MIT Open Access Articles

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

Citation: Neumayer, Sebastian, Alon Efrat, and Eytan Modiano. "Geographic max-flow and min-cut under a circular disk failure model." In 2012 Proceedings IEEE INFOCOM, 2736-2740. Institute of Electrical and Electronics Engineers, 2012.

As Published: http://dx.doi.org/10.1109/INFCOM.2012.6195690

Publisher: Institute of Electrical and Electronics Engineers (IEEE)

Persistent URL: http://hdl.handle.net/1721.1/81453

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Massachusetts Institute of Technology

Geographic Max-Flow and Min-Cut Under a Circular Disk Failure Model

Sebastian Neumayer
LIDS, MIT
Cambridge, MA 02139
bastian@mit.edu

Alon Efrat
University of Arizona
Tucson, AZ 85721
alon@cs.arizona.edu

Eytan Modiano
LIDS, MIT
Cambridge, MA 02139
modiano@mit.edu

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.

I. 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 [5], [12]. 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], [10]. The impact of a single randomly located disaster on network connectivity is considered in [8], [9], [15]. 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 [14] considered the problem of a geographic max-flow and min-cut in a wireless network setting. In [7] the problem of finding the maximum number of geographically
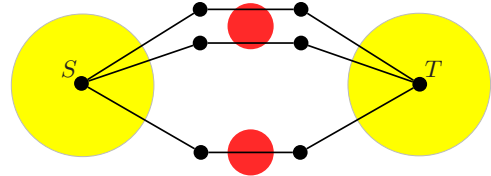
Fig. 1. The light gray area (yellow area in online color version) above represents the protected zone that no circular failure may be centered. The gray disks (red disks in online color version) represent disasters that remove links (of unit capacity) they intersect. Two disasters are required to disconnect 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.

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 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 [6], 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 $\leq$ min-cut $\leq$ max-flow $+1$.

II. Geographic Min-Cut

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

A. 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.*

B. 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 sized set by extending the methods in [10]. 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 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 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 [4]. Note that a hole of radius r_b is centered in $H(e, r_b)$ if and only if the hole intersects e.

In [10] we considered the same failure model without the protected zones. Under this model we found a polynomial size 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.

The polynomially sized set of potential failure locations found in [10] cannot be used for the GMCCD problem because of the restriction that holes cannot be placed inside the protected zones. For example, in a particular network some of the holes found using the methods in [10] may be centered inside the protected zone and thus cannot be considered. If we consider additional holes that are centered at the intersection of the boundaries of the protected zones and hippodromes associated with links, we can show that this expanded set of potential failure locations is sufficient. For brevity we omit the details, which can be found in [11]. Let this polynomially sized set of potential hole locations for the GMCCD problem be given by U.

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. 2 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. 2 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. 2 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. 3). This is similar to a known algorithm to find the min-cut in a planar graph (in the standard
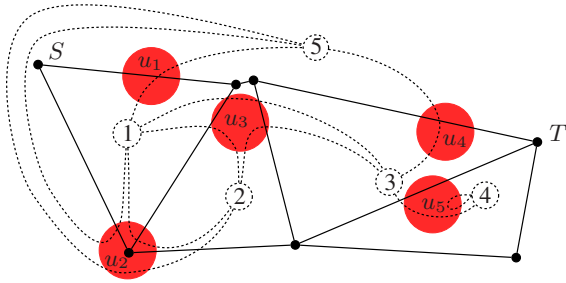
Fig. 2. The dual-like graph is shown by the dotted potion of the figure above. The solid dots and line segments represent the original network G. For ease of presentation, we take the set of gray disks (red disks in the online color version) above to be U. G has five faces; each of these faces represents 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. 4).
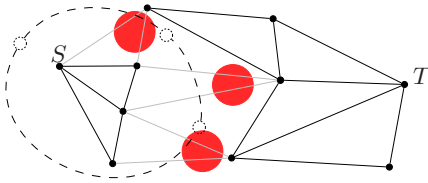

Fig. 3. The dashed links above represents a closed walk in K such that the corresponding holes (shown as disks) remove links which disconnect S and T. By searching over a set of closed walks in K, we will be able to find a solution to the GMCCD problem.

sense); the algorithm finds the shortest closed walk in the dual graph that disconnects S from T [13].

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. 4. 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
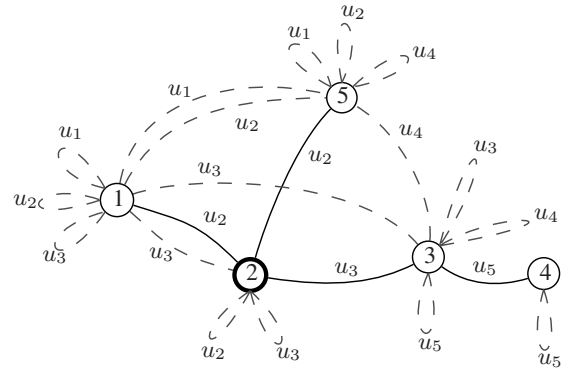

Fig. 4. The graph shown above is K from Fig. 2 complete with self-loops. Every link is marked with its respective hole.

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 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 polynomial in $|N|$. Since the algorithm considers a closed walk for every node-link pair in K, we know the algorithm runs in polynomial time in $|N|$.

C. Numerical Results

We used the algorithm in the previous section to solve the GMCCD problem for a major network provider [6]. 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. 5 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. 6 shows that when r_b is reduced slightly to 1 a total of four disasters are required to disconnect the two cities.

III. 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
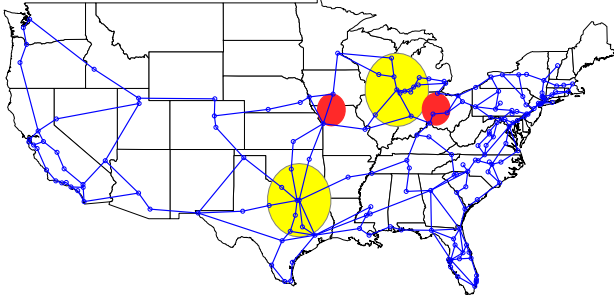
Fig. 5. A solution to the GMCCD problem when $r_b = 1.3$, $r_p = 3.0$, S = Dallas, and T = Chicago. The gray disks (red disks in the online color version) represent the hole locations and the light gray disks (yellow disks in the online color version) represent the protected zones. Only two disasters, located at 'choke' points to the east and west of Chicago, are required to disconnect these cities.
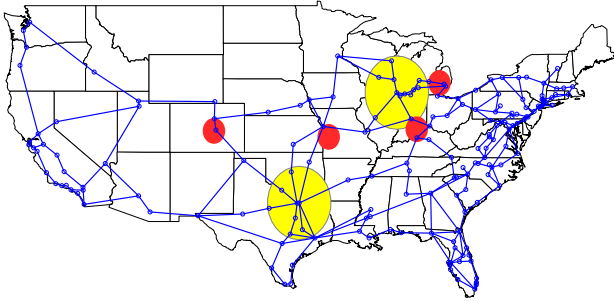

Fig. 6. A solution to the GMCCD problem when $r_b = 1$, $r_p = 3.0$, S = Dallas, and T = Chicago. The gray disks (red disks in the online color version) represent the hole locations and the light gray disks (yellow disks in the online color version) represent the protected zones. Note four disasters with $r_b = 1$ are required to disconnect the two cities, whereas only two disasters are required with $r_b = 1.3$ (see Fig. 5).

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.

A. 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.

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

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. Due to space constraints the formulation is not presented here but may be found in the technical report [11].

B. Bounds on C and F

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. 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. 7. 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 $F \leq C$ 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.

Since only a polynomial number of hole locations need to be considered (as discussed in step 1 of Section II), it follows that the GMFCD and GMCCD problems are special cases of the geographic max-flow and min-cut problems described in [2][1]. 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 $C \leq F + 1$. 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. 7.

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 the technical report [11]. These results are used to prove the correctness of the exact algorithm presented in the next subsection.

C. 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). 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 [11]). 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

[1]In particular, every disaster in [2] must have a shape that is homeomorphic to a disk.
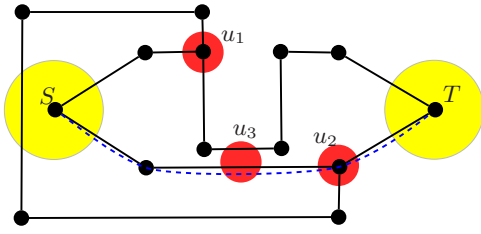
Fig. 7. A simple network where S and T lie in the same face (a version of this example may be found in [14]). All relevant holes are shown above in gray (red in online color version); others holes can effectively be ignored. Note $C = 2$ and $F = 1$ (a geographic min-cut is given by $\{u_1, u_2\}$ and 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 $C \leq F + 1$ (shown in [2]) is tight for our problem when S and T lie in the same face.

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. Due to space constraints, the details and proofs of the algorithm may be found in the technical report [11].

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.

D. 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 [11] for an explicit description.

E. Numerical Results

Similar to Section II-C, we discuss the results of our developed algorithms for the GMFCD problem when applied to a major network provider [6].

Fig. 8 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 $F \leq C$, we know the heuristic has found an optimal solution to the GMFCD problem in this case.

It is known $C \leq 2F + 2$ in the more general setting of [2]. However we believe our geographical setting allows for this bound to be tightened. We conjecture that $C \leq F + 1$. Using the algorithm in Section II and running CPLEX on the the ILP in Section III, we solve the GMCCD and GMFCD problems exactly for 1000 randomly generated geographic graphs consisting of 13 nodes. We found $C = F$ for 99%
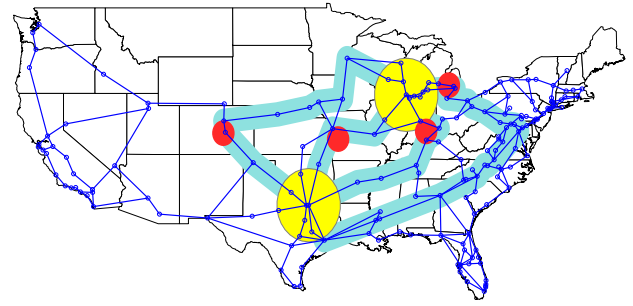

Fig. 8. Result of GMFCD heuristic algorithm when $r_b = 1.0$, $r_p = 3.0$, $S =$ Dallas, and $T =$ Chicago. The four gray disks (red disks in the online color version) represent the hole locations in a geographic min-cut and the light gray disks (yellow disks in the online color version) represent the 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 $F \leq C$, we know the heuristic has found an optimal solution to the GMFCD problem.

of the instances and $C = F + 1$ for the remaining 1%. There was not a single example where C exceeded F by more than 1, thus supporting our conjecture.

REFERENCES

[1] P. Agarwal, A. Efrat, S. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, "The resilience of wdm networks to probabilistic geographical failures," in *Proc. IEEE INFOCOM'11*, Apr. 2011.

[2] D. Bienstock, "Some generalized max-flow min-cut problems in the plane," *Math. Oper. Res.*, vol. 16, no. 2, pp. 310–333, 1991.

[3] J. Bondy and U. Murty, "R. Graph theory," 2008.

[4] A. Efrat and M. Sharir, "A near-linear algorithm for the planar segment-center problem," *Discrete and Computational Geometry*, vol. 16, no. 3, pp. 239–257, 1996.

[5] J. S. Foster Jr. et al., "Report of the commission to assess the threat to the United States from electromagnetic pulse (EMP) attack, Vol. I: Executive report," Apr. 2004.

[6] Level 3 Communications, Network Map, Sept. 2008. [Online]. Available: http://www.level3.com/Resource-Library/Maps/Level-3-Network-Map.aspx

[7] J. Mitchell and 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.

[8] S. Neumayer and E. Modiano, "Network reliability with geographically correlated failures," in *Proc. IEEE INFOCOM'10*, Mar. 2009.

[9] ——, "Network reliability under random circular cuts," in *Proc. IEEE GLOBECOM'11*, Dec. 2011.

[10] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the vulnerability of the fiber infrastructure to disasters," *IEEE/ACM Transactions on Networking*, no. 99, pp. 1–1, To Appear.

[11] S. Neumayer, A. Efrat, and E. Modiano, "Geographic max-flow and min-cut under a circular disk failure model," Massachusetts Institute of Technology, EE, Tech. Rep. 2011-07-29, Jan. 2012. [Online]. Available: http://www.mit.edu/%7Ebastian/INFOCOM12-Tech-Rpt-2012-01.pdf

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

[13] J. Reif, "Minimum st cut of a planar undirected network in $O(n \log^2(n))$ time," *Automata, Languages and Programming*, pp. 56–67, 1981.

[14] A. Sen, S. Murthy, and 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.

[15] W. Wu, B. Moran, J. Manton, and M. Zukerman, "Topology design of undersea cables considering survivability under major disasters," in *Proc. WAINA'09*, May 2009.