REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
14-04-2011	Proceedings	MAR 2011 - APR 2011

4. TITLE AND SUBTITLE

Dedicated vs Distributed: A Study of Mission Survivability Metrics

5a. CONTRACT NUMBER
FA8720-05-C-0002

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Hamed Okhravi, Andrew Johnson, Joshua Haines, Travis Mayberry, and Agnes Chan

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

MIT Lincoln Laboratory
244 Wood Street
Lexington, MA 02420

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

US STRATCOM
J843
Offutt Air Force Base, NE 68113

10. SPONSOR/MONITOR'S ACRONYM(S)

US STRATCOM

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A traditional trade-off when designing a mission critical network is whether to deploy a small, dedicated network of highly reliable links (e.g. dedicated fiber) or a large-scale, distributed network of less reliable links (e.g. a leased line over the Internet.) Previous work on this topic has widely focused on two approaches: probabilistic modeling of network reliabilities and graph theoretic properties (e.g. minimum cutset) The reliability metrics do not quantify the robustness, the ability to tolerate multiple link failures, in a distributed network. For example, 8 fully redundant network and a single link can have the same overall source-destination reliability (0.9999), but they have very different robustness. Many proposed graph theoretic metrics are not sufficient to capture network robustness either; i.e. two networks with identical metric values (e.g. minimum cutset) can have different resilience to link failures. More importantly, previous efforts have mainly focused on the source-destination connectivity and in many cases it is difficult to extend them to a general set of requirements. In this work, we study network-wide metrics to quantitatively compare the mission survivability of different network architectures when facing malicious cyber attacks. Specifically, we define a metric called relative importance (RI), a robustness metric for mission critical networks, and show how it can be used to both evaluate mission survivability and make recommendations for its improvement. Our metric can be evaluated for an arbitrarily general set of mission requirements (not just source-destination connectivity); hence, it quantifies the mission survivability of different network architectures. Finally, we study the probabilistic and deterministic algorithms to quantify the RI metric and empirically evaluate it for sample networks.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Zach Sweet
U	U	U	SAR	6	19b. TELEPHONE NUMBER *(include area code)* 781-981-5997

Dedicated vs Distributed: A Study of Mission Survivability Metrics*

Hamed Okhravi, Andrew Johnson, Joshua Haines
MIT Lincoln Laboratory
Massachusetts Institute of Technology
{hamed.okhravi, ajohnson, jhaines}@ll.mit.edu

Travis Mayberry, Agnes Chan
College of Computer and Information Science
Northeastern University
{travism, ahchan}@ccs.neu.edu

***Abstract*—A traditional trade-off when designing a mission critical network is whether to deploy a small, dedicated network of highly reliable links (e.g. dedicated fiber) or a large-scale, distributed network of less reliable links (e.g. a leased line over the Internet.) Previous work on this topic has widely focused on two approaches: probabilistic modeling of network reliabilities and graph theoretic properties (e.g. minimum cutset.) The reliability metrics do not quantify the robustness, the ability to tolerate multiple link failures, in a distributed network. For example, a fully redundant network and a single link can have the same overall source-destination reliability (0.9999), but they have very different robustness. Many proposed graph theoretic metrics are not sufficient to capture network robustness either; i.e. two networks with identical metric values (e.g. minimum cutset) can have different resilience to link failures. More importantly, previous efforts have mainly focused on the source-destination connectivity and in many cases it is difficult to extend them to a general set of requirements. In this work, we study network-wide metrics to quantitatively compare the mission survivability of different network architectures when facing malicious cyber attacks. Specifically, we define a metric called relative importance (RI), a robustness metric for mission critical networks, and show how it can be used to both evaluate mission survivability and make recommendations for its improvement. Our metric can be evaluated for an arbitrarily general set of mission requirements (not just source-destination connectivity); hence, it quantifies the mission survivability of different network architectures. Finally, we study the probabilistic and deterministic algorithms to quantify the RI metric and empirically evaluate it for sample networks.**

I. Introduction

When evaluating different architectures for a mission critical warfighting or military network, it is important to quantify the probability of mission success in the face of typical failures or cyber attacks. A traditional trade-off in this area is whether to deploy a dedicated network of hardened links (e.g. dedicated, protected fiber) or a distributed network of less reliable links (leased line over the Internet.) The dedicated network can provide higher assurance links, but it is costly. The distributed network, on the other hand, is cheaper since the infrastructure already exists, but its links can be less reliable and more easily attacked.

Given a set of complex mission requirements, however, one cannot easily compare two network architectures. Specifically we are interested in two quantities: 1) the probability that the mission requirements are satisfied under the typical failure rates (reliability), and 2) the probability that the mission requirements are satisfied given a number of components fail however low their typical failure rates are (robustness.) Note that the second quantity is especially useful when facing malicious cyber attacks because reliable links with low typical failure rates may still be attacked.

Related efforts on this topic have mainly focused on two approaches: probabilistic modeling and graph theoretic properties, but they are not sufficient for two reasons. First, different networks with identical metric values can behave differently in reality. For example, two networks with the same minimum cutset values can have different robustness (as defined above.) More importantly, they focus on a limited type of requirement (e.g. source-destination connectivity) and are not easily extensible. In this work, we define a mission success metric which can measure the robustness of a mission critical network given an arbitrarily general set of mission requirements. The metric which we call relative importance (RI) can quantify the mission survivability of a network when facing malicious cyber attacks. We study the deterministic and probabilistic algorithms to efficiently calculate the RI metric for relatively large networks. We also empirically evaluate this metric for sample networks. The most important aspect of this work is the focus on the mission. Contrary to other work, we evaluate the survivability metric for a given mission over the network, not for the network alone. This ensures that the focus is given to what the network *does* instead of what it *is*. Evaluating the mission survivability emphasizes those links in the networks that are crucial in performing the mission while it deemphasizes the others.

Our contributions are as follows:

1) We define a network mission survivability metric for arbitrarily general mission requirements.
2) We study efficient probabilistic and deterministic algorithms to calculate the metric.
3) We implement a fast algorithm and empirically evaluate the metric for a set of sample networks.

The rest of the paper is organized as follows. Section

*This work is sponsored by the Department of Defense under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

II provides an overview of the related work. Section III describes the relative importance metric and the probabilistic and deterministic algorithms to calculate it. We empirically evaluate the metric for sample networks in Section IV before concluding the paper in Section V.

II. Related Work

There have been many metrics posed to evaluate the survivability of networks based on different properties, but they can mostly be divided into two categories: probabilistic measures and those based on graph theoretic properties of a network.

A. Probabilistic Metrics

A well-studied and widely used metric for comparing the survivability of different networks is reliability. Consider a graph G, with edges $\{e_1, ..., e_n\}$. Each edge e_i is in an UP state with probability p_i. It is assumed that links in an UP state will successfully relay all communications going through them and links in a DOWN state will not relay any communications. That is, each link has a *reliability* of p_i. The probability that two chosen nodes s and t are connected at any given point in time is then called the 2-terminal reliability of G with respect to s and t. A generalization of the problem allows for a choice of k nodes and asks for the probability that all k nodes are connected.

This is a very useful and intuitive metric for measuring the survivability of a network under normal operating conditions. It allows for a whole-network reliability measure to be computed from individually defined link and node reliabilities which can be estimated through stress tests and benchmarks of their respective components.

A simple algorithm to solve this problem proceeds as follows:

1) Enumerate all paths $\{l_1, ..., l_n\}$ from s to t
2) For each path, its reliability $r_i = \prod_j p_j$
3) The 2-terminal reliability of G can then be calculated as $\phi(G) = 1 - \prod_i 1 - r_i$

Unfortunately, the number of paths from s to t is exponential in the size of the graph so this algorithm is not tractable for any graph of a useful size. It has been proven that both 2-terminal and k-terminal reliability are in NP-hard and thus unlikely to be solved exactly by any polynomial time algorithm [1]. There are, however, special types of graphs that can be solved efficiently. A tree graph has exactly one path from any node s to any other node t, so reliability is simply the product of the reliability of the edges along that path. It has also been shown that graphs which are not of a special form that is easily solved can sometimes be reduced to such a form through series-parallel reductions and application of a factoring theorem [2] [3]. These approaches do not work for all cases of graphs, and applying the factoring theorem successfully is itself a potentially difficult problem. It is also worth noting that the factoring theorem only holds for a small group of connectivity requirements such as source-to-sink and k-terminal. It does not hold for the more complicated notions of connectivity described above.

To make the problem tractable for larger graphs there have been Monte Carlo methods developed that can achieve a close approximation of reliability in much less time [4]. The basic approach is to instantiate the network by generating a random number n for each link in the range $\{0, 1\}$. $\forall i : e_i$is in G $\Leftrightarrow$ $n \leq p_i$ then e_i is added to the graph. In this way, a single moment in time for the network is simulated. If there are links up such that s and t are connected, then a counter U is incremented, otherwise D is incremented. This procedure is repeated some set number of times and the reliability is calculated as $U/D + U$.

B. Graph Theoretic Metrics

One of the earliest used graph theoretic metrics for assessing survivability is minimum vertex degree [5]. The node in a network with minimum degree is considered the weakest, and its degree is used as a comparable metric for determining how survivable the graph is (since at least that many links must be cut before any vertex becomes disconnected). Minimum degree can be found very quickly, but that is the only advantage minimum vertex degree has. This metric is very primitive and really only provides any insight if your network requires complete connectedness of all nodes to survive.

A similar but more advanced metric is the minimum cutset. A cutset is a set of links $c \in E$ such that$G = (G - c, V)$ is not connected. If C is the set of all cutsets, then the minimum cutset of a graph is $\hat{c} \in C : \forall x \in C|\hat{c}| \leq |x|$. The size of the minimum cutset in a graph with 2-terminal connectivity is equal to the maximum flow from one terminal to the other over an equivalent flow graph where all edges have unit capacity. Using the Ford-Fulkerson algorithm this can be found efficiently. The minimum cutset is particularly useful because it gives the minimum number of links that need to be subsumed in order to disrupt network connectivity [6]. However, two networks with the same size minimum cutset could have different degrees of survivability depending on how robust the remaining network is. If a network has one small size cutset but all other cutsets are much larger, it would be easier to shore up that weak spot and increase the survivability substantially than it would be in a network with many small cutsets.

There have also been other metrics based on cutsets that measure the survivability under specific scenarios. One such scenario is when the network is considered connected if it has at least a k of its nodes still connected to each other. In this case, a small set of weak cutsets can be found and used to to evaluate survivability without having to enumerate all cutsets [7] (of which there are an exponential number).

The main weakness of many graph theoretic measures, including those based on cutsets, is that the algorithms for determining them need to be tailored to the connectivity notion used, and may not be easily adapted for some more complicated schemes.

III. Relative Importance

Reliability can give the expected uptime of a network but it does not specify which links are most vulnerable or can be improved to gain survivability. Minimum cutset does produce a set of most vulnerable links in a network but it cannot be easily adapted to arbitrary notions of connectivity and only identifies a single set of vulnerable links.

Given the drawbacks of the existing methods discussed above, we propose a new metric, based on the BIM, that:

1) Can be efficiently computed for arbitrary graphs.
2) Is practical to implement and run on current hardware for graphs of usefully large size.
3) Shows relative importance of links to the connectivity of the graph.
4) Can be easily adapted to arbitrary notions of connectivity.

A. Birnbaum Importance Measure

The Birnbaum Importance Measure is a way of assessing the relative importance of links in a graph to its reliability. If Ψ is a function that calculates the reliability of a network given a list of individual link reliabilities $\{p_1, ..., p_n\}$, the BIM is defined as:

$$BIM_j = \frac{\partial \Psi(p_1, ..., p_n)}{\partial p_j} = \Psi(p_1, ..., p_{j-1}, 1, p_{j+1}, ..., p_n) - \Psi(p_1, ..., p_{j-1}, 0, p_{j+1}, ..., p_n) \quad (1)$$

Intuitively this is the difference in reliability between a network with component j replaced by an infallible component and one where j is removed entirely. Calculating the BIM for all values of j creates an importance spectrum whereby the contributions to network reliability of each link can be compared. This is a better metric than minimum cutset for determining network robustness because it allows for a more fine grained analysis of a graphs reliability. Directly calculating BIM would require $O(n)$ exact solutions for the reliability of the graph in question, which are each difficult to compute.

Fortunately, a connection between Ψ and the network spectrum (first described by Gertsbakh [8]) provides an alternate approach to efficiently estimate the BIM of a network. A network with n links has $n!$ possible permutations of those links, each of the form $\pi = l_1, l_2, ..., l_n$. If the network is thought of with all links starting off down, they can be brought up one by one according to the order they appear in π. There exists some link l_i such that before adding l_i the network is disconnected and becomes connected upon adding l_i. In this case, i is called the *anchor* of π. The *network spectrum* can then be written as

$$C = x_1, x_2, ..., x_n \quad (2)$$

such that x_i is the number of link permutations with anchor i. Then, the *cumulative spectrum* is

$$Y_b = \sum_{i=1}^{b} x_i \quad (3)$$

Gertsbakh has proven that the reliability of a network can be written in terms of the cumulative spectrum as

$$\Psi(G) = \sum_{i=1}^{n} Y_i \frac{p^i q^n - i}{i!(n-i)!} \quad (4)$$

This result can be used to express the BIM of individual components in terms of the cumulative spectrum as

$$BIM_j = \sum_{i=1}^{n} \frac{Z_{i,j} p^{n-i} - (Y_i - Z_{i,j}) p^i q^{n-i-1}}{i!(n-i)!} \quad (5)$$

where $Z_{i,j}$ is the number of permutations where the network is connected after the first i edges and component j is among those edges used. Intuitively, $Z_{i,j}$ will be close to Y_i for a component j that is very important to the network (i.e. it will be in most of the permutations of that size resulting in a connected network) and lower relative to Y_i for components that are less important.

At first, equation (5) not seem very helpful because there are an exponential number of permutations. However, Y and Z can be efficiently approximated with a Monte Carlo sampling method. Instead of enumerating all possible permutations (of which there are $n!$), $\hat{Y}$ and $\hat{Z}$ can be calculated with a random sampling of size m and then scaled by $\frac{n!}{m}$ before calculating BIM [9].

B. New Metric

Since we are concerned with unpredictable component failure (i.e. sabotage, attack), the reliability terms used in calculating BIM are not necessary for our new metric. In defining Relative Importance, we are interested in which links are necessary to put the network into a connected state for each permutation. Let X_j be the number of permutations where the index of component j is less than or equal to the anchor of that permutation. That is, the number of permutations where component j is required for the network to be connected. Relative Importance is then defined as

$$RI_j = \frac{X_j}{n!} \quad (6)$$

We propose that this is a useful metric for determining the importance of each link in the network because important links will be required more often than less important links to make the network connected, and therefore will be at a position less than the anchor in more permutations than other less important links. For instance, a link that is not part of any path that causes the network to be connected will be counted in a permutation only with a probability proportional to the average permutation anchor. That is, it will only appear coincidentally so if the average anchor is $\bar{\alpha}$ then the probability that it is in a location less than $\bar{\alpha}$ in a permutation is $\bar{\alpha}/n$. On the other hand, if a link is very important it will be part of the set of links that connect the network much more often.

C. Algorithms

The RI of a component can be calculated deterministically, like BIM, by enumerating all permutations and calculating X_j for each component. Also like BIM, RI can be approximated using Monte Carlo sampling by calculating $\hat{X}$ from m. The exact algorithm is as follows

1) Initialize all x_i to zero for $i = 1, ..., n$
2) Randomly sample permutation π from the set of all link permutations
3) Bring the network up one link at a time from π until the network is connected, after j links
4) Increment all x_i for all $i \leq j$
5) Repeat 2-4 m times
6) Approximate $\hat{X}_i = x_i * n!/m$

In practice, the $n!$ terms in the approximation and equation (6) will cancel and the RI can be calculated as

$$RI = \frac{x_i}{m} \tag{7}$$

Using this algorithm, RI can be computed with respect to any definition of connectivity by using an appropriate function to check in step 3. The runtime of each iteration of step 3 is $O(\alpha g)$ where α is the anchor of π and g is the running time of the function that checks for connectivity. For instance, using standard source-to-sink connectivity running time can be calculated with $g = |V|$ using breadth first search. Since α is at most E this makes the running time of each iteration $O(|E||V|)$ and the whole algorithm $O(|E||V|m)$. In practice, as with many sampling algorithms, a good approximation can be obtained with m being proportional to the size of the sample (in this case $|E|$) bringing the total time to $O(|E|^2|V|)$. Additionally, with 2-terminal connectivity (and some other similar schemes) a disjoint set data structure can be used instead of BFS to make $g = O(1)$ amortized. This makes the algorithm actually $O(|E|^2)$ which is very efficient compared to other metrics.

IV. Evaluation

In gauging the effectiveness of RI, we will present the RI spectrum for several small networks to demonstrate that it matches the intuition of robustness, as well as a larger network representing a possible real world situation. The first network is a star consisting of five nodes as shown in figure 1.

We would expect the RI spectrum to be flat since each node is isomorphic to the others. No link is any more important that the others since it is a fully connected network. Additionally, we would expect the RI of each link to be significantly less than one since there is so much redundancy in the network. The results of our algorithm can be seen in figure 2.

In presenting the problem we noted that there is a debate over the advantages of distributed and dedicated networks. Figure 3 shows a possible small scale dedicated network. With connectivity defined between nodes 1 and 6 it is easy to see that the (1,2) and (5,6) links are are more important than the links in the center. In fact, we would expect the RI of those links to be about twice that of the remaining links

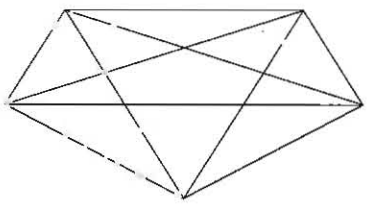

Fig. 1. Star network

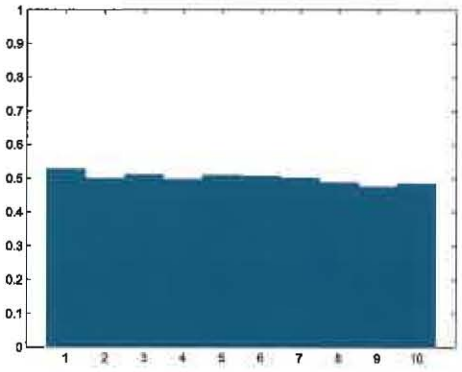

Fig. 2. Star graph RI

since there are two distinct paths through the center diamond portion. Figure 4 shows that this is in fact the case. Since (1,2) and (5,6) are necessary for every instance of the graph that is connected, their importance is one, with the remaining links being around .5.

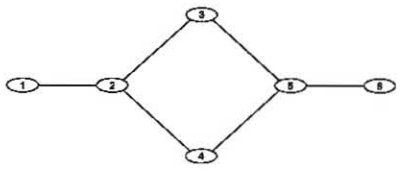

Fig. 3. Graph representing a possible "dedicated" network

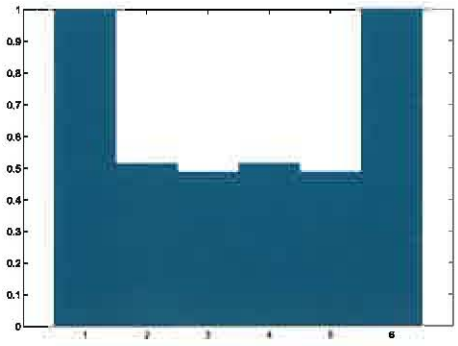

Fig. 4. Dedicated network RI

Fig. 5. Procedurally generated 2000 node network

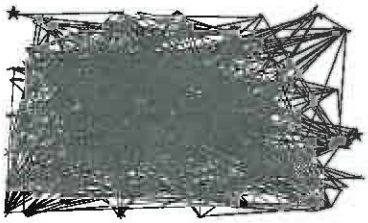

Figure 5 shows a large scale network that represents a possible real world situation. For our test, two nodes s and t (denoted by stars) were chosen on opposite sides of the network for the connectivity check. It is immediately apparent from figure 6 that one link is much more important that the others. Looking at the graph in detail shows that this link is the only one connected to t. Adding a redundant link as shown in figure 7 reduces the importance of that link significantly.

Figure 9 shows the relative importance of the same network with a different connectivity requirement. The network is considered connected if each of three individual connectivity requirements are met: the same original two nodes must be connected, a different set of three nodes must all be connected and another set of three nodes must have two of them out of three connected. This demonstrates our algorithms ability to work with arbitrary connectivity requirements.

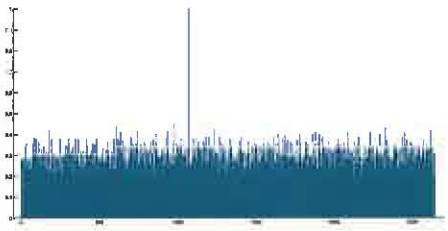

Fig. 6. BIM for 2000 node network

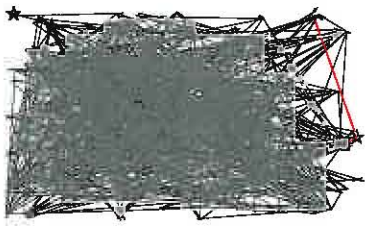

Fig. 7. Previous network with additional links added

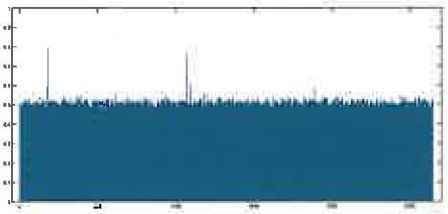

Fig. 8. New RI

V. Conclusion

We have introduced the problem of network robustness and discussed its applications to real world scenarios such as cyber attack, in the process examining existing network metrics and demonstrating why they are insufficient to fully capture the robustness of a network. From the Birnbaum Importance Measure we have adapted a new Relative Importance metric that better evaluates the robustness of a network and leads to a comparison spectrum that can be used to make command decisions about changes in network topology. We have shown that RI can be computed efficiently for arbitrary notions of connectivity and that its results match with intuitive ideas of network robustness.

References

[1] A. Rosenthal, "Computing the reliability of complex networks," *SIAM Journal on Applied Mathematics*, vol. 32, no. 2, pp. 384–393, Mar. 1977, ArticleType: research-article / Full publication date: Mar., 1977 / Copyright © 1977 Society for Industrial and Applied Mathematics. [Online]. Available: http://www.jstor.org/stable/2100423

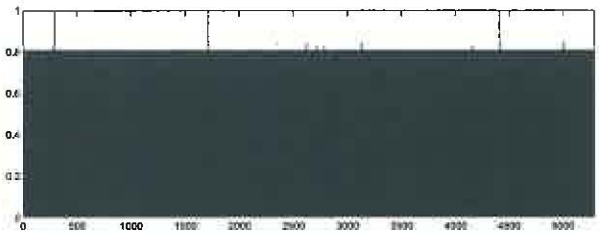

Fig. 9. RI with multiple connectivity requirements

[2] A. Satyanarayana and M. K. Chang, "Network reliability and the factoring theorem," *Networks*, vol. 13, no. 1, pp. 107–120, 1983. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/net.3230130107/abstract

[3] R. K. Wood, "Factoring algorithms for computing K-Terminal network reliability," *Reliability, IEEE Transactions on*, vol. 35, no. 3, pp. 269–278, 1986.

[4] H. Cancela and M. E. Khadiri, "A recursive variance-reduction algorithm for estimating communication-network reliability," *Reliability, IEEE Transactions on*, vol. 44, no. 4, pp. 595–602, 1995.

[5] H. Frank and I. Frisch, "Analysis and design of survivable networks," *Communication Technology, IEEE Transactions on*, vol. 18, no. 5, pp. 501 –519, october 1970.

[6] R. Wilkov, "Analysis and design of reliable computer networks," *Communications, IEEE Transactions on*, vol. 20, no. 3, pp. 660 – 678, jun 1972.

[7] L. Wu and P. Varshney, "On survivability measures for military networks," in *Military Communications Conference, 1990. MILCOM '90, Conference Record, A New Era. 1990 IEEE*, oct 1990, pp. 1120 –1124 vol.3.

[8] I. Gertsbakh and Y. Shpungin, "Network reliability importance measures: combinatorics and monte carlo based computations," *W. Trans. on Comp.*, vol. 7, pp. 216–227, Apr. 2008. [Online]. Available: http://portal.acm.org/citation.cfm?id=1457949.1457960

[9] I. Gertsbakh, *Models of network reliability : analysis, combinatorics, and Monte Carlo*. Boca Raton: CRC Press, 2010.