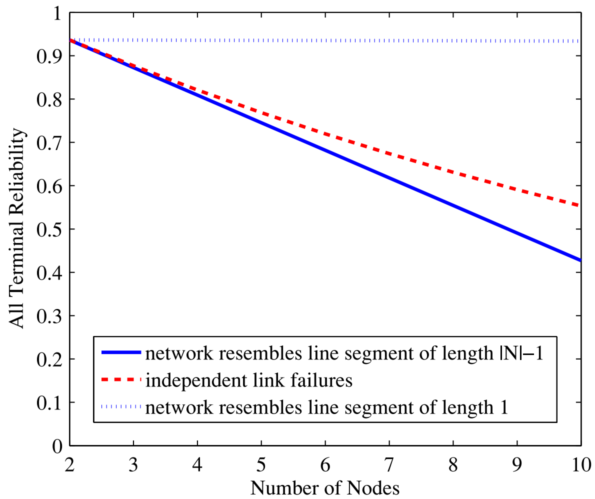

Fig. 22. Assuming $L_C = 10\pi$, the figure shows *ATR* versus number of nodes for different network configurations. The dotted line represents the network which resembles a line segment of length 1, the solid line represents the network which resembles a line segment of length $|N| - 1$, and the dashed line represents independent link failures.

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. 21). 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 4. 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. 22 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 0 otherwise). When links fail independently with probability $\frac{2}{10\pi}$, *ATR* is given by $Pr(\text{no links fail}) = (1 - \frac{2}{10\pi})^{|N|-1}$. 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 2 and thus the probability of any link being

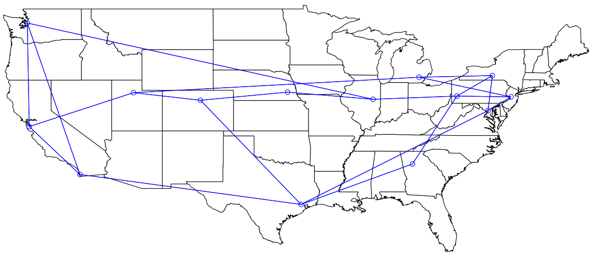

Fig. 23. This figure shows NSFNET from 1991 [28].

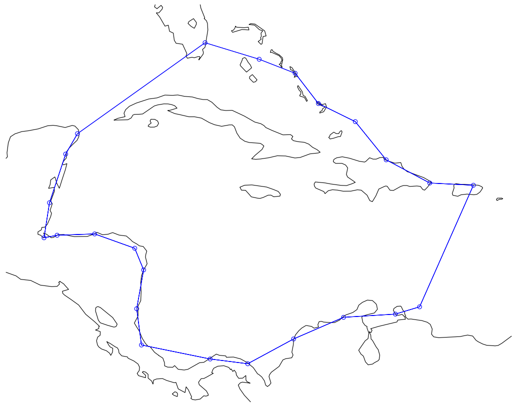

Fig. 24. ARCOS-1 network circa 2009 [14].

intersected is small. This example highlights the importance of node location on the survivability of the network.

5.2. Independent versus correlated failures

We consider NSFNET as found in 1991 [28] and the ARCOS-1 ring network [14]. The NSFNET network we consider has 14 nodes and connects major universities across the U.S. (see Fig. 23). ARCOS-1 has 24 nodes and connects regions on the Dominican Republic, Florida, Mexico, Panama, and Venezuela (see Fig. 24). All distance units mentioned in this section are in longitude and latitude coordinates (one unit is approximately 60 miles) and for simplicity we assume latitude and longitude coordinates are projected directly to [x, y] pairs on the plane. We assume that all the link capacities are equal to 1. We also assume each network is contained within a rectangular set *C*. We note that representing the physical fiber links by line segments is a first-order approximation.

Using the results of Section 4, we calculate *ATTR* of NSFNET and ARCOS-1 to random-disk cuts of $r_b = 2$ while the size of *C* varies. The size of *C* is varied to change the probability a unit of fiber is cut. So we can plot *ATTR* versus the probability a unit of fiber is cut. See Fig. 25 for results. Note the linear form of the result in the figure; this agrees with Eq. 3 since $1/area(C \oplus D)$ is proportional to the probability a unit of fiber is cut.

Next, we calculate *ATTR* of the networks assuming independent link failures such that links fail with the same probability as in the random disk-cut case. Thus the probability a link fails is still a function of its length, however links fail

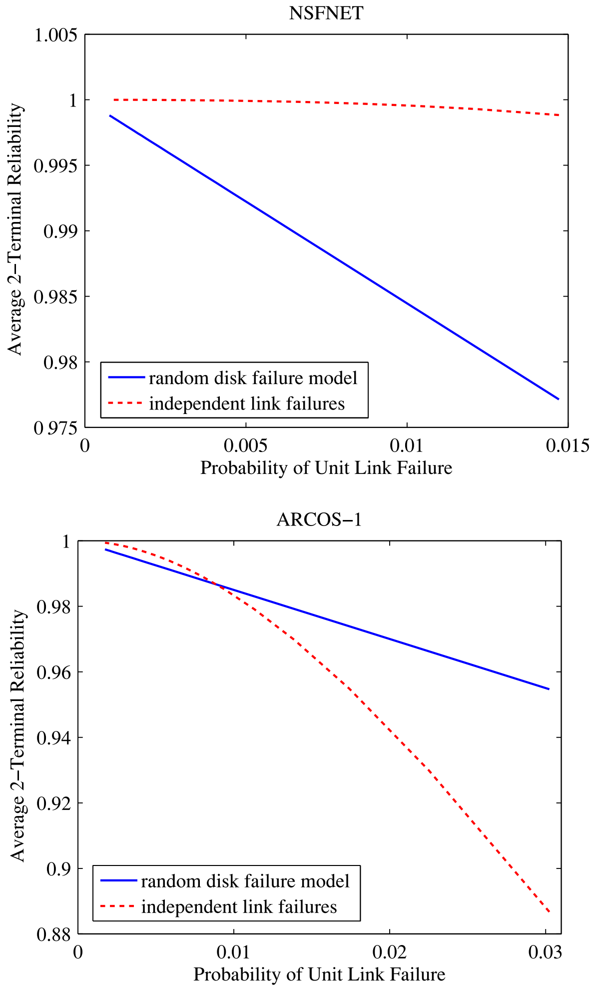

Fig. 25. The solid line shows *ATTR* versus the probability a unit (latitude/longitude) of fiber is cut by a random disk of $r_b = 2$ (approximately 120 miles). The dashed line shows *ATTR* assuming links fail *independently* such that links fail with the same probability as in the random disk-cut case.

independently. Since the total number of links is small in each network, calculating *ATTR* by enumerating all possible failures is still feasible (possible failures are exponential in number of links). Note the total expected number of removed links is the same for both the independent and geographically failure models. See Fig. 25 for results.

Notice that in NSFNET *ATTR* under independent failures is greater than in the case of random disk-cuts. Perhaps this is because in most cases at least three links must fail independently to disconnect the network; however a disk that intersects a node is guaranteed to disconnect the network. Since most backbone networks are likely to be well connected, we expect a random disk-cut to lead to lower *ATTR* than independent link failures in this type of mesh network setting. We also note that similar results were found for the random line-cut setting.

Looking at the results for the ARCOS-1 network we see the opposite tendency; *ATTR* under independent failures is typically less than the case of random disk-cuts. Perhaps this is

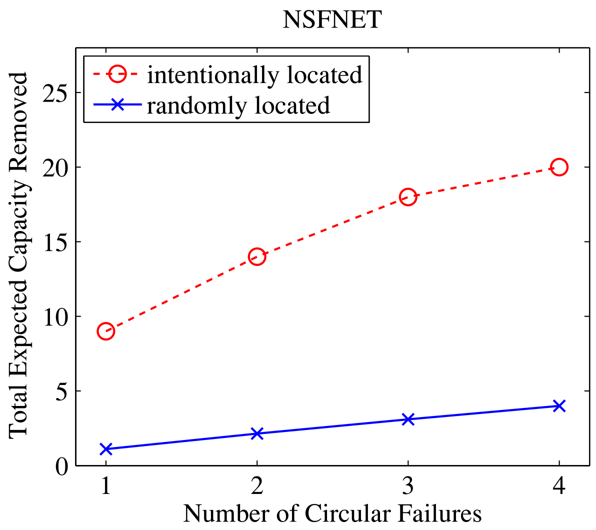

Fig. 26. The solid line shows the total expected capacity removed versus the number of *randomly* located circular disasters of $r_b = 2$. Using the algorithm in [34], the dashed line shows the total capacity removed versus the number of *intentionally* located circular disasters of $r_b = 2$ (approximately 120 miles).

because a single disk that intersects ARCOS-1 usually only removes two adjacent links creating components of size 1 and $|N| - 1$ (where $|N|$ is the number of nodes) whereas just two independent link failures on opposite sides of the ring create components of size $|N|/2$ and $|N|/2$ (which results in lower ATTR).

5.3. Multiple disk failures

Next we calculate the *TC* metric of the NSFNET and ARCOS-1 networks under sequential disk failures, both intentional and random. We assume every additional random failure is located independently of the previous failures. We first describe how to evaluate metrics after sequential failures, then we present some numerical results.

To calculate a network metric after two randomly located sequential disk failures, we simply evaluate the weighted average of the metric over each pair of possible areas (each area represents the set of centers of disks that remove exactly the same links). Eq. 3 then becomes $\sum_{K'\in P}\sum_{K\in P}\frac{area(K')\,area(K)}{area(C\oplus D_O)^2}Y(K\cup K')$. For n failures, Eq. 3 becomes

$$\sum_{K_1\in P}\cdots\sum_{K_n\in P}\left(\prod_{i=1}^{n}\frac{area(K_i)}{area(C\oplus D_O)}\right)Y\Big(\bigcup_{i=1}^{n}K_i\Big).$$

In [34] we propose an algorithm to evaluate network reliability metrics after an *intentional* disk failure. To calculate a network metric after sequential intentional failures, we simply apply the algorithm found in [34] iteratively.

Fig. 26 shows the results for multiple failures, both intentional and random for NSFNET (similar results for ARCOS-1 are not shown). As expected, the plots are sub-linear (the plot for randomly-located failures is barely sub-linear) since each additional failure is being placed on a smaller network. Note that random failures result in much less disruption than intentional failures.

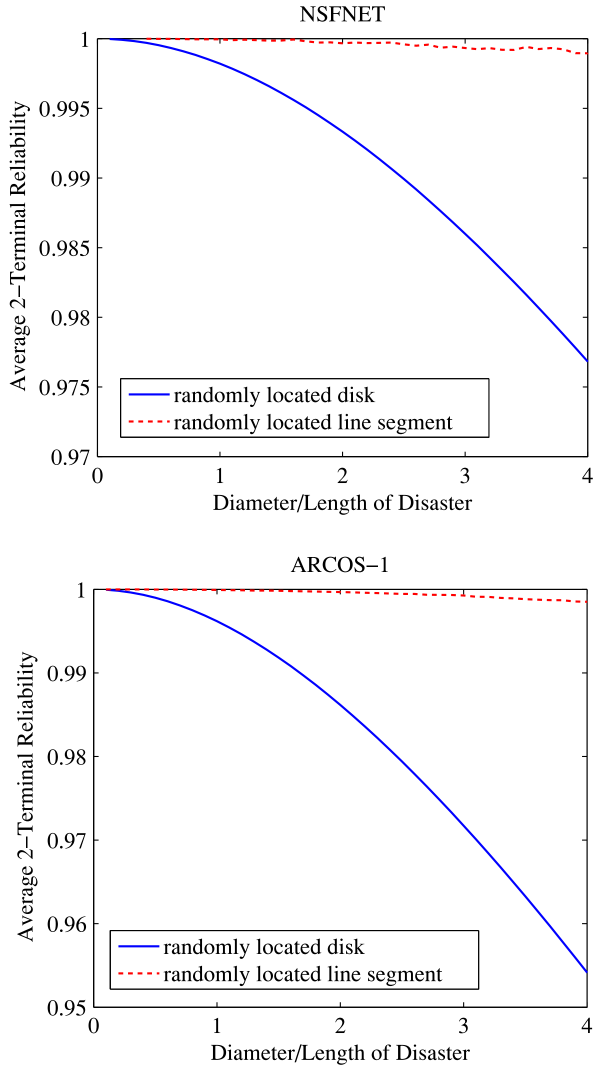

Fig. 27. The solid line shows *ATTR* versus the diameter of a randomly located disk failure event (one unit is approximately 60 miles). The dashed line shows *ATTR* versus the length of a randomly located line segment failure event.

5.4. Effect of size of disk and line segment failures

We now consider the effect of the size of randomly located failure events. We consider two failure models, the randomly located disk failure model and a randomly located line segment failure model. Although the tools to exactly compute the effect of a randomly located line segment disaster have not been developed, we consider this failure model here since it moves away from the idealized infinitely long line failure model assumptions. To compute the effect of a randomly located line segment disaster, we used a Monte Carlo simulation approach and so the associated line segment results are approximate.

See Fig. 27 for results showing the effect of the size of the disasters on the *ATTR* metric. Notice that for both the NSFNET and ARCOS-1 networks, the effect of a randomly located random line segment disaster is small compared to the effect of a

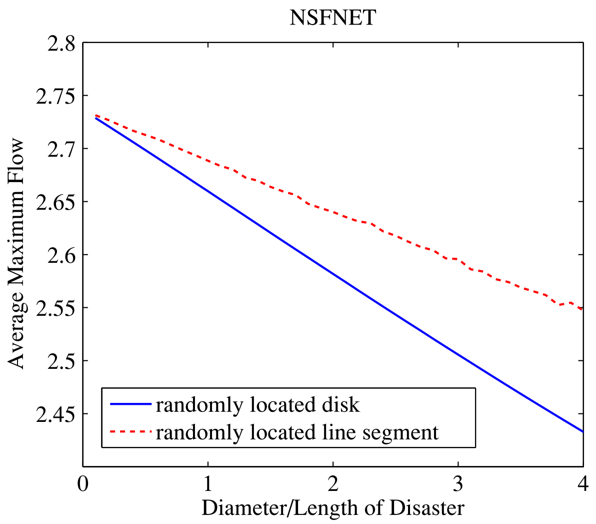

Fig. 28. The solid line shows *AMF* (average maximum flow between all pairs of nodes) versus the diameter of a randomly located disk failure event (one unit is approximately 60 miles). The dashed line shows *AMF* versus the length of a randomly located line segment failure event.

randomly located disk disaster. This agrees with our expectation since a random disk only needs to intersect a node to disconnect the network into at least two components (the node that is intersected is disconnected from the rest of the network), whereas a randomly located line segment has a negligible probability of intersecting a node and must completely 'slice' the network in order to disconnect it.

In Fig. 28 we present similar results for NSFNET with respect to the *AMF* (average maximum flow between all pairs of nodes) metric. Our results show that a randomly located line segment link is more likely to have an effect on the *AMF* metric as compared to *ATTR*. We believe this is because a random line segment that intersects links on the network is likely to affect the maximum flow between some pairs of nodes but may not affect the connectivity of the network. Additionally, we observe random disks of diameter d have a larger impact than a random line segment of length d. This agrees with our expectation since, loosely speaking, a line segment of length d causes less damage than a disk of diameter d.

6. Conclusions

Motivated by applications in the area of network robustness and survivability, 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 and disk cuts. Using tools from geometric-probability we demonstrated how to compute failure probabilities and showed how to calculate *ATTR* and other network performance metrics in polynomial time under these failure models. This result is significant because calculating this metric assuming independent link failures is known to be NP-hard [7]. We then presented some numerical results to demonstrate the significance of geometry on the survivability of the network.

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 multiple line-cuts (instead of a single line failure), convex cuts (e.g., oval cuts), and robust network design in the face of geographical failures.

References

[1] P. Agarwal, A. Efrat, S. Ganjugunte, D. Hay, S. Sankararaman, G. Zussman, The resilience of WDM networks to probabilistic geographical failures, in: Proceedings of IEEE INFOCOM'11, 2011.
[2] M. Sharir, P.K. Agarwal, Arrangements and their applications, in: J.-R. Sack, J. Urrutia (Eds.), Handbook of Computational Geometry, North-Holland, New York, 2000, pp. 49–119.
[3] P. Agarwal, A. Efrat, S. Ganjugunte, D. Hay, S. Sankararaman, G. Zussman, Network vulnerability to single, multiple, and probabilistic physical attacks, in: Proceedings of IEEE MILCOM'10, 2010.
[4] R. Ambartzumian, Combinatorial solution of the Buffon Sylvester problem, Probab. Theory Relat. Fields 29 (1) (1974) 25–31.
[5] R. Ambartzumian, A. Baddeley, Combinatorial Integral Geometry: With Applications to Mathematical Stereology, John Wiley & Sons, 1982.
[6] A. Antony, L. Cittadini, D. Karrenberg, R. Kisteleki, T. Refice, T. Vest, R. Wilhelm, Mediterranean Fiber Cable Cut (January-February 2008) Analysis of Network Dynamics, Technical Report RT-DIA-124-2008, Dept. of Computer Science and Automation, University of Roma Tre, 2008.
[7] M. Ball, Computational complexity of network reliability analysis: an overview, IEEE Trans. Reliab. 35 (3) (1986) 230–239.
[8] S. Bar, M. Gonen, A. Wool, A geographic directed preferential Internet topology model, in: Proceeding of IEEE MASCOTS'05, 2005, doi:10.1109/MASCOTS.2005.5.
[9] A.L. Barabasi, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.
[10] R. Bhandari, Survivable networks: algorithms for diverse routing, Kluwer, 1999.
[11] J. Borland, Analyzing the Internet collapse, MIT Technology Review, 2008 Editorial. [Online]. Available: http://www.technologyreview.com/Infotech/20152/?a=f.
[12] R. Cohen, K. Erez, D. Ben-Avraham, S. Havlin, Resilience of the Internet to random breakdowns, Phys. Rev. Lett. 85 (2000) 4626–4628, doi:10.1103/PhysRevLett.85.4626.
[13] R. Cohen, K. Erez, D. Ben-Avraham, S. Havlin, Breakdown of the Internet under intentional attack, Phys. Rev. Lett. 86 (16) (2001) 3682–3685, doi:10.1103/PhysRevLett.86.3682.
[14] Columbus Networks, Network Map. [Online]. Available: http://www.columbus-networks.com/LandingPoints.html.
[15] H. Edelsbrunner, L. Guibas, J. Pach, R. Pollack, R. Seidel, M. Sharir, Arrangements of curves in the planeùtopology, combinatorics, and algorithms, Automata Lang. Program. (1988) 214–229.
[16] A. Efrat, M. Sharir, A near-linear algorithm for the planar segment-center problem, Discrete Comput. Geometry 16 (3) (1996) 239–257.
[17] P. Francis, S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, L. Zhang, IDMaps: a global internet host distance estimation service, IEEE/ACM Trans. Netw. 9 (5) (2001) 525–540.
[18] L.K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, S. Havlin, Stability and topology of scale-free networks under attack and defense strategies, Phys. Rev. Lett. 94 (18) (2005), doi:10.1103/PhysRevLett.94.188701.
[19] O. Gerstel, R. Ramaswami, Optical layer survivability: a services perspective, IEEE Commun. 38 (3) (2000) 104–113, doi:10.1109/35.825647.
[20] M.F. Habib, M. Tornatore, F. Dikbiyik, B. Mukherjee, Disaster survivability in optical communication networks, Comput. Commun. 36 (6) (2013) 630–644.
[21] E. Harding, The number of partitions of a set of n points in k dimensions induced by hyperplanes, Proc. Edinb. Math. Soc. Ser. II 15 (285–289) (1966) 1967.
[22] 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 (2004).
[23] A. Lakhina, J. Byers, M. Crovella, I. Matta, On the geographic location of Internet resources, IEEE J. Sel. Areas Commun. 21 (6) (2003) 934–948.
[24] L. Lazos, R. Poovendran, J.A. Ritcey, Detection of mobile targets on the plane and in space using heterogeneous sensor networks, Wirel. Netw. 15 (5) (2009) 667–690, doi:10.1007/s11276-007-0092-z.
[25] D. Magoni, Tearing down the Internet, IEEE J. Sel. Areas Commun. 21 (6) (2003) 949–960, doi:10.1109/JSAC.2003.814364.
[26] J. Manchester, D. Saha, S.K. Tripathi (Eds.) Protection, restoration, and disaster recovery, IEEE Netw. 18 (2) (2004) special issue.

[27] A. Mathai, An Introduction to Geometrical Probability, CRC Press, 1999.
[28] Merit Networks, Network Map. [Online]. Available: http://www.merit.edu/networkresearch/projecthistory/nsfnet/nsfnet%5Fmaps.php.
[29] E. Modiano, A. Narula-Tam, Survivable lightpath routing: a new approach to the design of WDM-based networks, IEEE J. Sel. Areas Commun. 20 (4) (2002) 800–809.
[30] A. Narula-Tam, E. Modiano, A. Brzezinski, Physical topology design for survivable routing of logical rings in WDM-based networks, IEEE J. Sel. Areas Commun. 22 (8) (2004) 1525–1538.
[31] S. Neumayer, E. Modiano, Network reliability with geographically correlated failures, in: Proceedings of IEEE INFOCOM'10, 2010.
[32] S. Neumayer, E. Modiano, Network reliability under random circular cuts, in: Proceedings of IEEE GLOBECOM'11, 2011.
[33] S. Neumayer, E. Modiano, Network Reliability Under Geographically Correlated Line and Disk Failure Models, Technical Report, 2015-08-19, Massachusetts Institute of Technology, EE, 2015. http://www.mit.edu/%7Ebastian/GeographicalNetworkReliability-Tech-Report-2015-08-19.pdf
[34] S. Neumayer, G. Zussman, R. Cohen, E. Modiano, Assessing the vulnerability of the fiber infrastructure to disasters, IEEE/ACM Trans. Netw. 19 (6) (2011) 1610–1623.
[35] J. O'Rourke, Computational Geometry in C, Cambridge University Press, 1994.
[36] W. Radasky, High-altitude Electromagnetic Pulse (HEMP): A Threat to Our Way of Life, IEEE-USA Today's Engineer, 2007 http://te.ieeeusa.org/2007/Sep/HEMP.asp.
[37] S. Rai, D.P. Agrawal, Distributed Computing Network Reliability, IEEE Computer Society Press, Los Alamitos, CA, USA, 1990.
[38] L. Santalo, Integral Geometry and Geometric Probability, Cambridge University Press, 2004.
[39] A. Sen, B. Shen, L. Zhou, B. Hao, Fault-tolerance in sensor networks: a new evaluation metric, in: Proceedings of IEEE INFOCOM'06, 2006.
[40] J. Sylvester, On a funicular solution of Buffon's ôproblem of the needle in its most general form, Acta Math. 14 (1) (1890) 185–205.
[41] X. Wang, X. Jiang, A. Pattavina, Assessing network vulnerability under probabilistic region failure model, in: Proceedings of 2011 IEEE 12th International Conference on High Performance Switching and Routing (HPSR), IEEE, 2011, pp. 164–170.
[42] C. Wilson, High Altitude Electromagnetic Pulse (HEMP) and High Power Microwave (HPM) Devices: Threat Assessments, (CRS Report for Congress) 2004. [Online]. Available: http://www.fas.org/man/crs/RL32544.pdf.
[43] W. Wu, B. Moran, J. Manton, M. Zukerman, Topology design of undersea cables considering survivability under major disasters, in: Proceedings of WAINA'09, 2009.
[44] S. Yook, H. Jeong, A. Barabasi, Modeling the Internet's large-scale topology, Proc. Natl. Acad. Sci. 99 (21) (2002) 13382–13386.
[45] D. Zhou, S. Subramaniam, Survivability in optical networks, IEEE Netw. 14 (6) (2000) 16–23, doi:10.1109/65.885666.

Sebastian Neumayer (M10) received the B.S. degree (with Highest Honors) in electrical engineering from the University of Illinois at UrbanaChampaign in 2004, the M.S. degree in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 2007, and Ph.D. degree in electrical engineering at MIT in 2013. Dr. Neumayer is a recipient of the NDSEG Fellowship.

Eytan Modiano (S90-M93-S00) received his B.S. degree in Electrical Engineering and Computer Science from the University of Connecticut at Storrs in 1986 and his M.S. and Ph.D. degrees, both in Electrical Engineering, from the University of Maryland, College Park, MD, in 1989 and 1992, respectively. He was a Naval Research Laboratory Fellow between 1987 and 1992 and a National Research Council Post Doctoral Fellow during 1992–1993. Between 1993 and 1999 he was with MIT Lincoln Laboratory where he was a project leader for MIT Lincoln Laboratory's Next Generation Internet (NGI) project. Since 1999 he has been on the faculty at MIT; where he is a Professor in the Department of Aeronautics and Astronautics and the Laboratory for Information and Decision Systems (LIDS). His research is on communication networks and protocols with emphasis on satellite, wireless, and optical networks. He is the co-recipient of the Sigmetrics 2006 Best paper award for the paper "Maximizing Throughput in Wireless Networks via Gossiping," and the Wiopt 2005 best student paper award for the paper "Minimum Energy Transmission Scheduling Subject to Deadline Constraints." He is currently an Associate Editor for IEEE/ACM Transactions on Networking, and had served as Associate Editor for IEEE Transactions on Information Theory. He was the Technical Program co-chair for IEEE Wiopt 2006, IEEE Infocom 2007, and ACM MobiHoc 2007.