Computer Networks 77 (2015) 117–127

Contents lists available at ScienceDirect

Computer Networks

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

Geographic max-flow and min-cut under a circular disk failure model

Sebastian Neumayer [a,*], Alon Efrat [b], Eytan Modiano [a]

[a] LIDS, MIT, Cambridge, MA 02139, United States
[b] University of Arizona, Tucson, AZ 85721, United States

ARTICLE INFO

Article history:
Received 30 January 2014
Received in revised form 25 September 2014
Accepted 29 October 2014
Available online 27 November 2014

Keywords:
Network survivability
Min-cut max-flow
Geographically correlated failures
Fiber-optic
Electromagnetic Pulse (EMP)

ABSTRACT

Failures in fiber-optic networks may be caused by natural disasters, such as floods or earthquakes, as well as other events, such as an Electromagnetic Pulse (EMP) attack. These events occur in specific geographical locations, therefore the geography of the network determines the effect of failure events on the network's connectivity and capacity.

In this paper we consider a generalization of the min-cut and max-flow problems under a geographic failure model. Specifically, we consider the problem of finding the minimum number of failures, modeled as circular disks, to disconnect a pair of nodes and the maximum number of failure disjoint paths between pairs of nodes. This model applies to the scenario where an adversary is attacking the network *multiple* times with intention to reduce its connectivity. We present a polynomial time algorithm to solve the geographic min-cut problem and develop an ILP formulation, an exact algorithm, and a heuristic algorithm for the geographic max-flow problem.

1. Introduction

Fibers in optical networks are laid out along physical paths, hence they are susceptible to geographical physical events such as earthquakes and Electromagnetic Pulse (EMP) attacks [6,13]. These types of disasters may lead to multiple geographically correlated failures on the fiber infrastructure. Thus, the survivability of the fiber network is affected by its geographical layout. In this paper, we attempt to account for geographically correlated failures on network connectivity and flow.

Previous works considered the problem of finding the worst-case location for a failure in a geographic network with respect to certain network connectivity measures [1,11]. The impact of a single randomly located disaster on network connectivity is considered in [9,10,16]. In this work we consider the problem of finding the minimum number of failures, modeled as circular disks, to disconnect a pair of nodes and the maximum number of failure disjoint paths between pairs of nodes.

Min-cut and max-flow problems similar to the ones presented here have also received some attention in the literature. Recently [15] considered the problem of a geographic max-flow and min-cut in a wireless network setting. In [8] the problem of finding the maximum number of geographically disjoint paths with minimum total cost is discussed in a continuous setting where paths may be placed anywhere within a polygonal domain. Finally, [2] considers a related problem to the geographic max-flow and min-cut, where failures of nearly arbitrary shape occur at a finite set of candidate locations. Here we take the geography into account by allowing failures to take place at any location, yet restricting the shape of a failure to a geometric disk.

We first consider a geographical variant of the min-cut problem. Given a set of points on the plane, each of which

* Corresponding author.
E-mail addresses: bastian@mit.edu (S. Neumayer), alon@cs.arizona.edu (A. Efrat), modiano@mit.edu (E. Modiano).

http://dx.doi.org/10.1016/j.comnet.2014.10.026
1389-1286/

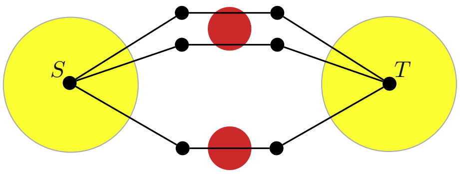

Fig. 1. The light gray area (yellow area in online color version) above represents the protected zone that no circular failure may be centered. The gray disks (red disks in online color version) represent disasters that remove links (of unit capacity) they intersect. Two disasters are required to disconnect the two nodes S and T (shown above), so the geographic min-cut is two. Also, since the top pair of paths can be intersected by the same failure, geographic max-flow is two; two failure disjoint paths are given by the topmost and bottommost path. In contrast, the standard min-cut and max-flow is three. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

represents a node, and non-crossing line segments between these points representing links, what is the minimum number of circular attacks such that two nodes, S and T, are disconnected from each other. If applied to the national fiber plant, the solution to this problem is the number of geographic failures required to disconnect two cities. If we do not restrict the locations of potential failure sites, the geographic min-cut will be at most one because nodes S or T can trivially be eliminated with a single failure. In order to make the problem more relevant and realistic we restrict potential failure locations (see Fig. 1). This can represent fibers that have been hardened against EMP attacks or a well defended city. We show that we only need to consider a polynomial number of possible failure sites, thus reducing the geographic min-cut to a discrete problem. Then applying methods from [2], we show how to find a solution in polynomial time. We obtain numerical results for a specific backbone network [7], thereby demonstrating the applicability of our min-cut algorithm to a real-world network.

Next, in the context of geographic attacks and path-protection algorithms we study a geographic max-flow problem: the largest set of paths between nodes S and T such that no two paths can be intersected by the same failure. The solution to this problem gives the maximum number of paths that are geographically disjoint with respect to disasters of a particular radius. See Fig. 1 for an example. We then develop an ILP formulation, an exact algorithm, and a heuristic algorithm for this geographic max-flow problem.

Finally, we explore the analogue to the min-cut max-flow theorem in the geographic setting. In particular, we show that the cardinality of the solutions to these geographic min-cut and max-flow problems are not the same. Supported by simulation results, we conjecture this difference is no greater than one, i.e. max-flow $\leqslant$ min-cut $\leqslant$ max-flow $+1$.

2. Geographic min-cut

We start by formulating the geographic min-cut problem and presenting an algorithm to solve this problem in polynomial time.

2.1. Network model and problem formulation

Let N be an ordered set of points in the plane representing nodes. Assume the points representing the nodes are in general position, that is no three points are collinear. A link from node i to node j is represented as a line segment in the plane with endpoints at node i and node j. Let the set of undirected links be given by E. We assume that the graph is simple (contains no self-loops or multiple edges) and connected, and links do not intersect each other except at node locations.

We model disasters as disks of radius r_b and refer to these disks as holes. We assume a hole removes all links that intersect it. We also assume a hole may be centered anywhere in the plane, except inside a protective disk of radius r_p centered at nodes S and T.

We now define the following problem and demonstrate its formulation.

Geographical Min-Cut By Circular Disasters (GMCCD) Problem: *Given a graph drawn in the plane $G = \{N, E\}$, two distinct nodes S and T, hole radius r_b, and protection radius r_p, find a minimum cardinality set of holes that disconnect S from T.*

2.2. Algorithm to solve GMCCD problem

We describe an algorithm that finds a solution to the GMCCD problem. For clarity of presentation we break down the algorithm into steps. We initially note that holes may be centered anywhere not inside the protective disks; thus there are an infinitely uncountable number of holes to consider in general. The first step (step 1) of the algorithm reduces this infinitely sized set of potential holes to a polynomial (quadratic) sized set by extending the methods in [11]. Once this set of holes is enumerated, we can apply a simplified algorithm for computing geographic min-cut based on [2]. We do this by first creating a dual-like graph (step 2) and then running an algorithm based on shortest closed walks[1] on this new dual-like graph to solve the GMCCD problem (step 3).

Step 1: There are an infinite number of hole locations centered outside the protective disks; in this step we find a polynomial (quadratic) sized set of holes from which we can construct a solution to the GMCCD problem.

Before proceeding, we introduce some notation. Let $H(e, r_b)$ be the set of points whose shortest distance to line segment e is less than or equal to r_b. Such a shape is known as a hippodrome [5]. Note that a hole of radius r_b is centered in $H(e, r_b)$ if and only if the hole intersects e (see Fig. 2).

In [11] we considered the same failure model without the protected zone. Under this model we found a polynomial (quadratic) sized set of hole locations such that every hole in the plane can be represented by one of these locations and intersects at least the same set of links. For example, any hole centered in the intersection of the two hippodromes in Fig. 3 can be represented by a hole

[1] In this paper we define a closed walk as a sequence of nodes where the first and last node are the same and every pair of consecutive nodes are connected by a link.

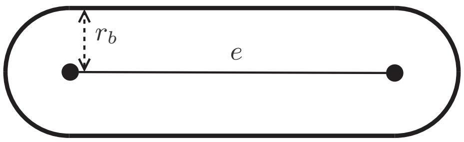

Fig. 2. Let $H(e, r_b)$ be the set of points whose distance to link l is less than or equal to r_b. Such a shape (shown above) is known as a *hippodrome* [5]. A hole of radius r_b is centered in $H(e, r_b)$ if and only if the hole intersects l.

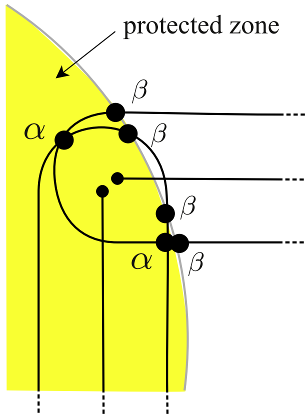

Fig. 3. The light gray area (yellow area in the online color version) above represents the protected zone. α represent centers of some holes given by the algorithm in [11]. These holes intersect both links above, however they are centered in the protected zone. We consider additional holes centered at the points labeled β. Note two of these points correspond to holes that intersect both links and are not centered within the protected zone. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

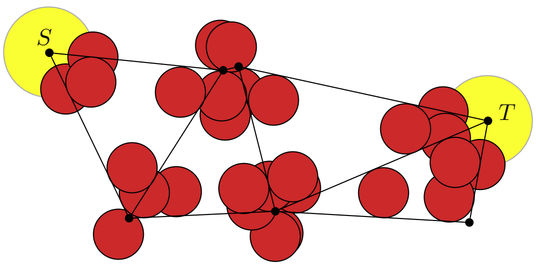

Fig. 4. The red disks above represent most of the holes considered in step 1 of the algorithm for this particular graph (some holes are omitted for clarity). Note that many of these disks intersect a pair of links at exactly one location. This is an effect of the algorithm used to reduce the set of candidate hole locations presented in [11]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

centered on one of the two points labeled α. Holes centered at one of these locations will intersect the same links as any hole centered in the intersection of the hippodromes.

The polynomially sized set of potential failure locations found in [11] cannot be used for the GMCCD problem because of the restrictions that holes cannot be placed inside the protected zones. For example, the set of holes found in [11] would have us consider the holes marked with α in Fig. 3. However, these holes are centered inside the protected zone and cannot be considered. If we consider additional holes that are centered at the intersection of the boundaries of the protected zones and hippodromes (shown by points labeled by β in Fig. 3), we can show that this expanded set of potential failure locations is sufficient. The details can be found in Appendix A. Let this polynomially (quadratic) sized set of potential hole locations for the GMCCD problem be given by U. A subset of potential holes for a particular graph is shown in Fig. 4.

Step 2: We construct an undirected dual-like graph from G, the original graph, and U, the polynomially sized set of potential hole locations. Let this dual-like graph be denoted by K.

We first introduce some notation. The drawing of G in the plane partitions the rest of the plane into connected regions called *faces* (including the outer, infinitely large region). For example, the graph in Fig. 5 divides the plane into five faces, four bounded faces and one infinitely large face.

We now describe the dual-like graph K. Every node in the dual-like graph K corresponds to a face in G. For example, in Fig. 5 G has five faces; each of these faces represents a node in K (shown as dashed circles). There exists a link between two nodes in K for each hole $u \in U$ that intersects the faces they represent. For example, in Fig. 5 there exist two holes intersecting face one and face five, u_1 and u_2. So there exist two links between node one and node five in K; one corresponding to u_1 and the other corresponding to u_2. Note, because every link in K is associated with a hole, there exist more than one edge between two nodes in K if more than one hole intersects their corresponding faces.

Step 3: The final step finds a solution to the GMCCD problem by considering a set of closed walks in K and then from this set finds the shortest walk whose corresponding holes disconnect S from T (see Fig. 6). This is similar to a known algorithm to find the min-cut in a planar graph

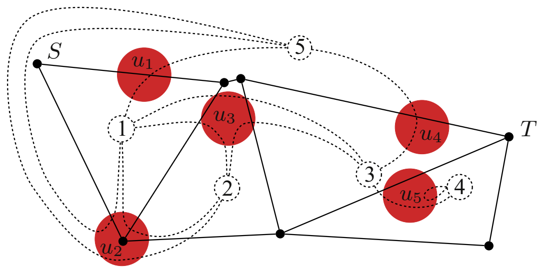

Fig. 5. The dual-like graph is shown by the dotted portion of the figure above. The solid dots and line segments represent the original network G. For ease of presentation, we take the set of gray disks (red disks in the online color version) above to be U. G has five faces; each of these faces represents a node in K (shown as dashed circles). There exists a link between two nodes in K for each hole in U that intersects the faces they represent. Note, there exist two holes intersecting face one and face five, u_1 and u_2. So there exist two links between node one and node five in K; one corresponding to u_1 and the other corresponding to u_2. Also, for presentation purposes the only self-loop in K shown is located at node 4 and corresponds to u_5; there are more self-loops in K (see Fig. 7). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

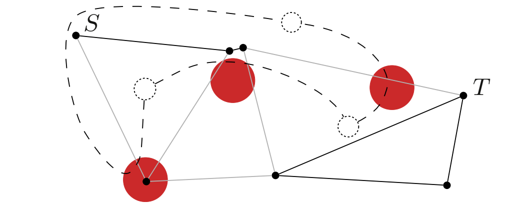

Fig. 6. The dashed links above represents a closed walk in K such that the corresponding holes (shown as disks) remove links which disconnect S and T. By searching over a set of closed walks in K, we will be able to find a solution to the GMCCD problem.

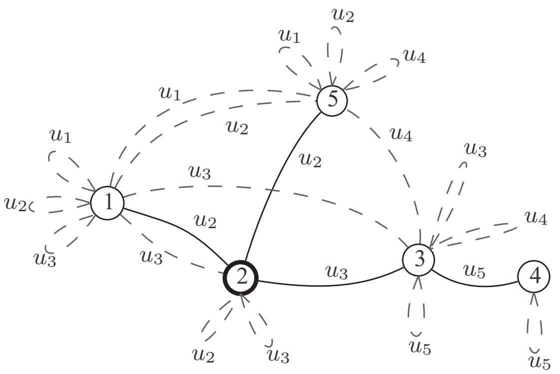

Fig. 7. The graph shown above is K from Fig. 5 complete with self-loops. Every link is marked with its respective hole.

(in the standard sense); the algorithm finds the shortest closed walk in the dual graph that disconnects S from T [14].

We now describe the algorithm. First, for all nodes in the dual-like graph run Dijkstra's algorithm [3]. This gives a shortest path tree rooted at each node. Denote links in this tree for node n by C_n. Notice that when a set of links is removed from the graph new faces are created. Intuitively, a shortest path in K between two nodes gives the minimum number of disasters such that the faces corresponding to these nodes will be contained in a larger face after the disaster. It is worth emphasizing that this face is not necessarily the outer face of the new graph.

Next, for every link e in K consider the closed walk in $C_n \cup e$ which contains node n and link e. A solution to the GMCCD problem is given by finding the closed walk in $C_n \cup e$ for all nodes n and links e in K and then searching over these walks for the shortest one whose corresponding holes disconnect S from T.

For example, consider Fig. 7. Let the link from node n_i to n_j associated with hole u be given by $\{(n_i, n_j), u\}$. The solid links are the links in the shortest path tree rooted at node 2, C_2. Consider the link $\{(1,5), u_2\}$. Now $C_2 \cup \{(1,5), u_2\}$ contains a closed walk given by $\{\{(1,2), u_2\}, \{(2,5), u_2\}, \{(1,5), u_2\}\}$. Since hole u_2 does not disconnect S and T (every hole in this cycle is marked with u_2), $\{u_2\}$ is not a candidate solution. Now consider the link $\{(1,5), u_1\}$. The resulting closed walk is given by $\{\{(1,2), u_2\}, \{(2,5), u_2\}, \{(1,5), u_1\}\}$. Since disasters u_1 and u_2 disconnect S and T, $\{u_1, u_2\}$ is a candidate solution. Enumerating over all nodes and edges in K and finding the minimum cardinality candidate solution solves the GMCCD problem (in this example, a solution is given by $\{u_1, u_2\}$).

Theorem 1. *The algorithm described in steps 1–3 finds a solution to the GMCCD problem.*

Proof. In step 1 we identify a polynomial (quadratic) sized set of locations such that we can find a geographic min-cut considering only holes placed at these locations. Once these locations have been identified the correctness of steps 2 and 3 follow from [2]. □

Let M be the set of nodes in K. As a result of Euler's formula [3] $|N| - |E| + |M| = 2$, we know $|M|$ is linear in $|N|$. Since each hole generates up to $O(|M|^2)$ edges in K (imagine a hole that intersects every face) and there are $O(|M|^2)$ potential holes, we know there are at most $O(|M|^4)$ edges in K. The algorithm creates a shortest path tree for every node in K and considers a closed walk for every node-link pair in K of which there are $O(|M|^5)$. For every closed walk considered, the algorithm explicitly checks if the set of holes corresponding to the closed walk disconnect S and T. A closed walk may correspond to $O(|M|)$ holes and a hole may remove up to $O(|M|)$ links in the original non-dual

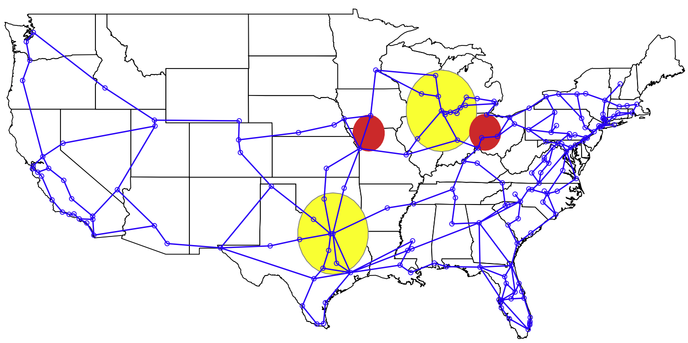

Fig. 8. A solution to the GMCCD problem when $r_b = 1.3$, $r_p = 3.0$, S = Dallas, and T = Chicago. The gray disks (red disks in the online color version) represent the hole locations and the light gray disks (yellow disks in the online color version) represent the protected zones. Only two disasters, located at 'choke' points to the east and west of Chicago, are required to disconnect these cities. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

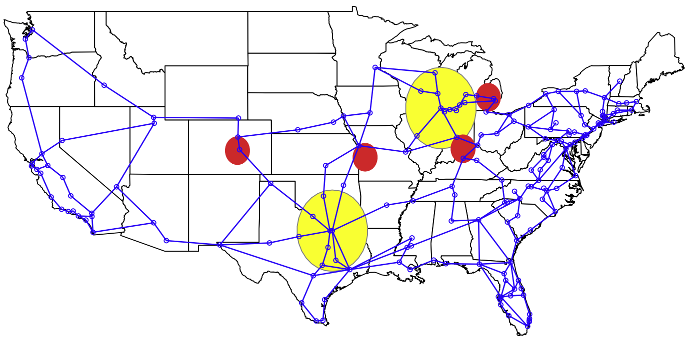

Fig. 9. A solution to the GMCCD problem when $r_b = 1,\ r_p = 3.0$, S = Dallas, and T = Chicago. The gray disks (red disks in the online color version) represent the hole locations and the light gray disks (yellow disks in the online color version) represent the protected zones. Note four disasters with $r_b = 1$ are required to disconnect the two cities, whereas only two disasters are required with $r_b = 1.3$ (see Fig. 8). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

graph. So naively checking *ST* connectivity for each closed walk takes up to $O(|M|^2)$. Thus, the algorithm runs in time at most $O(|M|^7)$.

2.3. Numerical results

We used the algorithm in the previous section to solve the GMCCD problem for a major network provider [7]. We replace every link intersection with a node in this network which allows our algorithm to be applied. All distance units mentioned here 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.

Fig. 8 shows a solution to the GMCCD problem when $r_b = 1.3,\ r_p = 3,\ S$ = Dallas, and T = Chicago. Only two disasters, located at 'choke' points to the east and west of Chicago, are required to disconnect these cities. Fig. 9 shows that when r_b is reduced slightly to 1 a total of four disasters are required to disconnect the two cities.

3. Geographic max-flow

In the context of geographic attacks and path-protection algorithms we consider the geographic max-flow problem: the maximum number of paths between nodes *S* and *T* such that no two paths can be disconnected by the same hole. The solution to this problem gives the maximum number of paths which are geographically disjoint with respect to disks of a particular radius.

In this section we formulate the geographic max-flow problem, develop an exact algorithm as well as a low complexity heuristic algorithm for the problem, and present numerical results based on real-world networks.

3.1. Problem formulation

We use the network and disaster model from the last section.

Geographical Max-Flow By Circular Disasters (GMFCD) Problem: *Given a graph drawn in the plane* $G = \{N, E\}$*, two distinct nodes S and T, hole radius* r_b*, and protection radius* r_p*, find the maximum cardinality set of paths connecting S and T such that no hole intersects a pair of these paths.*

Let *P* be a set of paths from *S* to *T*. Let *H* be the set of all holes in the plane centered outside the open disks of radius r_p centered at *S* and *T* (centered outside the protected zone). The solution to the GMFCD optimization problem below is a geographical max-flow.

$$\begin{aligned} &\max \quad |P| \\ &\text{such that} \quad \nexists h \in\ H \text{ where} \\ &\qquad p_i \cap h \neq \emptyset \text{ and } p_j \cap h \neq \emptyset \ \begin{array}{l} \forall\, p_i \in P \\ \forall\, p_j \in P \end{array},\ i \neq j \end{aligned}$$

We are able to find an ILP formulation of the GMFCD problem with a polynomial number of constraints. The idea for this formulation is to find paths, each with a different 'label', such that each one of these paths obeys some flow constraints and every pair of these paths is failure disjoint.

Denote the cardinality of a solution to the GMCCD problem by *C* and denote the cardinality of a solution to the GMFCD problem by *F*. Note that $F \leqslant C$ since every path in a GMFCD solution must be intersected by a hole in order to disconnect the network and there exists no hole that intersects a pair of paths in a GMCCD solution. Let $A = \{1, 2, \ldots, C\}$. We use this set to limit the number of variables in the ILP formulation.

Define the following {0, 1} variables for all (undirected) links $(i,j) \in E$ and for all $a \in A$:

$$x_{ij}^a = \begin{cases} 1 & \text{if } (i,j) \text{ has label } a \\ 0 & \text{otherwise} \end{cases}$$

We call link (i,j) *active with label a* if $x_{ij}^a = 1$. In the ILP constraints below we ensure that sets of active links with the same label obey flow conservation constraints.

Define the following {0, 1} variables for all nodes $i \in N$ and for all $a \in A$:

$$y_i^a = \begin{cases} 1 & \text{if there exists a node } j \text{ such that } x_{ij}^a = 1 \\ 0 & \text{otherwise} \end{cases}$$

Intuitively, y_i^a is 1 if any active links with label a have an endpoint at i. This variable allows us to write the flow constraints in (3) below.

Define the following {0, 1} constants for all links $(i,j) \in E$ and for all links $(k,l) \in E$:

$$\beta_{kl}^{ij} = \begin{cases} 1 & \text{if } \exists h \in H \text{ that intersects both } (i,j) \text{ and } (k,l) \\ 0 & \text{otherwise} \end{cases}$$

Effectively, $\beta_{ij}^{kl} = 1$ if links (i,j) and (k,l) intersect the same hole. If a pair of active links is intersected by the same hole then constraint (2) below ensures they cannot have different labels, and thus the paths they form are failure disjoint.

The solution to the ILP below is a solution to the GMFCD problem.

$$\max \sum_{a \in A} \sum_{(S,j) \in E} x_{Sj}^a \tag{1}$$

such that

$$x_{ij}^a + x_{kl}^{a'} \leqslant 1 \ \forall \begin{matrix} a \in A \\ (i,j) \in E \end{matrix} \text{ and } \forall \begin{matrix} a' \in A \\ (k,l) \in E \end{matrix} \tag{2}$$
$$\text{where } \beta_{kl}^{ij} = 1 \text{ and } a \neq a'$$

$$\sum_{j:(i,j) \in E} x_{ij}^a = \begin{cases} y_i^a & \text{if } i = S \text{ or } T \\ 2y_i^a & \text{otherwise} \end{cases} \ \forall \begin{matrix} i \in N \\ a \in A \end{matrix} \tag{3}$$

$$x_{ij}^a, y_i^a \in \{0,1\} \qquad \forall i \in N, \quad \forall (i,j) \in E, \quad \text{and } \forall a \in A$$

Constraint (2) above ensures that a pair of active links with differing labels cannot be intersected by the same hole. Note that constraint (2) is generated only for x_{ij}^a and $x_{ij}^{a'}$ pairs where $\beta_{ij}^{kl} = 1$ and $a \neq a'$ (this reduces the total number of constraints), so there must be some offline computation done to find β_{ij}^{kl}.

Constraint (3) consists of flow conservation equations that ensure the total number of active links with a particular label and endpoints is either 0 or 2 except for nodes S and T (0 or 1 for nodes S and T). This ensures active links with a particular label form an ST path (or a cycle not including S or T).

The objective function in the above ILP maximizes the total number of active label-link pairs that have an endpoint at S. Since the flow constraint (3) above ensures an active link with label a and endpoint at S must be part of an ST path consisting of links active with label a and constraint (2) ensures differently labeled links do not interfere, this ILP will give us the maximum number of failure disjoint paths (i.e. a solution to the GMFCD problem). In Section 3.5 we obtain numerical results using this ILP and its relaxations, and compare these results to heuristic algorithms.

3.2. Bounds on C and F

We now present a few bounds on C and F. We first note that $C \neq F$. A simple example demonstrating this is given in Fig. 12. Note in this example $C = 2$ and $F = 1$; a geographic min-cut is given by $\{u_1, u_2\}$ and the max-flow is given by the path corresponding to the dashed curve. This is interesting as it shows the analogue to the max-flow min-cut theorem [3] does not hold in our setting. Also we know that $F \leqslant C$ because every geographic max-flow path must be intersected by a hole in a geographic min-cut or otherwise there would remain a path from S to T after the removal of holes on the min-cut.

We now discuss the relationship between our problems and the ones found in [2]. The max-flow and min-cut problem in [2] differs from the GMFCD and GMCCD problem in two key aspects. First, a hole in [2] need not be a disk; the only requirement is that every hole be homeomorphic to a disk. Second, in [2] holes may only be placed in a finite number of set locations (as opposed to our case where there exists an infinite number of holes outside the protected zones). This is a crucial difference because under the model of [2] some nodes or links may not be intersected by a hole. This means that it is possible for a pair of geographic max-flow paths to intersect each other. In contrast, in the context of our geometric problems, since holes can be centered anywhere on the plane outside the protected zone, we know that a pair of geographic max-flow paths must be node disjoint outside the protected zone (see Fig. 10).

Since only a polynomial number of hole locations need to be considered (as discussed in step 1 of Section 2), it follows that the GMFCD and GMCCD problems are special cases of the geographic max-flow and min-cut problems described in [2]. Thus, some of the results presented in [2] can be applied to our setting. For example, in the special case where S and T share a common face (that is, S and T are both nodes on the same face) it is known that $C \leqslant F + 1$. Moreover, in our setting there exists a case where this bound is tight (i.e., can be met with equality) as demonstrated by the example in Fig. 12.

There exists a family of graphs for which it is known that $C = F$ [2]. These graphs do not contain what are known as 'bad' holes. Applying a type of greedy algorithm to these graphs results in a solution to the GMFCD problem. Details can be found in below. These results are used to prove the correctness of the exact algorithm presented in the next subsection.

We now describe a family of graphs for which [2] shows that $C = F$. In order to describe these graphs we introduce some notation. Assume S and T are on a same face, denoted by B. Consider the two paths between S and T that form the

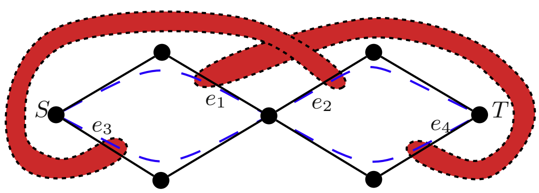

Fig. 10. A graph and holes in the context of [2]. There are two holes shown in gray above (red in online color version). One hole intersects e_2 and e_3 and the other intersects e_1 and e_4. So the two dashed paths in above constitute a geographic max-flow in this setting. Note however that in our context, there exists a hole centered at the middle node that intersects the middle four links so $F = 1$. This example makes clear the key difference between the two settings; in our setting geographic max-flow paths must be node disjoint except perhaps for nodes located inside the protected zone. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

boundary of B. Denote them by q and r respectively. We now define a *bad* hole with respect to face B.

Definition 1. A bad hole with respect to face B is a hole that intersects both q and r but does not contain a curve with endpoints on q and r that only intersects faces other than B.

For an example of some bad holes, see the holes in Fig. 11.

Lemma 1 [2]. *If there does not exist a bad hole with respect to a common ST face, then $C = F$.*

In fact, when there is no bad hole a simple greedy algorithm is optimal. The greedy algorithm starts with path q and removes all links not hole disjoint with path q. The common face will now be a subset of a larger face for which a new q and r are defined. We remove all links not hole disjoint with respect to this new path q and repeat until S and T are no longer connected.

Lemma 2 [2]. *If there does not exist a bad hole with respect to a common ST face, then greedy algorithm returns a solution to the GMFCD problem.*

It is interesting to note that the greedy algorithm is not always guaranteed to give an optimal solution when there exist bad holes (in contrast the analogous greedy algorithm always works in the non-geographic setting [4]). Fig. 11 shows an example of the greedy approach failing. The greedy algorithm outputs just one path whereas the optimal solution is given by the two paths that form a rectangle.

3.3. Exact algorithm

Next we present an algorithm to solve the GMFCD problem exactly that works by applying a greedy routine to every ST path. We give a brief overview of the algorithm. Let p be a ST path in G. We remove every link that is not hole disjoint with p (effectively, every link outside the protected zone that intersects a 'worm' around p is removed).

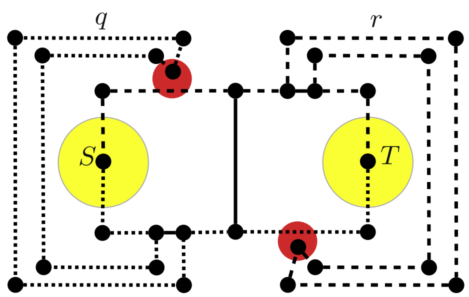

Fig. 11. The gray disks (red disks in the online color version) above are examples of *bad* holes on this graph. Note S and T share the outer infinitely large face. Consider the paths that form the boundary of this common face, shown as the dotted path labeled q and the dashed path labeled r. Each of these holes intersects both q and r but does not contain a curve with endpoints on q and r that only intersects the inside faces, thus they are bad holes. We also note that the greedy algorithm outputs a single path, q, whereas the optimal solution is given by the two center paths which form a rectangle. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

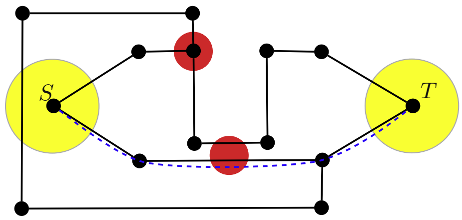

Fig. 12. A simple network where S and T lie in the same face (a version of this example may be found in [15]). All relevant holes are shown above in gray (red in online color version); others holes can effectively be ignored. Note $C = 2$ and $F = 1$ (a geographic min-cut is given by $\{u_1, u_2\}$ and max-flow given by path corresponding to the blue dotted curve). This shows the analogue to the max-flow min-cut theorem [3] does not hold in our setting. Also, it shows that there exists a case where the bound $C \leqslant F + 1$ (shown in [2]) is tight for our problem when S and T lie in the same face. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Denote the resulting graph by G' and let F' denote the cardinality of the geographic max-flow for G'. S and T now share a common face on G' (with a caveat described in Appendix B). It can be shown that the greedy algorithm on G' finds the geographic max-flow for G'. Additionally, if p belongs to a solution to the GMFCD problem, then $F = F' + 1$ and so p combined with the set of paths found by the greedy approach is an optimal solution. Thus, by considering all ST paths we can find a set of paths that includes p that is an optimal solution to the GMFCD problem. See Algorithm 1 for an explicit description.

Algorithm 1. Exact Algorithm to Solve GMFCD Problem.

1: maxDisjointPaths $\leftarrow \emptyset$
2: **for** every ST path p **do**
3: **call greedyRoutine**(p)
maxDisjointPaths
Procedure greedyRoutine(p)
4: disjointPaths $\leftarrow p$
5: $G' \leftarrow G$ except for links that intersect a hole that intersects p
6: **while** S and T in same component of G' **do**
7: **call removeQ**
8: **if** $|$disjointPaths$| > |$maxDisjointPaths$|$ **then**
9: maxDisjointPaths $\leftarrow$ disjointPaths
Procedure removeQ
10: $\{q, r\} \leftarrow ST$ paths that form the boundary of the new face
11: disjointPaths $\leftarrow$ disjointPaths $\cup\, q$
12: $G' \leftarrow G'$ except for links that intersect a hole that intersects q

---

We now present a few lemmas which help prove that Algorithm 1 solves the GMFCD problem.

Lemma 3. *G' contains no bad holes with respect to the new common face S and T share.*

Proof. Let B denote the new common face S and T share after the removal of links not hole disjoint with path p.

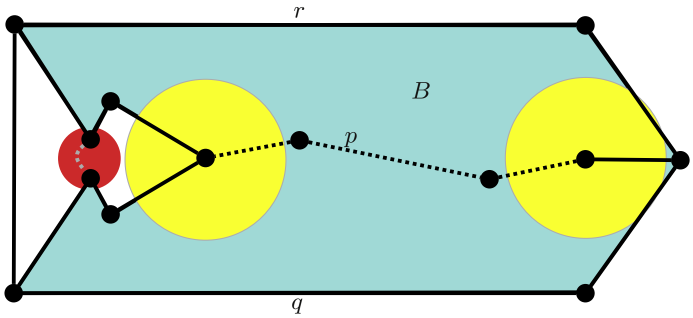

Fig. 13. The dashed links represent path p. Every link not hole disjoint with path p is removed. The gray face (teal face in the online color version) B above represents the new common face S and T share. q and r are the two ST paths that form the boundary of B. If a hole intersects both q and r, it must not intersect p (because all links intersecting a hole that intersects p are removed). This hole must then contain a curve with endpoints on q and r that only intersects faces other than B (shown as the gray dashed curve above).

Let q and r be two ST paths that form the boundary of face B (see Fig. 13).

We will use proof by contradiction. Assume there exists a bad hole with respect to B, denoted by h. Hole h must intersect both q and r but not p (because all links intersecting a hole that intersects p are removed). Thus, h must contain a curve with endpoints on q and r that only intersects faces other than B (see Fig. 13). So, h is not a bad hole, a contradiction. □

Lemma 4. *If p is a path in a solution to the GMFCD problem, then $F' + 1 = F$.*

Proof. G' necessarily contains all the other disjoint paths in the solution to the GMFCD problem because only links not hole disjoint from p were removed from G. Since p is a single path, we have $F' + 1 = F$.

Theorem 2. *Algorithm 1 finds an exact solution for the GMFCD problem.*

Proof. A path in a solution to the GMFCD problem will be considered by Algorithm 1 since every ST path is enumerated. Let p denote one of these paths. By Lemma 3 we know G' contains no bad holes with respect to the new common face S and T share. So, by Lemma 2 the greedy algorithm obtains a geographic max-flow for G'. Since p is assumed to be in the solution, by Lemma 4 we know $F' + 1 = F$. Thus, path p combined with the result of the greedy algorithm on G' is a solution to the GMFCD problem. □

This algorithm may not be practical since typically the number of ST paths grows exponentially with the size of a graph, however, it gives insight to the development of a good heuristic algorithm.

3.4. Heuristic

The basis of the heuristic algorithm presented here is to try to identify the paths that are likely to be in the geographic max-flow. The algorithm works similarly to the exact algorithm above except we apply the greedy routine to a subset of paths, instead of every ST path. In particular, the subset of paths considered consists of those found by a standard (node disjoint) max-flow algorithm on the original topology. We apply the greedy routine on every one of these paths and return the largest set of disjoint paths found. In the next section we provide some numerical results using this heuristic. See Algorithm 2 for an explicit description.

Algorithm 2. Heuristic Algorithm For GMFCD Problem.

1: maxDisjointPaths $\leftarrow \emptyset$
2: $P \leftarrow$ max-flow ST paths (non-geographic)
3: **for** every path $p \in P$ **do**
4: **call greedyRoutine**(p)
5: **return** maxDisjointPaths

3.5. Numerical results

Similar to Section 2.3, we discuss the results of our developed algorithms for the GMFCD problem when applied to a major network provider [7].

Fig. 14 shows a result of the GMFCD heuristic algorithm. The four disks represent hole locations in a geographic min-cut. The four 'worms' correspond to hole disjoint paths found using the GMFCD heuristic algorithm. Since the cardinality of the geographic max-flow and min-cut solutions is the same and $F \leqslant C$, we know the heuristic has found an optimal solution to the GMFCD problem in this case.

Fig. 15 shows the cardinality of a solution to the GMCCD problem and result of the GMFCD heuristic algorithm as a function of hole radius. As expected, the cardinality of the results decreases as hole radius increases. Additionally, since $F \leqslant C$, these results show the GMFCD heuristic algorithm gives an optimal solution to the GMFCD problem for nearly all radii considered.

It is known $C \leqslant 2F + 2$ in the more general setting of [2]. However we believe our geographical setting allows for this bound to be tightened. We conjecture that $C \leqslant F + 1$.

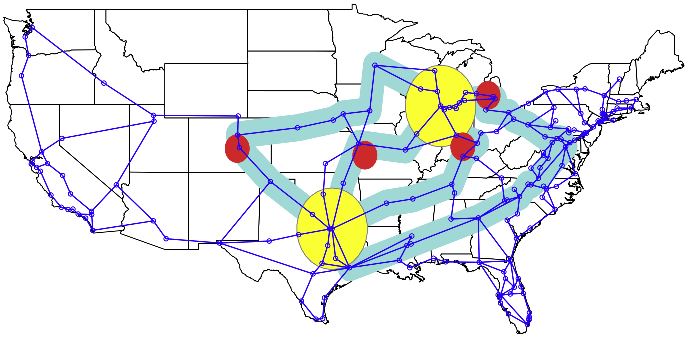

Fig. 14. Result of GMFCD heuristic algorithm when $r_b = 1.0$, $r_p = 3.0$, S = Dallas, and T = Chicago. The four gray disks (red disks in the online color version) represent the hole locations in a geographic min-cut and the light gray disks (yellow disks in the online color version) represent the protected zones. The four light gray 'worms' (teal 'worms' in the online color version) correspond to hole disjoint paths found using the heuristic algorithm. Since the cardinality of the max-flow and min-cut solutions is the same and $F \leqslant C$, we know the heuristic has found an optimal solution to the GMFCD problem. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

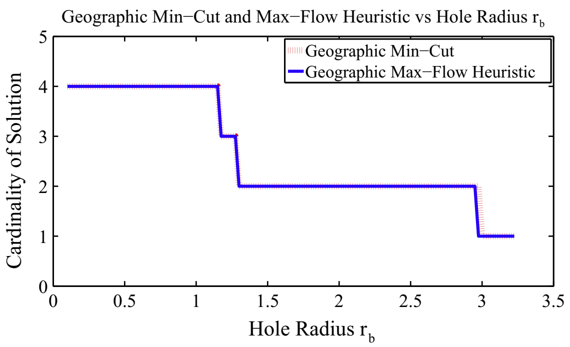

Fig. 15. Cardinality of a solution to the GMCCD problem and result of the GMFCD heuristic algorithm as a function of hole radius r_b when $r_p = 3.0$, S = Dallas, and T = Chicago. As expected, the cardinality of the results decreases as hole radius increases. Additionally, since $F \leqslant C$, these results show the GMFCD heuristic algorithm gives an optimal solution to the GMFCD problem for nearly all radii considered (the plots are nearly the same and largely overlap).

Using the algorithm in Section 2 and running CPLEX on the ILP in Section 3, we solve the GMCCD and GMFCD problems exactly for 1000 randomly generated geographic graphs consisting of 13 nodes. We found $C = F$ for 98.8% of the instances and $C = F + 1$ for the remaining 1.2%. There was not a single example where C exceeded F by more than 1, thus supporting our conjecture.

We also compared the result of the GMFCD heuristic algorithm to the GMFCD optimal solution for these 1000 randomly generated graphs. The GMFCD heuristic algorithm returned an optimal solution for 98.8% of these instances and contained one less path than the GMFCD optimal solution for the remainder.

3.6. Complexity of the GMFCD problem

The max-flow problem in [2] is shown to be NP-hard, however, the proof does not directly transfer to our setting since in our setting geographic max-flow paths cannot intersect outside protected zones. We believe a polynomial time solution may be possible and this is a subject of future work.

4. Conclusions and future work

Motivated by applications in the area of network survivability, in this paper we present a geographic max-flow and min-cut problem where failures, modeled as disks, may be placed anywhere in the graph except for certain protected zones. We show these problems can be reduced to discrete ones and present a polynomial time algorithm for the GMCCD problem based on ideas from [2,11]. We then develop an ILP formulation, an exact algorithm, and a heuristic algorithm for the GMFCD problem. Using these algorithms, we obtain numerical results for a specific backbone network, thereby demonstrating the applicability of our algorithms to a real-world network.

Our approach provides a way to look at network survivability in the face of multiple disasters or attacks that takes into account the geographical correlation between links. Some future directions include application of this approach to the electric power transmission network, finding a tight bound on the difference between geographic min-cut and max-flow (i.e. the analog to the max-flow min-cut theorem), finding the complexity of the GMFCD problem, and the development of network design tools (e.g. how to build a network under some constraints such that geographic min-cut is maximized).

Acknowledgements

This work was supported by National Science Foundation – United States Grants CNS-0830961, CNS-1017714, CNS-1017800, NSF CAREER Grant 0348000, and by DTRA Grants HDTRA1-07-1-0004 and HDTRA-09-1-005. Preliminary and partial versions of this paper appeared in Proc. IEEE INFOCOM'12, March 2012 [12].

Appendix A. Details of Step 1

There are an infinite number of hole locations centered outside the protective disks; in this step we find a polynomial (quadratic) sized set of candidate holes for the GMCCD problem. We first make a note about holes. Let h and h' be holes such that h' intersects every link h does in addition to possibly other links. We note that if h belongs to a set of holes that disconnects S and T, then replacing h with h' will still result in S and T being disconnected (if S and T are disconnected, removing additional links will also leave them disconnected).

We now describe how to find a set of potential failure locations for the GMCCD problem. First, we apply an algorithm from [11] where the graph and disaster model is the same except that there exists no protected zones. The algorithm creates a polynomial (quadratic) sized set of holes, denoted by A, such that for every hole in the plane there exists a hole $a \in A$ that intersects at least the same set of links. Note however, one of these holes may be centered inside one of the protective disks around S or T (see Fig. 3), and so we must consider additional holes to solve the GMCCD problem. Let A' be the set of disks in A not centered in the protected zone.

Let h be a hole not centered in the protected zone and let a be a hole in A that intersects at least the same links as hole h. We will show there exists a polynomially (quadratic) sized set of holes centered outside the protected zone, denoted by M, such that if hole a is in the protected zone, then a hole $m \in M$ intersects at least the same links as hole h. So, for every hole not centered in the protected zone there exists a hole in $A' \cup M$ that intersects at least the same set of links. Thus, $A' \cup M$ gives us a polynomial (quadratic) sized set of candidate holes for the GMCCD problem.

In the following we present notation that allows us to describe the locations of the holes in M. Let O^S denote the circle centered at S with radius r_p and let O^T denote the circle centered at T with radius r_p. Let e denote a link and let $\overleftrightarrow{e}$ denote the line that contains e. Let $\partial H(e, r_b)$ denote the boundary of $H(e, r_b)$.

Lemma 5. *Let M be the set of all holes of radius r_b centered on at least one of the following sets:*

(i) $\cup_{e \in E}\{O^S \cap \partial H(e, r_b) | e$ *does not have an endpoint at* S *or* $r_p \neq r_b\}$,
(ii) $\cup_{e \in E}\{O^T \cap \partial H(e, r_b) | e$ *does not have an endpoint at* T *or* $r_p \neq r_b\}$,
(iii) $\cup_{e \in E}\{O^S \cap \overleftrightarrow{e} | e$ *has an endpoint at* S *and* $r_p = r_b\}$,
(iv) $\cup_{e \in E}\{O^T \cap \overleftrightarrow{e} | e$ *has an endpoint at* T *and* $r_p = r_b\}$.

Assume there exists a polynomial (quadratic) sized set of holes A such that for every hole in the plane there exists a hole $a \in A$ that intersects at least the same links. Given a hole h not centered in the protected zone, if a hole $a \in A$ is in the protected zone and intersects at least the same set of links as h, then there exists a hole $m \in M$ that intersects at least the same links as h.

Proof. Let β denote the center of hole h and α denote the center of a. Let Z denote the set of all links that intersect h. Note that $x \in \cap_{z \in Z} H(z, r_b)$ iff x is the center of a hole that intersects every $z \in Z$. So $\beta \in \cap_{z \in Z} H(z, r_b)$ and $\alpha \in \cap_{z \in Z} H(z, r_b)$. Also note that $\cap_{z \in Z} H(z, r_b)$ is convex and thus path-connected, so there exists a path p in $\cap_{z \in Z} H(z, r_b)$ from α to β that necessarily intersects O^S or O^T. Let y be this intersection point. The hole centered at y must necessarily intersect at least the same links as h since $y \in \cap_{z \in Z} H(z, r_b)$.

W.l.o.g. assume y lies on O^S. Note every point in the sets (i) and (iii) above lies on O^S. Let y' be the first point in (i) or (iii) in the clockwise direction from y on O^S. Now, $y' \in \cap_{z \in Z} H(z, r_b)$ since every hippodrome y intersects is also intersected by y'. Therefore a hole centered on (i) or (iii) must intersect every link that intersects h. $\square$

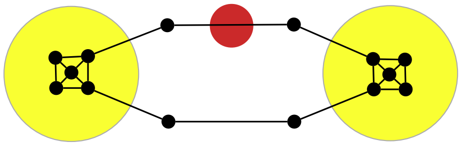

Fig. 16. No hole intersects the nodes located in the protected zone, so removing links not hole disjoint with some path will never result in a graph where S and T are on the same face.

Since for each link we consider at most eight holes ($|O^S \cap \partial H| \leqslant 4$ and $|O^T \cap \partial H| \leqslant 4$ under conditions above), M is of polynomial (quadratic) size. Since $|A'|$ is polynomial (quadratic), the set of potential holes for the GMCCD problem, $A' \cup M$, is of polynomial (quadratic) size. For the remainder of the section let $U = A' \cup M$.

Appendix B. Modifying G

Here we discuss a caveat for applying Algorithms 1 and 2s. We first note that removing every link not hole disjoint with some ST path does not ensure S and T lie on the same face. See Fig. 16 for an example. Here we note how to modify G so that when links around a ST path are removed S and T are guaranteed to share a common face. Let O be a circle of radius $r_p - r_b$ centered at S (assume $r_p - r_b \geqslant 0$). Note no hole may intersect a link anywhere within this circle. Place nodes everywhere O intersects a link. Consider the links forming paths that lie entirely within O that have endpoints at S and these new nodes. Replace these links such that there exists a path inside O from S to each of these new nodes such that these paths do not intersect each other except at S. Repeat this process for T. Now when links within r_b of a particular path are removed (outside the protected zone) S and T are guaranteed to be on the same face. Since the removed links do not intersect any hole and since the connectivity of G is unchanged outside the protected zone, the solution to the GMFCD problem remains the same. We assume the exact algorithm is applied after this modification.

References

[1] P. Agarwal, A. Efrat, S. Ganjugunte, D. Hay, S. Sankararaman, G. Zussman, The resilience of WDM networks to probabilistic geographical failures, in: Proc. IEEE INFOCOM'11, April 2011.
[2] D. Bienstock, Some generalized max-flow min-cut problems in the plane, Math. Oper. Res. 16 (2) (1991) 310–333.
[3] J. Bondy, U.R. Murty, Graph Theory, 2008.
[4] G. Borradaile, P. Klein, An $O(n \log n)$ algorithm for maximum st-flow in a directed planar graph, J. ACM (JACM) 56 (2) (2009) 9.
[5] A. Efrat, M. Sharir, A near-linear algorithm for the planar segment-center problem, Discrete Comput. Geom. 16 (3) (1996) 239–257.

[6] J.S. Foster Jr., et al., Report of the Commission to Assess the Threat to the United States from Electromagnetic Pulse (EMP) Attack, Executive Report, vol. I, April 2004.
[7] Level 3 Communications, Network Map, September 2008. <http://www.level3.com/Resource-Library/Maps/Level-3-Network-Map.aspx>.
[8] J. Mitchell, V. Polishchuk, Thick non-crossing paths and minimum-cost flows in polygonal domains, in: Proceedings 23rd ACM Symposium on Computational Geometry, Citeseer, 2007, pp. 56–65.
[9] S. Neumayer, E. Modiano, Network reliability with geographically correlated failures, in: Proc. IEEE INFOCOM'10, March 2010.
[10] S. Neumayer, E. Modiano, Network reliability under random circular cuts, in: Proc. IEEE GLOBECOM'11, December 2011.
[11] 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.
[12] S. Neumayer, A. Efrat, E. Modiano, Geographic max-flow and min-cut under a circular disk failure model, in: INFOCOM, 2012 Proceedings IEEE, IEEE, 2012, pp. 2736–2740.
[13] W. Radasky, High-altitude Electromagnetic Pulse (HEMP): A Threat to Our Way of Life, IEEE-USA Today's Engineer, September 2007. <http://www.todaysengineer.org/2007/Sep/HEMP.asp>.
[14] J. Reif, Minimum st cut of a planar undirected network in $O(n\log^2(n))$ time, Automata Lang. Program. (1981) 56–67.
[15] A. Sen, S. Murthy, S. Banerjee, Region-based connectivity: a new paradigm for design of fault-tolerant networks, in: Proceedings of the 15th International Conference on High Performance Switching and Routing, IEEE Press, 2009, pp. 94–100.
[16] W. Wu, B. Moran, J. Manton, M. Zukerman, Topology design of undersea cables considering survivability under major disasters, in: Proc. WAINA'09, May 2009.

Sebastian Neumayer received the B.S. degree (with Highest Honors) in electrical engineering from the University of Illinois at Urbana–Champaign 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.

Alon Efrat is an associate professor in the Department of Computer Science at the University of Arizona since 2000. He has earned his PhD from Tel-Aviv University under the supervision of Prof. Micha Sharir at 1998. He was also a postdoctorate research assistant at Stanford University, and at IBM Almaden Research Center. He won the NSF CAREER award in 2004. He serves on the editorial board of the International Journal of Computational Geometry and its Application (IJCGA). He has published over 110 papers. His focuses in optimizations problems from geometric algorithms. In particular, he is intrigued by reliability problems in networks, in algorithms for sensors networks and sensors locations, and by browsing, indexing, and manipulations of videos of educational content.

Eytan Modiano 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 PhD 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.