Network Vulnerability to Single, Multiple, and Probabilistic Physical Attacks

Pankaj K. Agarwal*, Alon Efrat†, Shashidhara K. Ganjugunte*,
David Hay‡, Swaminathan Sankararaman†, and Gil Zussman‡

*Computer Science, Duke University. {pankaj, shashigk}@cs.duke.edu
†Computer Science, University of Arizona. {alon, swami}@cs.arizona.edu
‡Electrical Engineering, Columbia University. {hdavid, gil}@ee.columbia.edu

***Abstract*—Telecommunications networks heavily rely on the physical infrastructure and, are therefore, vulnerable to natural disasters, such as earthquakes or floods, as well as to physical attacks, such as an Electromagnetic Pulse (EMP) attack. Large-scale disasters are likely to destroy network equipment and to severely affect interdependent systems such as the power-grid. In turn, long-term outage of the power-grid might cause additional failures to the telecommunication network.**

In this paper, we model an attack as a disk around its epicenter, and provide efficient algorithms to find vulnerable points within the network, under various metrics. In addition, we consider the case in which *multiple* disasters happen simultaneously and provide an approximation algorithm to find the points which cause the most significant destruction. Finally, since a network element does not always fail, even when it is close to the attack's epicenter, we consider a simple *probabilistic* model in which the probability of a network element failure is given. Under this model, we tackle the cases of single and multiple attacks and develop algorithms that identify potential points where an attack is *likely* to cause a significant damage.

***Index Terms*—Network survivability, geographic networks, network design, Electromagnetic Pulse (EMP), computational geometry.**

I. Introduction

Telecommunication networks are crucial for the normal operations of all sectors of our society. During a crisis, telecommunications is essential to facilitate the control of physically remote agents, provides connections between emergency response personnel, and eventually enables reconstitution of societal functions. However, telecommunication networks heavily rely on physical infrastructure (such as optical fibers, amplifiers, routers, and switches), making them *vulnerable to physical attacks, such as an Electromagnetic Pulse (EMP) attack, as well as natural disasters, such as earthquakes, hurricanes or floods* [1]–[4]. Increasingly, networks use a shared infrastructure to carry voice, data, and video simultaneously; hence, failures in the physical infrastructure will lead to a breakdown of vital services.

Physical attacks or disasters affect a specific geographical area and will result in failures of neighboring components. Therefore, it is crucial to consider the effect of disasters on the *physical (fiber) layer* as well as on the (logical) network layer. Although there has been a significant amount of work on network survivability, most previous work considered a small number of isolated failures. In contrast, and similarly to [6]–[8], in this paper we consider events that cause a large number of failures in a specific geographical region. Furthermore, while in [6]–[8] only a single disaster was considered, we consider a case in which *several disasters happen simultaneously* and show how this affects the vulnerability of the network.

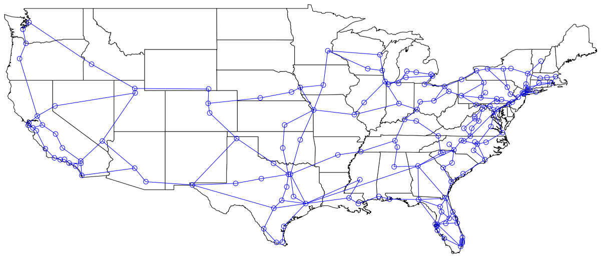
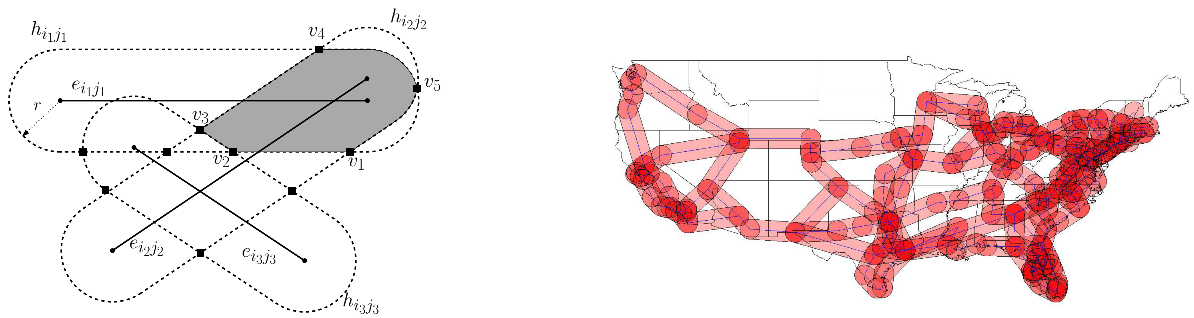


Fig. 1. The fiber backbone operated by a major U.S. network provider [5].

The resilience of the network clearly depends on its topology as well as the shape of the disaster. This work focuses on *circular cut* failures, where all components within a pre-determined disk around the epicenter of the disaster may fail (while all components outside this disk continue to operate normally). In addition, we consider two network topologies: a general graph, in which no assumption is made, and a *planar graph*, in which we assume that the links intersect each other only at the end-points. This assumption is satisfied in practice.

Moreover, we present a new *probabilistic failure model*, in which network components in the vicinity of the disaster fail with probability p while other components further away do not fail. We then provide efficient algorithms to find the vulnerable points under this model. Most algorithms are precise in case of a single disaster/attack, although, in order to reduce computation time we also provide *approximation algorithms* for some instances. For multiple simultaneous disasters, the problem is known to be NP-hard, both in the probabilistic and the deterministic setting [9]. Thus, we provide only efficient approximation algorithms. Interestingly, our approximation algorithms for the probabilistic failure model has the same—and in some instances even superior—performance than its deterministic counterpart. Finally, we present numerical results that demonstrate the use of our algorithms.



(a) The hippodromes $h_{i_1j_1}$, $h_{i_2j_2}$, and $h_{i_3j_3}$ which correspond to 3 links e_{i_1,j_1}, e_{i_2,j_2}, and $e_{i_3j_3}$.

(b) The hippodromes, corresponding to attacks with radius $r = 60$ miles, of the fiber backbone network shown in Fig. 1.

Fig. 2. Examples of hippodromes induced by various networks. Each hippodrome indicates the region around a link where an attack needs to occur, i.e., a cut needs to be centered, in order to affect that link. The arrangement of the hippodromes for 3 links is shown in Fig 2a. A face of this arrangement is shown shaded and the vertices marked as squares. Each arc is the portion of the boundary of a hippodrome from one vertex to another.

II. Related Work

Network survivability and resilience is a well-established research area (e.g., [10], [11] and references therein). However, most of the previous work in this area and in particular in the area of physical topology and fiber networks (e.g., [12], [13]) focused on a small number of fiber failures (e.g., simultaneous failures of links sharing a common physical resource, such as a cable, conduit, etc.). Such correlated link failures are often addressed systematically by the concept of *Shared Risk Link Group* (SRLG) [14]. In contrast with these works, we focus on failures within a specific geographical *region* (e.g., [2], [3], [15]), implying that the failed components do not necessarily share the same physical *resource*. To the best of our knowledge, geographically correlated failures have been rarely considered [6], [7], [16]–[18] and in most cases, the assumption is that the failures of the components are deterministic.

Another closely related theoretical problem is the *network inhibition problem* [19]–[22], in which the objective is to minimize the value of a maximum flow in the graph, where there is a cost associated with destroying each link, and a fixed budget is given for an orchestrated attack (namely, removing a set of links whose total destruction cost is less than the budget). However, previous works dealing with this setting and its variants (e.g., [22], [23]) did not study the removal of (geographically) neighboring links.

Notice that when the logical (i.e., IP) topology is considered, wide-spread failures have been extensively studied [23]–[26]. Most of these works consider the topology of the Internet as a random graph [27] and use percolation theory to study the effects of random link and node failures on these graphs. These studies are motivated by failures of routers due to attacks by viruses and worms rather than physical attacks.

Finally, we note that results regarding the resilience of fiber networks to geographically correlated failures were recently obtained in [6], [7]. Most related to our paper are results in [7] where, as this paper, disasters are modeled as circular areas in which the links and nodes are affected. However, [7] considers only a *single disaster* scenario where failures are *deterministic*. Moreover, the algorithms presented in [7] (for identifying vulnerable points) have a substantially higher complexity than the algorithm presented in this paper.

III. Problems Statement

The network topology is given by a geometric graph model $G = (V, E)$ where V is the set of n disjoint nodes in the plane (representing routers location) and E is the set of m links between the nodes (representing fiber links between routers). The location of node i is given by the coordinates (x_i, y_i). A link from i to j is modeled as a straight line segment from i to j and is denoted by e_{ij}. Further, we assume that every pair of links intersect in at most one point. Following [7], a disaster or attack results in a circular cut which is modeled as a disk of radius r which is centered at point b and is denoted by $cut_r(b)$. Let the minimum distance from a point $b \in \mathbb{R}^2$ to a link e be denoted by $d(b, e)$. A link $e_{ij} \in E$ is *affected* by $cut_r(b)$ centered at some point b, if and only if it is within its impact radius r. In other words, b is within the *hippodrome* h_{ij} defined as $h_{ij} = \{x \in \mathbb{R}^2 \mid d(x, e_{ij}) \le r\}$. Examples of such hippodromes are depicted in Fig. 2a. Let $\mathcal{H} = \{h_{ij} \mid e_{ij} \in E\}$ be the set of all m hippodromes.

In this paper, we consider the following two ways of measuring the impact of (one or many) attacks. First, we consider the *number of link failures* caused by the attacks. Second, we consider the *terminal reliability* that measures the effect on the global connectivity of the network caused by a cut. Namely, for a pair of nodes $v_i, v_j \in V$ and a cut $cut_r(b)$ centered at the point $b \in \mathbb{R}^2$, let $z_{ij}(b) = 1$, if there is an undirected path from v_i to v_j, even in the presence of the cut $cut_r(b)$, and let $z_{ij}(b) = 0$ otherwise. Then, the *average two-terminal reliability (ATTR)* is

$$\chi(b) = \frac{1}{\binom{n}{2}} \sum_{i \neq j} z_{ij}(b).$$

For example, if the graph G is connected even after an attack at b, then $\chi(b) = 1$.

The following observation captures the connection between the average two-terminal reliability and the partitioning of the network graph.

Observation 1. *Assume G is partitioned to (maximally) connected components of sizes $n_1 \le n_2 \le \ldots \le n_k$, and let j*

be the first index for which $n_j \geq 2$. *Thus, the average two-terminal reliability of the graph is* $\frac{1}{\binom{n}{2}} \sum_{i=j}^{k} \binom{n_i}{2}$.

In addition, when considering multiple simultaneous attacks, it turns out that identifying the most vulnerable point becomes NP-hard. Since it is intractable to find the optimal solutions in these settings, we aim at devising *constant-factor approximation algorithms*. Namely, algorithms which provide a solution within a (known) constant multiplicative factor of the optimal solution.

In order to provide a framework in which the effect on a link due to an attack is probabilistic, we also consider a probabilistic model for the cuts. In this model, we assume that links within $cut_r(b)$ fail with some probability p. Then, we measure the *expected* number of failed links, denoted by $\Phi(b)$. More formally, $f_{ij}(b)$ is a random variable representing the probability that the link between v_i and v_j is destroyed by a cut whose epicenter is in b:

$$f_{ij}(b) = \begin{cases} p & \text{if } b \in h_{ij}, \\ 0 & \text{otherwise.} \end{cases}$$

and $\Phi(b) = \sum_{e_{ij} \in E} f_{ij}(b)$. Notice that the objective of an adversary is to find a point b that either minimizes $\chi(b)$ or maximizes $\Phi(b)$.

IV. Attacks with Deterministic Outcome

In this section we assume that the response of a link to an attack at b is completely predictable. Namely, a link e_{ij} is destroyed by an attack at b, if and only if it is *affected* by $cut_r(b)$, i.e., $b \in h_{ij}$.

A. The average two-terminal reliability — single cut case

We present a fast algorithm to compute a point b^* that minimizes, among all points in the plane, the average two-terminal reliability after the links in $cut_r(b^*)$ are removed from the graph. We next present a fast algorithm to compute b^*.

For a set $\mathcal{H}$ of hippodromes, the *arrangement* $\mathcal{A}(\mathcal{H})$ is the subdivision of the plane $\mathbb{R}^2$, into vertices, arcs, and faces. The vertices are the intersection points of the hippodromes, the arcs are either maximally connected circular arcs or straight line segments of the boundaries of hippodromes that occur between the vertices, and faces are maximally connected regions bounded by arcs (see Fig. 2). Let $\mathcal{A} = \mathcal{A}(\mathcal{H})$ be the arrangement of hippodromes of all links, and let $|\mathcal{A}|$ denote the number of vertices, arcs, and faces in $\mathcal{A}$. In the worst case $|\mathcal{A}|$ can be quadratic in m but in practice it is near-linear in m. The arrangement can be computed in expected time $O(m \log m + |\mathcal{A}|)$ [28]. Our algorithm to compute the point b^* takes $O(|\mathcal{A}| \log^2 m)$ expected time, thus improves the time complexity of the algorithm presented in [7], by a factor of $\approx m^3/|\mathcal{A}|$.

For a face $\varphi \in \mathcal{A}$, let $E_\varphi \subseteq E$ be the set of links e_{ij} such that h_{ij} does not contain φ. Note that for two adjacent faces φ, φ' in $\mathcal{A}$, $|E_\varphi \oplus E_{\varphi'}| \leq 1$, where the operator $\oplus$ represents the symmetric difference between the set E_φ, and $E_{\varphi'}$. Moreover, for all points b in a face φ of $\mathcal{A}$, $\chi(b)$ remains the same, which we denote by $\chi(\varphi)$. We traverse the arrangement $\mathcal{A}$, compute $\chi(\varphi)$ for each face, and return a point from a face that maximizes the value of χ. To facilitate the traversal of $\mathcal{A}$ and the computation of $\chi(\varphi)$, we maintain a union-find data structure $\mathcal{U}$ on the nodes of G. Whenever we visit a face φ, $\mathcal{U}$ maintains the connected components of the graph (V, E_φ) along with the size of the connected component from which $\chi(\varphi)$ can be updated quickly. $\mathcal{U}$ supports the following three operations each in $O(\log m)$ time.

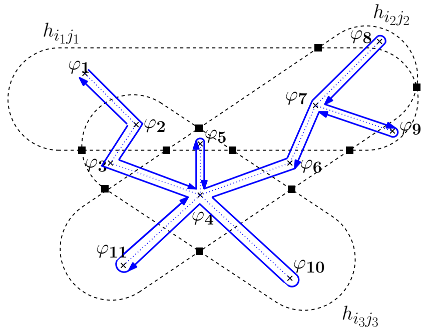


Fig. 3. Illustration of a spanning tree (dotted lines) on the dual graph of the arrangement shown in Fig. 2a, whose faces are $\mathcal{F} = \{\varphi_1, \ldots, \varphi_{11}\}$. The spanning path (solid lines) is $\Pi = \langle \varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_4, \varphi_6, \varphi_7, \varphi_8, \varphi_7, \varphi_9, \varphi_7, \varphi_6, \varphi_4, \varphi_{10}, \varphi_4, \varphi_{11}, \varphi_4, \varphi_3, \varphi_2, \varphi_1 \rangle$. Let u_i be the i-th face in Π ($1 \leq i \leq 21$). The intervals associated with the hippodromes are therefore $\mathcal{I}_{i_1j_1} = \{[u_3, u_4], [u_6, u_7], [u_9, u_9], [u_{13}, u_{19}]\}$, $\mathcal{I}_{i_2j_2} = \{[u_1, u_3], [u_{11}, u_{11}], [u_{15}, u_{15}], [u_{19}, u_{21}]\}$, and $\mathcal{I}_{i_3j_3} = \{[u_1, u_1], [u_7, u_{13}], [u_{17}, u_{17}], [u_{21}, u_{21}]\}$.

1) Merge (u, v): Merge the components of nodes u, v.
2) Find (v): Return the connected component of a node v.
3) Make_Component (v): Create a component for a node v.

We now describe the algorithm in detail.

1) We first compute the arrangement $\mathcal{A}$. Let $\mathcal{F}$ be the set of faces of $\mathcal{A}$ and let $\mathcal{A}^* = (\mathcal{F}, \Gamma)$ be the *dual graph* of $\mathcal{A}$; $\gamma_{ij} = (\varphi_i, \varphi_j) \in \Gamma$ if the faces φ_i and φ_j share an arc in $\mathcal{A}$. If the arc is a portion of the boundary of a hippodrome h_{kl}, then we label γ_{ij} with the link e_{kl}. Initialize $\mathcal{U}$ by invoking Make_Component (v) for each node $v \in V$.
2) Compute a spanning tree of the vertices of $\mathcal{A}^*$, and convert the spanning tree into a *spanning path* Π by starting from a leaf of the tree and visiting each edge of tree twice; see Fig. 3 and [29] for details. The length of the spanning path is $|\Pi| = O(|\mathcal{A}|)$.
3) Fix a hippodrome h_{ij}. If we remove the edges of Π labeled with the link e_{ij}, Π is decomposed into maximal subpaths, called intervals, such that either all faces in a subpath lie in h_{ij} or none of them lie inside h_{ij}. Let $\mathcal{I}_{ij}$ be the set of intervals of the latter type — none of the faces in any interval of $\mathcal{I}_{ij}$ destroy the link e_{ij}. Set $\mathcal{I} = \bigcup_{h_{ij} \in \mathcal{H}} \mathcal{I}_{ij}$; see Fig. 3. By construction, $|\mathcal{I}| \leq |\Pi| = O(|\mathcal{A}|)$.
4) Construct a minimum-height binary tree $\mathcal{T}$ on Π, i.e., the i-th leaf of $\mathcal{T}$ corresponds to the i-th vertex in Π. Each vertex $\xi \in \mathcal{T}$ is associated with a subpath Π_ξ of Π, spanned by the leaves in the subtree rooted at ξ. We store

an interval $I \in \mathcal{I}$ at $\xi \in \mathcal{T}$ if $\Pi_{\xi} \subseteq I$, and $\Pi_{p(\xi)} \not\subseteq I$, where $p(\xi)$ is the parent of ξ. Let $\mathcal{I}_{\xi} \subseteq \mathcal{I}$ be the set of intervals stored at ξ. We also associate a subset $E_{\xi} \subseteq E$ of links with ξ — if an interval I_{ij} belongs to $\mathcal{I}_{\xi}$ we add the link e_{ij} to E_{ξ}. Let $E^{*}_{\xi} = \bigcup_{\lambda \in \Lambda(\xi)} E_{\lambda}$, where $\Lambda(\xi)$ is the set of ancestors of ξ, including ξ itself. Let $\hat{\chi}(\xi)$ be the average number of pairs of nodes in the subgraph (V, E^{*}_{ξ}) that are connected. For a leaf $\eta \in \mathcal{T}$, if η corresponds to a face φ of $\mathcal{A}$, then $E_{\varphi} = E^{*}_{\xi}$, so $\chi(\varphi) = \hat{\chi}(\eta)$. Note that, $\sum_{\xi \in \mathcal{T}} |\mathcal{I}_{\xi}| = O(|\mathcal{I}| \log m)$.

5) Traverse $\mathcal{T}$ in a *top-down manner*, i.e., recursively visit first a vertex, then its children. At each vertex ξ, we update the $\mathcal{U}$ so that it stores the connected components of the subgraph (V, E^{*}_{ξ}). More precisely, when we reach ξ, $\mathcal{U}$ stores the connected components of $(V, E^{*}_{p(\xi)})$, so we insert the links E_{ξ} to get the subgraph (V, E^{*}_{ξ}), and for each link $e_{ij} \in E_{\xi}$, we perform MERGE (i, j) on $\mathcal{U}$, and update the value of $\hat{\chi}(\xi)$. We also maintain a copy of the changes into the data structure $\mathcal{U}$ on a stack. If ξ is a leaf corresponding to a face, then we have the subgraph (V, E^{*}_{ξ}) at our disposal. So, we output $\hat{\chi}(\xi) = \chi(\varphi)$. When we finish traversing the subtree rooted at ξ, we undo all the changes made to $\mathcal{U}$ by MERGE $(\cdot, \cdot)$ procedures.

6) Return a point from a face φ of $\mathcal{A}$ that has the minimum value of $\chi(\varphi)$.

The correctness of the algorithm follows from the invariant that for each leaf of $\mathcal{T}$ corresponding to a face ϕ of $\mathcal{A}$, the algorithm maintains E_{φ} and computes $\chi(\varphi)$. Moreover, there is a leaf in $\mathcal{T}$ for each face of $\mathcal{A}$. The expected running time of the algorithm is $\sum_{\xi \in \mathcal{T}} O(|E_{\xi}| \log m) = O(|\mathcal{A}| \log^2 m)$.

B. The maximum number of affected links — single cut case

In this section our goal is to find a point b^* for which the number of the links affected by $cut_r(b^*)$ is maximized. Finding b^* can be done using standard techniques, such as constructing $\mathcal{A}(\mathcal{H})$ explicitly. This can be obtained in time $O(m^2)$, or more precisely, in time $O(m \log m + |\mathcal{A}|)$. Similar approach was taken in [7] and thus we omit the details from this paper.

If the network topology is *planar*, a much faster *approximation* algorithm is obtainable. Recall that planarity of the graphs implies that links might meet at their endpoints but do not cross each other. Note that even in the case of a planar graph, the arrangement can have $\Omega(m^2)$ vertices. Specifically, we denote by $\Delta(b, H)$, the *depth* of a point $b \in \mathbb{R}^2$ with respect to a subset $H \subseteq \mathcal{H}$ to be the number of hippodromes in H that contain b. The depth of a set of hippodromes H is $\Delta(H) = \max_{b \in \mathbb{R}^2} \Delta(H, b)$, and let $\Delta = \Delta(\mathcal{H})$. Notice that Δ is the the number of links affected by $cut_r(b^*)$.

Using the technique of Aronov and Har-Peled [30], one can find a point b whose depth is close to Δ: For any $\varepsilon > 0$, the algorithm finds in $O(\frac{m}{\varepsilon^2} \log^2 m)$ expected time and with probability $1 - 1/m^{O(1)}$, a point b, such that $\Delta(b, \mathcal{H}) \geq (1 - \varepsilon)\Delta$. However, this technique requires an oracle procedure, named DEPTH_THRESHOLD(H, k, ℓ), whose input is a set H of k hippodromes and an integer $\ell \leq k$; if $\Delta(H, k) \leq \ell$, the procedure returns TRUE along with the point z of the maximum depth. Otherwise, it returns FALSE and a point z with $\Delta(z, H) > \ell$. Such a procedure can be implemented using a *randomized divide-and-conquer* approach that runs in $O(k\ell \log k)$ time; see [29] for details.

C. The maximum number of affected links — multiple cuts case

This section considers the maximum number of links that can be destroyed by a set of k simultaneous attacks. This is an instance of the k set cover problem [9] which is known to be NP-hard. However, a greedy strategy that picks a point that destroys as many links as possible (or, equivalently, picks a point that intersects as many hippodromes as possible), and processes the remaining links for k iterations destroys at least $1 - 1/e$ fraction of the links destroyed by any set of k attacks.

V. MULTIPLE ATTACKS WITH PROBABILISTIC OUTCOMES

In this section we consider our probabilistic failure model which assumes that a link is destroyed only with some prespecified probability $p > 0$. We then measure the *expected* number of links destroyed.

When considering only a *single* attack, the algorithms presented in Section IV-B achieve the same performance also for this model. The exact details, which are omitted from this paper for brevity, are straightforward and based on the linearity of expectation.

Thus, we study the impact on the network of k simultaneous circular cuts, whose centers are at locations $\mathcal{B} = \{b_1 \dots b_k\}$. Recall that $f_{ij}(b)$ is the probability that link e_{ij} is affected by an attack whose epicenter is b. Hence, the total number of links that are destroyed by at least one of the cuts is

$$\Phi(B) = \sum_{e_{ij} \in E} \left[1 - \prod_{b \in B} (1 - f_{ij}(b)) \right] = m - \sum_{e_{ij} \in E} \Pi(e_{ij}, B)$$

where $\Pi(e_{ij}, B) = \Pi_{b \in B}(1 - f_{ij}(b))$. Let $\mathcal{B}^* = \{b^*_1 \dots b^*_k\}$ denote the optimal solution, i.e., $\mathcal{B}^* = \arg\max_B \Phi(B)$.

Since finding $\mathcal{B}^*$ is at least as hard as the set cover problem [9], we propose the following *greedy* heuristic. Assume that $j - 1$ cuts are already present, and define the *revenue* of the cut j at location b_j to be the expected number of links that $cut_r(b_j)$ damages, among the links that survived cuts $1 \dots j - 1$; namely, $Rev(x \mid b_1 \dots b_{j-1}) = \sum_{e_{ij} \in E} [f_{ij}(x) \cdot \Pi(e_{ij}, \{b_1, \dots, b_{j-1}\})]$. Our greedy strategy is to pick, at each iteration $j = 1 \dots k$, the point b_j that maximizes the revenue, i.e., $b_j = \arg\max_{x \in \mathbb{R}^2} Rev(x \mid b_1, b_2, \dots, b_{j-1})$. Thus, our algorithm executed k times the algorithm to identify a single vulnerable point (see Section IV-B), and therefore, its running time is $O(km \log m + k|\mathcal{A}|)$. Let $\hat{\mathcal{B}} = \{\hat{b}_1 \dots \hat{b}_k\}$ denote the resulting sequence and let $\alpha = \frac{\Phi(\hat{\mathcal{B}})}{\Phi(B^*)}$ be the *approximation ratio* of this greedy strategy; Fig. 4 shows specific locations of $\hat{B}$ and B^* for $p = 1.0$ and $r = 300$ miles. We next show that $\alpha \leq (4 - p)/4$ when $k = 2$. We leave the evaluation of the case for $k > 2$ for future research.

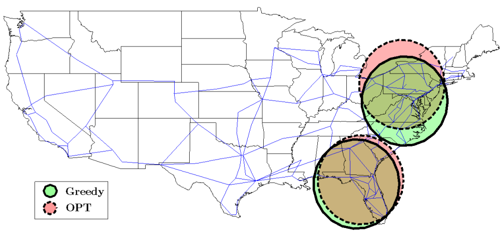


Fig. 4. Comparison of two optimal locations B^*, for p=1.0 and r=300 and the two locations $\hat{B}$, selected by the greedy algorithm.

Theorem 1. *The greedy heuristic achieves an approximation ratio of* $(4-p)/4$ *for any* p *and* $k=2$.

Proof: The proof follows by applying the inclusion-exclusion principle, where for any set of locations $\mathcal{B} = \{b_x, b_y\}$, $\Phi(\{b_x, b_y\}) = p(|S_x| + |S_y| - p|S_x \cap S_y|)$, where S_x (resp. S_y) is the set of links affected by $cut_r(b_x)$ (resp. $cut_r(b_y)$).

Recall that $\mathcal{B}^*$ is the set of (two) optimal locations and $\hat{\mathcal{B}}$ is the set of locations picked by our greedy strategy, and denote by $S_1^*, S_2^*, \hat{S}_1$, and $\hat{S}_2$ the corresponding affected link sets. By the greedy choice of $\hat{b}_1$, we get that $|\hat{S}_1| \geq \frac{1}{2}(|S_1^*| + |S_2^*|)$. Thus,

$$|\hat{S}_1| \geq \frac{1}{2}(|S_1^*| + |S_2^*| - p|S_1^* \cap S_2^*|) = \frac{1}{2p}\Phi(\mathcal{B}^*). \qquad (1)$$

Now, without loss of generality, let $\Phi(\{\hat{b}_1, b_1^*\}) \geq \Phi(\{\hat{b}_1, b_2^*\})$. Thus, $\Phi(\{\hat{b}_1, b_1^*\}) \geq \frac{1}{2}(\Phi(\{\hat{b}_1, b_1^*\}) + \Phi(\{\hat{b}_1, b_2^*\}))$, implying that $\Phi(\{\hat{b}_1, b_1^*\}) \geq \frac{p}{2}\Big(|S_1^*| + |\hat{S}_1| - p|S_1^* \cap \hat{S}_1| + |S_2^*| + |\hat{S}_1| - p|S_2^* \cap \hat{S}_1|\Big)$. However, since $|S_1^* \cap \hat{S}_1| + |S_2^* \cap \hat{S}_1| \leq |\hat{S}_1| + |S_1^* \cap S_2^*|$, we get the following approximation ratio:

$$\begin{aligned}\Phi(\hat{\mathcal{B}}) = \Phi(\{\hat{b}_1, \hat{b}_2\}) \;&\geq\; \Phi(\{\hat{b}_1, b_1^*\}) \\ &\geq p|\hat{S}_1| + \frac{p}{2}\left(|S_1^*| + |S_2^*| - p|S_1^* \cap S_2^*|\right) - \frac{p^2}{2}|\hat{S}_1| \\ &= \frac{p(2-p)}{2}|\hat{S}_1| + \frac{1}{2}\Phi(\{b_1^*, b_2^*\}) \;\geq\; \frac{4-p}{4}\Phi(\mathcal{B}^*),\end{aligned}$$

where the last inequality follows by (1).

To show that the bound on the ratio is tight, consider the example in Fig. 5. The links are all parallel and also perpendicular to a line l as shown. The two points, b_1^*, and b_2^*, that maximize Φ, affect disjoint set of links, and hence, the expected number of links destroyed is $12p$. For picking the first cut, any point b on the line l for which the corresponding $cut_r(b)$ affects six links has equal revenue. Hence, the greedy approach picks such a point arbitrarily. If it picks the point $\hat{b}_1$ shown, then the ratio $\frac{\Phi(\hat{\mathcal{B}})}{\Phi(B^*)} \geq (4-p)/4$. This is because b_1^* or b_2^* should be picked at the second step of the greedy approach to maximize the revenue. Hence, $\Phi(\{\hat{b}_1, b_1^*\}) = (12p - 3p^2)$ which implies that $\alpha = (4-p)/4$. ∎

VI. Numerical Results

We have conducted a numerical study to measure the expected number of failed links Φ when the number of cuts is 2, the failure model is probabilistic with $p = \{0.5, 0.75, 1\}$, and the cut radius ranges from 60 to 300 miles. The network under consideration is the fiber-optic network depicted in Fig. 1, and therefore, the possible locations of the two attacks of radius 60 miles are the vertices of the arrangement of the hippodromes shown in Fig. 2b. Fig. 6 shows the number of intersecting pairs of hippodromes in the arrangement and indicates that most of the intersecting hippodromes are on the east coast (recall that this number determines the running time of our algorithms). Table I shows the expected number of links destroyed by the optimal choice of two attacks (namely, $\Phi(B^*)$), by a greedy choice of two attacks (namely, $\Phi(\hat{B})$), and the resulting ratio $\alpha = \Phi(\hat{B})/\Phi(B^*)$. As can be seen, the ratio α is 1 for radii 60 through 240 miles in all the three cases of p, and it is very close to 1 for radius of 300 miles. This shows that in this fiber-optic network, the greedy algorithm produces a solution which is practically as good as the optimal one. A possible explanation for this phenomenon is the fact that in the network of Fig. 1, the links are clustered along the east coast in both the north and the south. Hence, the optimal locations tend to be disjoint for radii 60 to 180 miles and the greedy is able to successfully find these disjoint locations.

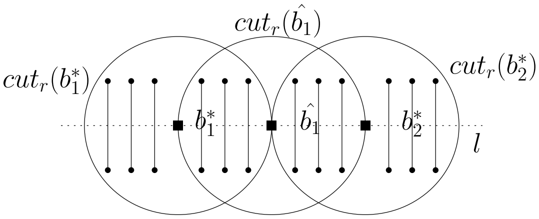


Fig. 5. Worst case performance of the greedy approach.

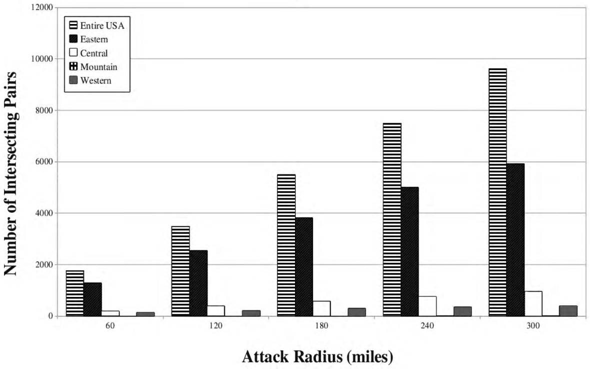


Fig. 6. The number of intersecting pairs of hippodromes for attack radii 60–300 miles. The results are based on the entire fiber-optic network presented in Fig. 1 as well as various sub-networks (determined by the different US time-zones).

VII. Conclusion

In this paper, we provided algorithms to detect the most vulnerable point(s), given a network embedded in the Euclidean plane and circular-shaped attack(s). We considered both deterministic and probabilistic scenarios as well as situations in which several attacks happen simultaneously. All our algorithms run in low-complexity polynomial-time and

TABLE I

NUMERICAL RESULTS FOR TWO ATTACKS IN THE FIBER-OPTIC NETWORK OF FIG. 1. THE EXPECTED NUMBER OF FAILED LINKS HAS BEEN EVALUATED FOR $p \in \{0.5, 0.75, 1\}$ AND THE RATIO BETWEEN THE GREEDY AND OPTIMUM α IS SHOWN. THESE MEASUREMENTS ARE TAKEN FOR CUTS OF RADIUS 60 TO 300 MILES.

r	$\mathbf{\Phi(B^*)}$	$\mathbf{\Phi(\hat{B})}$	α
60	17.25	17.25	1
120	28.5	28.5	1
180	41	41	1
240	51.5	51.5	1
300	62.75	62.5	0.996

(a) $p = 0.5$

r	$\mathbf{\Phi(B^*)}$	$\mathbf{\Phi(\hat{B})}$	α
60	25.688	25.688	1
120	42.75	42.75	1
180	61.5	61.5	1
240	77.25	77.25	1
300	92.625	90.938	0.982

(b) $p = 0.75$

r	$\mathbf{\Phi(B^*)}$	$\mathbf{\Phi(\hat{B})}$	α
60	34	34	1
120	57	57	1
180	82	82	1
240	103	103	1
300	122	120	0.984

(c) $p = 1$

significantly improve upon prior state of the art. Since these algorithms must be executed a large number of times for various possible scenarios, reducing their complexity and guaranteeing low approximation ratios significantly improves our ability to understand the possible impacts of geographically correlated attacks or natural disasters. Future research directions include generalizing our framework to different shapes of attacks (e.g. splines and polygons), considering more realistic probability models (e.g., models in which the failure probability of each link is inversely proportional to its distance from the attack's epicenter), consider more than two simultaneous attacks, and improving the running time of our algorithms using finer analysis and additional tools from computational geometry.

VIII. ACKNOWLEDGMENTS

P.A. and S.G. are supported by NSF under grants CNS-05-40347, CCF-06-35000, IIS-07-13498, and CCF-09-40671, by ARO grants W911NF-07-1-0376 and W911NF-08-1-0452, by an NIH grant 1P50-GM-08183-01, by a DOE grant OEG-P200A070505, and by a grant from the U.S.–Israel Binational Science Foundation.

A.E. and S.S. are supported by NSF CAREER grant 0348000 and NSF grant CNS-1017714.

G.Z. and D.H. are supported by DTRA grant HDTRA1-09-1-0057, NSF grant CNS-1018379, and the Legacy Heritage Fund program of the Israel Science Foundation (Grant No. 1816/10).

REFERENCES

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

[2] J. S. Foster, E. Gjelde, W. R. Graham, R. J. Hermann, H. M. Kluepfel, R. L. Lawson, G. K. Soper, L. L. Wood, and J. B. Woodard, "Report of the commission to assess the threat to the United States from electromagnetic pulse (EMP) attack, critical national infrastructures," Apr. 2008.

[3] C. Wilson, "High altitude electromagnetic pulse (HEMP) and high power microwave (HPM) devices: Threat assessments," CRS Report for Congress, July 2008. [Online]. Available: http://www.ntia.doc.gov/broadbandgrants/comments/7926.pdf

[4] W. R. Forstchen, *One Second After*. Tom Doherty Associates, LLC, 2009.

[5] Level 3 Communications, Network Map. [Online]. Available: http://www.level3.com/interacts/map.html

[6] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the impact of geographically correlated network failures," in *Proc. IEEE MILCOM*, Nov. 2008.

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

[8] S. Neumayer and E. Modiano, "Network reliability with geographically correlated failures," in *Proc. IEEE INFOCOM*, Mar. 2010.

[9] D. Hochbaum and A. Pathria, "Analysis of the greedy approach in problems of maximum k-coverage," *Naval Research Logistics (NRL)*, vol. 45, no. 6, pp. 615–627, 1998.

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

[11] C. Ou and B. Mukherjee, *Survivable Optical WDM Networks*. Springer-Verlag, 2005.

[12] O. Crochat, J.-Y. Le Boudec, and O. Gerstel, "Protection interoperability for WDM optical networks," *IEEE/ACM Trans. Netw.*, vol. 8, no. 3, pp. 384–395, 2000.

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

[14] IETF Internet Working Group , "Inference of Shared Risk Link Groups," November 2001, Internet Draft. [Online]. Available: http://tools.ietf.org/html/draft-many-inference-srlg-02

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

[16] A. F. Hansen, A. Kvalbein, T. Cicic, and S. Gjessing, "Resilient routing layers for network disaster planning," in *Proc. ICN*. Springer-Verlag, Apr. 2005.

[17] M. M. Hayat, J. E. Pezoa, D. Dietz, and S. Dhakal, "Dynamic load balancing for robust distributed computing in the presence of topological impairments," *Wiley Handbook of Science and Technology for Homeland Security*, 2009.

[18] K. Atkins, J. Chen, V. S. A. Kumar, and A. Marathe, "The structure of electrical networks: a graph theory-based analysis," *Int. J. Critical Infrastructures*, vol. 5, no. 3, pp. 265–284, 2009.

[19] C. A. Phillips, "The network inhibition problem," in *Proc. ACM STOC*, 1993.

[20] C. Burch, R. Carr, S. Krumke, M. Marathe, C. Phillips, and E. Sundberg, "A decomposition-based pseudoapproximation algorithm for network flow inhibition," in *Network Interdiction and Stochastic Integer Programming*. Springer, 2003, ch. 3, pp. 51–68.

[21] A. Schrijver, "On the history of combinatorial optimization (till 1960)," in *Handbook of Discrete Optimization*. Elsevier, 2005, pp. 1–68.

[22] A. Pinar, Y. Fogel, and B. Lesieutre, "The inhibiting bisection problem," in *Proc. ACM SPAA*, June 2007.

[23] R. L. Church, M. P. Scaparra, and R. S. Middleton, "Identifying critical infrastructure: the median and covering facility interdiction problems," *Ann. Assoc. Amer. Geographers*, vol. 94, no. 3, pp. 491–502, 2004.

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

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

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

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

[28] M. Sharir and P. Agarwal, *Davenport-Schinzel Sequences and their Geometric Applications*. Cambridge University Press, 1995.

[29] P. K. Agarwal, D. Z. Chen, S. K. Ganjugunte, E. Misiołek, M. Sharir, and K. Tang, "Stabbing convex polygons with a segment or a polygon," in *Proc. ESA*, Sept. 2008.

[30] B. Aronov and S. Har-Peled, "On approximating the depth and related problems," in *Proc. ACM-SIAM SODA*, Jan. 2005.