Graph theoretic analysis of Ad-Hoc network vulnerability

András Faragó

Graph Theoretic Analysis of Ad Hoc Network Vulnerability*

András Faragó
Department of Computer Science
The University of Texas at Dallas
Richardson, Texas, U.S.A.
E-mail: farago@utdallas.edu

Extended Abstract

1 Introduction

Currently, large ad hoc networks are applied predominantly in the tactical field. Therefore, these networks often have to be designed for operation in a hostile environment. In such an environment it is critically important that the network has as low vulnerability as possible.

How can we measure the vulnerability of the network? Attack or jamming can make some links/nodes fail and this can destroy the connectivity of the network. If the network topology (in a static snapshot) is modeled by an undirected graph, then a conventional measure is the size of a *minimum cut*, that is, the minimum number of edges that have to be removed to make the graph disconnected, which is the (edge-)*connectivity* of the graph. This number, at the same time, is the guaranteed number of edge disjoint paths between any given pair of nodes (by the classical theorem of Menger, see, e.g., [7]).

Clearly, the fewer link failures can disconnect the network, the more vulnerable it is. A minimum cut acts here as a bottleneck, a weakest part with respect to network connectivity.

Much work exists regarding network connectivity, both on the probabilistic and combinatorial aspects of the subject. In network reliability

*Supported in part by NSF Grant ANI-0220001.

investigations it is a standard approach to assume independent link failures in a fixed network topology and then compute or approximate the probability that the network stays connected (k-connected) [9]. A celebrated result on the algorithmic aspects of reliability is due to Karger [4]: despite the inherent complexity of the problem, there is a *randomized fully polynomial approximation scheme (FPRAS)* to compute the probability that the network gets disconnected.

In the ad hoc networking literature there are investigations on how the network connectivity depends on basic parameters, such as the transmission range, assuming a stochastic model on the spatial distribution of nodes, see, e.g., [1, 8]. An essential difference here to the network reliability setting is that in an ad hoc network any link can exist if its end-nodes get close enough to each other, but the existence of the various links is not independent.

On the graph theoretic side, there are known efficient algorithms to compute graph connectivity. The size of a minimum cut (which equals to the edge-connectivity of the graph) can be found in nearly linear time: the randomized algorithm of Karger [5] finds it in $O(m \log^3 n)$ time in a graph of n nodes and m edges.

In the present paper we initiate a new model of graph theoretic flavor to capture network vulnerability, motivated by new factors in ad hoc networks, as explained in the next section. The emphasys is on the general graph theoretic aspects of capturing complex failure scenarios. As a first step on the algorithmic side, we focus on the centralized computation of the relevant parameters, their distributed evaluation is left for future research.

2 Motivation for a New Model

An essential new factor in ad hoc networks is that links usually do not fail independently. On the contrary, in a number of situations various *sets* of links are prone to simultaneous failure, especially when it is due to intentional actions. A few examples:

- If an adversary broadcasts a jamming signal in a given area, then all links that are close enough are likely to fail together.
- If a node is destroyed, then all adjacent links fail.
- If various codes are used for different links (such as in CDMA), then an adversary who was able to find out one or more codes can jam the set of links that use the leaked codes.

- If links work in different frequency bands, then jamming a frequency band in a given area can inhibit all links in the area that use the given frequency band, while others may remain intact.

- If overlay logical networks exist on top of the physical network, then several logical links may use the same physical link. Then the failure of a single physical link can result in the failure of a (possibly large) set of logical links.

The above scenarios (and many conceivable similar ones) motivate the new model in the next section to capture network vulnerability. The model generalizes the concept of the minimum cut and may also be capable of capturing a number of more complex situations in other areas, as well. (For example, the vulnerability/reliability of logical overlay networks on top of any physical network.)

3 Model for Network Vulnerability

Let $G = (V, E)$ be a graph and let $\mathcal{F} = \{E_1, \ldots, E_k\}$, $\forall E_i \subseteq E$, be a family of edge sets in G. Let us call them *failure sets* (each set represents a set of links that are expected to fail together). It is assumed that each edge is contained in at least one failure set.

Let us call a subfamily $\mathcal{C} = \{E_{i_1}, \ldots, E_{i_r}\} \subseteq \mathcal{F}$ an *$\mathcal{F}$-cut*, if the removal of the union of all edge sets in $\mathcal{C}$ disconnects the network. A minimum $\mathcal{F}$-cut is an $\mathcal{F}$-cut with the minimum number of sets in it.

When we look for a minimum $\mathcal{F}$-cut, it means we look for the minimum number of failure sets that can disconnect the network. Clearly, this is a generalization of the conventional minimum cut, as if each failure set is a single edge, then the minimum $\mathcal{F}$-cut directly reduces to a minimum cut. On the other hand, if the family $\mathcal{F}$ of edge sets is more complex, then it can express much more sophisticated failure conditions.

Examples:

- Assume an adversary can place nodes that try to disconnect the network by broadcasting a jamming signal. There are k locations where the jamming nodes can be placed. Let E_i be the set of links that are disconnected if a jamming node is placed at location i. Then the size of a minimum $\mathcal{F}$-cut is the minimum number of jamming nodes that can disonnect the network.

- Let E_i be the set of links adjacent to node i. Then the minimum $\mathcal{F}$-cut is equal to the minimum number of nodes that needs to be removed in order to disconnect the network. Thus, the minimum $\mathcal{F}$-cut provides a common generalization of edge- and vertex-connectivity.

- Assume there is an overlay logical network configured on top of the physical network. Each logical link is implemented as a path in the underlying physical network. These paths can share physical links. Let E_i be the set of logical links (i.e., paths) that use physical link i. Then the minimum $\mathcal{F}$-cut is the minimum number of physical links that have to be disconnected to disconnect the logical network. Note that, because of the sharing of the physical links, potentially fewer physical links may disconnect the logical network than the logical connectivity.

- Various combinations of the above can also occur. For example, the jamming situation can be considered jointly with the overlay network scenario.

Thus, we characterize network vulnerability by the size of a minimum $\mathcal{F}$-cut, i.e., the minimum number of failure sets that can jointly disconnect the network. Further, we may also consider a *weighted version*, when a non-negative weight is assigned to each failure set, representing that the different failure sets may have different cost or importance. Then, for example, if the weights represent costs, then a minimum weight $\mathcal{F}$-cut can characterize the minimum cost needed to disconnect the network, under a complex system of potential failure events.

The first natural question is: how to find a minimum $\mathcal{F}$-cut? While conventional minimum cuts can be found by well known efficient algorithms, the new scenario raises a number of challenging novel questions. In the next section we present some positive initial results for the cardinality (i.e., unweighted) case.

4 Results on minimum $\mathcal{F}$-cuts

For notational convenience let us introduce the following notation. If C is a conventional cut, let $\mathcal{F}(C)$ be the minimum number of failure sets that cover C, we call this the $\mathcal{F}$-size of C. Clearly, the minimum $\mathcal{F}$-size that a conventional cut can have is equal to the weight of a minimum $\mathcal{F}$-cut.

Let us note first that, as one can expect, it is hard to find a minimum $\mathcal{F}$-cut if the failure sets are arbitrary, as shown by the following theorem.

Theorem 1. *For an arbitrary failure set system $\mathcal{F}$ and an input number s it is NP-complete to decide if an $\mathcal{F}$-cut of size less than or equal to s exists.*

Proof. We show that the well known *NP*-complete problem SET COVER (see, e.g., [3]) can be reduced to the task of deciding whether an $\mathcal{F}$-cut C with $\mathcal{F}(C) \leq s$ exists. Consider the following instance of SET COVER. Given a set $A = \{a_1, \ldots, a_n\}$, a family $\mathcal{H} = \{H_1, \ldots, H_k\}$ of subsets of A and a natural number ℓ, does there exist a subsystem $H_{i_1}, \ldots, H_{i_m}$, with $m \leq \ell$, such that $\cup_{j=1}^{m} H_{i_j} = A$? We can directly construct an instance of the minimum $\mathcal{F}$-cut problem to which this can be reduced in polynomial time. The simplest contruction is obtained if allow parallel edges. Let the graph have only two nodes and n parallel edges between them corresponding to the set elements $a_1, \ldots, a_n$. Let the failure set system be given by $\mathcal{F} = \{H_1, \ldots, H_k\}$ and let $s = \ell$. Since in this graph the only cut is the set of all edges, therefore, clearly, there exists an $\mathcal{F}$-cut with $\mathcal{F}(C) \leq s$ if and only if the answer to the question of the SET COVER instance is yes.

□

A next natural step is to look for restrictions that make the minimum $\mathcal{F}$-cut problem solvable in polynomial time. A trivial restriction is when each failure set consists of a single edge, then we get back the traditional minimum cut problem that can be solved very efficiently. Now we show that the problem still remains efficiently solvable if the size of each failure set is bounded by a constant, but otherwise the failure set system can be arbitrary. We have to assume, however, that a subroutine is available to compare the $\mathcal{F}$-sizes of any two cuts in the input graph.

Theorem 2. *Assume that a subroutine is available to compare the $\mathcal{F}$-sizes of any two given cut in the input graph and calling this subroutine is counted as a single step. Then, if the size of each failure set is bounded by a number r, a minimum $\mathcal{F}$-cut can be found by a polynomial-time algorithm of running time $O(n^{2r} \log^2 n)$ in a graph of n nodes. In particular, if the sizes of failure sets is bounded by a constant, then a minimum $\mathcal{F}$-cut can be found in polynomial time.*

Proof. Let $\mathcal{F} = \{E_1, \ldots, E_k\}$ be the failure set system. By assumption, $|E_i| \leq r$, $i = 1, \ldots, k$, holds for the constant r. Let C_0 be a cut with minimum $\mathcal{F}$-size and let C_1 be a minimum cut in the traditional sense. Since C_0 has minimum $\mathcal{F}$-size, we have $\mathcal{F}(C_0) \leq \mathcal{F}(C_1)$. As each failure set can cover at most r edges, we also have $|C_0| \leq r\mathcal{F}(C_0)$. Finally, for any cut C, no more than $|C|$ failure sets are needed to cover it, i.e., $\mathcal{F}(C) \leq |C|$ holds.

Combining these inequalities, we obtain

$$|C_0| \leq r\mathcal{F}(C_0) \leq r\mathcal{F}(C_1) \leq r|C_1|. \tag{1}$$

Let us call a cut *r-minimum* if its size is at most r times the size of a minimum cut (in the traditional sense). Then inequality (1) implies that, under the conditions of the theorem, any cut with minimum $\mathcal{F}$-size must be among the r-minimum cuts. Now we can use the algorithm of Karger and Stein [6] that can find all r-minimum cuts in $O(n^{2r}\log^2 n)$ time. It also involves the remarkable fact that there are only $O(n^{2r})$ r-minimum cuts for any $r \geq 1$. Note that $O(n^{2r})$ is polynomially bounded for constant r, even though the total number of cuts may well be exponential. Having obtained all r-minimum cuts, we can select one of them with minimum $\mathcal{F}$-size in linear time using the available subroutine, thus, the entire running time remains within the $O(n^{2r}\log^2 n)$ bound, whcih is polynomial for constant r.

□

Now let us turn to the general case, when the failure set system $\mathcal{F}$ is arbitrary. One could expect that in this case we cannot do any better than essentially performing an exhaustive search. Quite interestingly, this is not the case. Below we show that even for the general case (despite its *NP*-hardness) one can find a minimum $\mathcal{F}$-cut faster than exponential time. Essentially, the running time exponent will grow only with the square root of the graph size, rather than linearly, which is much faster than exhaustive search.

Theorem 3. *Assume that a subroutine is available that can compare the $\mathcal{F}$-sizes of any two cuts in the input graph and calling this subroutine is counted as a single step. Then, If the graph has m edges and $k = O(m^b)$ arbitrary failure sets, then a cut with minimum $\mathcal{F}$-size can be found in $2^{O(b\sqrt{m}\log m)}$ time.*

Proof. The algorithm works as follows. Let us do first an exhaustive search among all possible systems of at most $\sqrt{m}$ failure sets. The number of such systems is bounded by $k^{\sqrt{m}} = 2^{O(b\sqrt{m}\log m)}$. If one of them separates the graph, then the smallest such system is a minimum $\mathcal{F}$-cut. The reason is that we search all possible families consisting of at most $\sqrt{m}$ failure sets, so if there is a smaller $\mathcal{F}$-cut, that one would also be found in this search.

Thus, if the above search returns any $\mathcal{F}$-cut, then we are done. It remains to consider the case when the minimum $\mathcal{F}$-cut has more than $\sqrt{m}$ failure sets, in which case the above search does not return any $\mathcal{F}$-cut.

Let C be a any cut. If the first search did not find any $\mathcal{F}$-cut, then $\mathcal{F}(C) > \sqrt{m}$ must hold, as we have searched all systems with $\leq \sqrt{m}$ failure sets and none of them contained a cut. Since $\mathcal{F}(C) \leq |C|$ always holds, therefore, we have $|C| > \sqrt{m}$. Thus, all cuts have more than $\sqrt{m}$ edges. Just as in the proof of Theorem 2, let us call a cut r-minimum if it has at most r times more edges than a (conventional) minimum cut. Since obviously no cut can have more than m edges, therefore, we have that now all cuts fall in the set of r-minimum cuts with $r = m/\sqrt{m} = \sqrt{m}$. All these cuts can again be generated by the Karger-Stein algorithm [6] in $O(n^{2r} \log^2 n) = 2^{O(b\sqrt{m} \log m)}$ time and a smallest one among them can be found in linear time using the comparison subroutine. Since both the initial exhaustive search and the second phase have the same complexity and the needed side computations clearly fit in this complexity, the proof is completed.

□

5 Path Vulnerability

So far we have dealt with the genralization of the minimum cut concept, which is a *global* characterization of network vulnerability. Often, however, it is important to capture the vulnerability of a *path*, rather than the entire network. Then we may want to find a *least vulnerable path* for safely routing the messages.

Let us consider again a system $\mathcal{F} = \{E_1, \ldots, E_k\}$ of failure sets, just as in the previous sections. Let P be a path. Observe that the definition of the $\mathcal{F}$-size now does not make sense in the same form as we used for cuts. The reason is that for a path the important thing is not that how many failure sets cover the entire path. It is more important that how many failure sets *intersects* with it, i.e., contains a link from the path. The reason is that if a failure set intersects with the path, then the failure of this set will disconnect the path. Thus, the more failure sets itersects with the path, the more vulnerable it is.

According to the above reasoning, let us define the vulnerability of a path P as the number of failure sets that intersect with the path and let us call it the *vulnerability measure* of P, denoted by $\mathcal{F}(P)$. In general, we consider the weighted case, when a positive weight is assigned to each failure set and each contributes to $\mathcal{F}(P)$ with its weight.

Having defined the vulnerability measure of paths, it is natural to consider the search for a least vulnerable path. This generalizes the shortest

path problem, since if each failure set is a single edge. then $\mathcal{F}(P)$ is the number of hops in th eunweighted case and the weight (length, cost) of the path in the weighted case.

At this point one may wonder how expressive is the vulnerability measure, that is, how complex can be the path metrics that are expressible this way? Note that a path metric can be very sophisticated, it may depend on complicated properties of the entire graph. To answer this question we state an interesting result below. This result says that essentially *any* path metric, no matter how complicated, can be expressed as a vulnerability metric, that is, in the relatively simple form introduced above.

Theorem 4 (Path Metric Representation Theorem) *Let G be a (directed or undirected) graph with two distinguished vertices u, v that serve as endpoints of the considered paths. Let h be an arbitrary path metric that assigns a positive, but otherwise arbitrary, value $h(P)$ to every u-v path P in G. Then there exists a family $\mathcal{F}$ of weighted failure sets in G and a constant M, such that the vulnerability measure $\mathcal{F}(P)$ represents the path metric $h(P)$ for every u-v path in G, up to constant translation, i.e.,*

$$\mathcal{F}(P) = h(P) + M$$

holds for every u-v path in G.

In other words, the theorem says that every path metric, no matter how comlicated, is a vulnerability measure, apart from a cosntant translation.

Due to space limitations, we omit the proof of Theorem 4 in this extended abstract. The interested reader can find it, along with more material and open questions on path problems in [2].

6 Open Problems

There are a number of intriguing open questions regarding the introduced model, both for cut and path problems. A few examples are listed below.

1. Can the running time exponent in the general minimum $\mathcal{F}$-cut algorithm (addressed in Theorem 3) be further decreased, given a subroutine that can compare the $\mathcal{F}$-sizes of any two cuts in the input graph?

2. Consider the paths between two nodes u, v. Two such paths are called $\mathcal{F}$-disjoint if no failure set contains an edge from both, that is, a single

failure set E_i cannot disconnect *both* paths. Then a natural question is to search for a maximum set of pairwise $\mathcal{F}$-disjoint u-v paths. It is a direct generalization of both conventional edge- and vertex-disjoint paths and, in a sense, this would be a least vulnerable route system between u and v. Is it possible to prove similar positive results for the maximum number of $\mathcal{F}$-disjoint paths as for the minimum $\mathcal{F}$-sized cuts or is the path problem fundamentally harder?

3. Between two given nodes u, v, the minimum $\mathcal{F}$-cut must be at least as large as the maximum number of pairwise $\mathcal{F}$-disjoint paths, since each failure set can disconnect at most one of the paths. Which are the cases when equality holds here, thus providing an appealing max flow–min cut type of characterization for those cases?

4. For which failure set systems can the least vulnerable path problem be solved in polynomial time? A trivial example is when each failure set is a single edge, then we get back the classical shortest path problem. If the size of each failure set is bounded by a number r, then an r-approximation can be found in polynomial time [2]. Compare this with Theorem 2, which finds the exact optimum for cuts for bounded failure sets. Can the same be achieved for paths?

References

[1] C. Bettstetter, "On the Minimum Node Degree and Connectivity of a Wireless Multihop Network", *MOBIHOC'02*, June 9-11, Lausanne, Switzerland, pp. 80-91.

[2] A. Faragó, "Algorithmic Challenges in Ad Hoc Networks", to appear as a book chapter in *Ad Hoc Networking*, IEEE Press and Kluwer, Eds. S. Basagni and M. Conti.

[3] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, San Francisco, CA, 1979.

[4] D.R. Karger, "A Randomized Fully Polynomial Approximation Scheme for the All Terminal Network Reliability Problem", 27^{th} *Annual ACM Symp. on the Theory of Computing*, Las Vegas, Nevada, May 1995, pp. 11-17.

[5] D.R. Karger, "Minimum Cuts in Near-Linear Time", *Journal of the ACM*, 47(2000/1), pp. 46-76.

[6] D.R. Karger and C. STein, "A New Approach to teh Minimum Cut Problem", *Journal of the ACM*, 43(1996/4), pp. 601-640.

[7] L. Lovász and M.D. Plummer, *Matching Theory*, North-Holland and Academic Publisher, Amsterdam, Netherlands, and Budapest, Hungary, 1986.

[8] T.K. Philips, S.S. Panwar and A.N. Tantawi, "Connectivity Properties of a Packet Radio Model", *IEEE Trans. Inf. Theory*, 35(1989), pp. 1044-1046.

[9] F. Roberts, F. Hwang and C. Monma (Eds.), *Reliability of Computer and Communication Networks,* DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 5, American Mathematical Society, 1991.