Computing the Impact of Disasters on Networks

Oostenbrink, Jorik; Kuipers, Fernando

DOI
10.1145/3152042.3152075

Publication date
2017

Document Version
Accepted author manuscript

Published in
ACM SIGMETRICS Performance Evaluation Review

Citation (APA)
Oostenbrink, J., & Kuipers, F. (2017). Computing the Impact of Disasters on Networks. *ACM SIGMETRICS Performance Evaluation Review*, *45*(2), 107-110. https://doi.org/10.1145/3152042.3152075

Computing the Impact of Disasters on Networks

Jorik Oostenbrink and Fernando Kuipers
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
jorik@oostenbrink.org, F.A.Kuipers@tudelft.nl

ABSTRACT

In this paper, we consider the vulnerability of a network to disasters, in particular earthquakes, and we propose an efficient method to compute the distribution of a network performance measure, based on a finite set of disaster areas and occurrence probabilities. Our approach has been implemented as a tool to help visualize the vulnerability of a network to disasters. With that tool, we demonstrate our methods on an official set of Japanese earthquake scenarios.

1. INTRODUCTION

Over the past few decades, communication networks have been used more and more in commerce, government, and in our personal lives. But with an increased usage also comes an increased dependency. Currently, a failure in our communication systems can have a significant impact on society. During and following a disaster, communication networks become even more important, as they are used for timely communication between emergency services and for deploying and coordinating relief operations.

In 2011, a massive earthquake struck Japan; the earthquake and subsequent tsunami not only resulted in a massive loss of human lives, but also caused widespread connectivity problems. Meanwhile, peak communications traffic was 9 times as high as normal. Clearly, large-scale disasters, like earthquakes, form a formidable challenge in network design. Not only do they cause hardware failures spread over a large area, but they also prevent repairs for a significant amount of time. Computing the vulnerability of a network against earthquakes can help in (1) preparing for potential earthquakes and (2) in designing or modifying a network to become more robust against them.

Although considered as early as 1991 by D. Bienstock [2], research on so-called geographically correlated challenges has only really taken of in the last decade. Often it is assumed that the disaster area takes a fixed shape (e.g., a line or a circle with fixed radius), after which the amount of disasters required to disconnect two nodes (e.g., [6]), the most vulnerable spot(s) of the network against this type of disaster (e.g., [7]), or the impact after a randomly placed disaster (e.g., [8]) are computed. Typically, the vulnerability of the network is reflected by a single value.

We take a different approach: based on a set of possible disasters (of varying shapes), we compute the distribution of the measure after one of these disasters randomly occurs. We show that this distribution can be efficiently calculated, and that it provides more information to a network operator or designer than any single (e.g., expected) value could.

Our main contributions in this paper are threefold:

- We propose an efficient method to compute the distribution of a network performance measure, based on a finite set of disaster areas and occurrence probabilities.
- We describe our tool to compute and visualize such distributions for any network topology and disaster set.
- We apply our method to a set of Japanese earthquake scenarios, although our methodology is applicable to other types of disasters as well.

2. MODEL

We assume the network $G = \{V, E\}$, consisting of a set V of N nodes connected by a set E of links, is embedded in a plane, and lies completely in a bounded convex region $R \subseteq \mathbb{R}^2$. The network can either be directed or undirected. Nodes $v \in V$ are modeled as points $p \in R$. Instead of modeling them as straight line segments, each link is modeled as a finite sequence of line segments connecting their nodes.

We model earthquakes deterministically, i.e., we assume that all links intersecting a disaster area, which we take as the area(s) in which ground motions exceed a specific level, fail. If a node lies within a disaster area, all of its links must have at least one endpoint in the disaster area and therefore would fail.

Earthquakes typically occur at faults, and thus can not occur everywhere in R. In addition, the ground motion, and thus the disaster area after an earthquake, depends on the earthquake's magnitude, as well as the properties of the rocks and sediments that earthquake waves travel through. Many earthquakes with similar locations affect the same links of the network, even though their exact disaster areas may differ. We therefore argue that it makes sense to take a finite representative set of earthquakes and use it to calculate the network's vulnerability.

We assume that we are given a finite set of possible disasters D. We further assume that exactly one of these disasters will manifest at a time. The probability of multiple (independent) earthquakes occurring simultaneously is generally very small and thus is ignored in this paper. Earthquakes that trigger other disasters (e.g., aftershocks) can still be modeled, by combining their disaster areas. Each disaster

$d \in D$ has a disaster area $A(d) \subseteq \mathbb{R}^2$ and an occurrence probability $P(d)$. Note that $\sum_{d \in D} P(d) = 1$.

We model a disaster area as either a circle, line segment, simple polygon, or a finite union of these. However, our model and methods can be used with any shape of disaster area, as long as it is possible to calculate if a line segment intersects it.

There are multiple ways to obtain the set D. One can generate potential earthquakes in a Monte Carlo approach based on fault parameters. The United States Geological Survey (USGS) provides tools to, given parameters, compute detailed intensity maps, which can be used for this purpose (`usgs.github.io/shakemap/shakemap_archives.html#generating-earthquake-scenarios`). Another approach is to take a historic set of the last N earthquakes above a certain magnitude. Finally, one can use a given set of earthquake scenarios as input. The last two methods have the advantage that ground motion data can be more accurately calculated by incorporating more details (e.g., on ground properties) than the automated tool would.

As an example, in the following section, we will convert Japanese J-SHIS earthquake scenarios to our disaster model.

3. J-SHIS EARTHQUAKE SCENARIOS

Japan has one of the highest earthquake rates in the world and thus needs to be especially prepared for major earthquakes. The National Research Institute for Earth Science and Disaster Resilience (NIED) provides much information about potential earthquakes through the Japan Seismic Hazard Information Station (J-SHIS, `http://www.j-shis.bosai.go.jp/en/`). Of particular interest to us are their Seismic Hazard Map and Scenario Earthquake Shaking Maps.

The Seismic Hazard Map gives probabilities for significant ground motion for all of Japan. These probabilities are calculated in a very similar method as our approach: by aggregating over a set of (representative) modeled earthquakes [3]. Unfortunately, as the end result is an aggregation, and the intermediate results are not publicly available, this map was not usable for our purposes.

Instead, we made use of the Scenario Earthquake Shaking Maps. Of special interest are the earthquakes occurring in major active fault zones, as these are the highly active fault zones that cause earthquakes that have large social and economical impact. These scenario maps contain, among other data, (JMA) seismic intensities for each affected grid in Japan, using Divided Quarter Grid Square Codes [1]. By converting these to geographical coordinates, and only keeping those grids with an intensity above a specific threshold, a disaster area (of a union of rectangles) can be obtained for every single scenario on the dataset. The resulting disaster areas are not contiguous, as there are gaps where the seismic intensity is below the threshold.

The scenarios do not contain occurrence probabilities. To obtain these probabilities, we take the mean recurrence intervals for each fault from the parameter dataset for the Seismic Hazard Map. If a fault segment has N scenarios and mean recurrence interval i, the occurrence probability of all its disasters is taken to be

$$\frac{1}{iNT},$$

where T is the sum of the inverses of all recurrence intervals of fault segments with $N > 0$.

4. VULNERABILITY DISTRIBUTIONS

In 1994, S. C. Liew et al. proposed characterizing network survivability by a survivability function, rather than by a single value (like the expected value after a random disaster) [5]. In essence, their survivability function is the probability mass function of a given survivability measure after a random disaster. Some interesting values can easily be derived from this function, for example the worst-case survivability, r-percentile survivability, or the probability of zero survivability. As far as we know, this concept has not yet been applied to geographically correlated failures. In this section, we propose a method to efficiently compute these distributions.

4.1 Failure States

As an intermediate step towards computing measure distributions, we first consider the probability distribution over the state of the network after a random disaster.

Let a failure state s be defined as a set $s \subseteq E$, where $e \in s$ if and only if e is down.

Let S be the random value indicating the failure state after the disaster and let $S(d)$ be the failure state after disaster $d \in D$. Thus $S(d)$ is the set of all links intersecting the disaster area $A(d)$.

Because we assume exactly one disaster occurs, we have

$$P(S = s) = \sum_{d \in D | S(d) = s} P(d) \tag{1}$$

The distribution over S can now be computed as follows:

1. $\forall d \in D$, compute $S(d)$
2. $\forall s \in S[D]$ (the image of S), store $S^{-1}(s) = \{d \in D | S(d) = s\}$
3. $\forall s \in S$, $P(S = s) = \sum_{d \in S^{-1}(s)} P(d)$

Note that $|S[D]| \leq |D|$ (trivially), and can be significantly smaller when many disasters occur in the same small region. The value of a measure only depends on the state of the network, and thus it only needs to be computed once per possible failure state, instead of once for each $d \in D$. By iterating over possible failure states instead of disasters, we can potentially significantly reduce computation time, when computing the distribution over a measure.

4.2 Measures

Consider a measure M. Let $M(d)$ be the value of the measure after disaster d, and $M(s)$ be the value of the measure in failure state s. Note that $M(d) = M(S(d))$.

Similarly as in equation 1, we have

$$\begin{aligned} P(M = m) &= \sum_{d \in D | M(d) = m} P(d) \\ &= \sum_{s \in S[D] | M(s) = m} \left(\sum_{d \in D | S(d) = s} P(d) \right) \\ &= \sum_{s \in S[D] | M(s) = m} P(S = s) \end{aligned} \tag{2}$$

The distribution over M can now be calculated as follows:

1. $\forall s \in S[D]$, compute $P(S = s)$ as described in section 4.1

2. $\forall s \in S[D]$, compute $M(s)$

3. $\forall m \in M[S[D]]$, store $\{s \in S[D] | M(s) = m\}$

4. $\forall m$, $P(M = m) = \sum_{s \in S[D] | M(s) = m} P(S = s)$

Note that this method can be performed in parallel, to further increase performance.

5. DISASTER IMPACT VISUALIZATION

The disadvantage of computing a distribution instead of a single value is that one may be overwhelmed by the amount of data. Thus it is important to properly visualize the results in a useful fashion.

The distributions over a measure can be clearly visualized with a histogram of the cumulative distribution function (CDF), for example, as in figures 2 and 3.

The intermediate results of the computations in section 4.2, such as the distribution over failure states and the coupling of disasters with their resulting state and measure, can also greatly help in preparing the network against disasters.

To this end, we have created the Disaster Impact Visualization Tool (DIVT). This tool can, given any network topology and disaster set, compute and visualize the vulnerability distribution and many intermediate results. DIVT maps the network on a world map using the NASA World Wind library (`worldwind.arc.nasa.gov`). By drawing disaster areas over the network, users can clearly see which links are affected by a disaster and why.

The measure distribution, state distribution, and the coupling between these distributions and the disasters themselves, are visualized in a tree structure (see figure 1).

At the top level one can see and select the values of the measure with their corresponding probability. Their child nodes show the probabilities of the states resulting in these values. Finally, at the lowest level are the individual disasters causing these states. By selecting one or more of these tree nodes, all corresponding disaster areas are drawn in red on the map. Failing links are colored pink.

An example is given in figure 1. We first expanded all failure states that result in an Average 2-Terminal Reliability – the number of connected node pairs divided by the total amount of node pairs – of 0.87512. Subsequently, we expanded a specific failure state with 4 failed links. This failure state is the result of either disaster scenario "F006104 (case 5)" or "F006104 (case 6)" (names are assigned based on J-SHIS Fault Code and Case Number). "F006104 (case 6)" was selected and is drawn on the map. Some basic statistics, like the expected value, variance, worst case, and all CDF values, are computed and displayed via the Statistics tab.

6. EXPERIMENTAL RESULTS

In this section, we demonstrate the use of our methods on two Japanese network topologies: JGN2plus-Japan and Sinet. Both were downloaded from the Topology Zoo [4]. As these files only contain (broad) geographical coordinates for the nodes, and not the links, all links are assumed to be straight line segments directly connecting their endpoints. Mercator projection was used to map all geographical coordinates to the 2-dimensional plane. Nodes without any geographical information were ignored.

JGN2plus-Japan spans almost all of Japan, but only has 11 nodes and 10 links. In contrast, Sinet spans a slightly smaller area, but consists of 47 nodes and 49 links.

As disasters, we took the J-SHIS earthquake scenarios described in section 3, specifically those from the 2016 dataset. These comprise 655 scenarios for 189 fault segments. The JMA seismic intensity threshold was set to 5.5. We chose the average 2-terminal reliability (ATTR) as our measure.

In figures 2 and 3, the cumulative distribution functions of the ATTR of both networks, after one of the earthquake scenarios, has been plotted. One may notice a difference between the two: while JGN2plus-Japan has a much lower probability of becoming disconnected than Sinet (0.352 and 0.673, respectively), its probability of incurring a large ATTR impact is much higher than for Sinet. $P(ATTR \leq 0.7)$ is 0.224 for JGN2plus-Japan and 0.049 for Sinet.

This is probably caused by the large difference in network size between both networks. As JGN2plus-Japan consists of fewer nodes and links, it has a higher probability that it will not be hit by the earthquake at all. However, in the case that the network does get hit, it lacks the backup paths to keep most of its connections. We can confirm this by inspecting P(No Link Failures) in DIVT. Indeed, the probability of all links of JGN2plus-Japan being unaffected is 0.648, and there are no possible states in which any link fails, but the network stays connected. The comparatively low P(No Link Failures) of Sinet is 0.263.

The worst-case disasters for JGN2plus-Japan all occur around Tokyo, resulting in an ATTR of 0.291 with probability 0.007. The worst-case disasters for Sinet are located around Osaka, and result in an ATTR of 0.362 with probability 0.009. JGN2plus-Japan has an expected ATTR value of 0.866 with a variance of 0.044 and Sinet an expected ATTR value of 0.920 with a variance of 0.016.

For both networks, only computing the ATTR for each possible failure state, instead of for each disaster, had a large effect on performance, reducing the number of times ATTR had to be computed from 655 to 22 and 93 for JGN2plus-Japan and Sinet, respectively.

Acknowledgement

Part of this work has been supported by SURFnet and COST Action CA15127 (RECODIS).

7. REFERENCES

[1] Standard grid square and grid square code used for the statistics (announcement no. 143 by the administrative management agency on july, 12, 1973). `www.stat.go.jp/english/data/mesh/02.htm`. Accessed: 07-05-2017.

[2] D. Bienstock. Some generalized max-flow min-cut problems in the plane. *Mathematics of Operations Research*, 16(2):310–333, 1991.

[3] H. Fujiwara, S. Kawai, N. M. Shin Aoi, S. Senna, N. Kudo, M. Ooi, K. X. Hao, K. Wakamatsu, Y. Ishikawa, T. Okumura, T. Ishii, S. Matsushima, Y. Hayakawa, N. Toyama, and A. Narita. Technical reports on national seismic hazard maps for japan. Technical Report 336, NIED, November 2009.

[4] S. Knight, H. Nguyen, N. Falkner, R. Bowden, and M. Roughan. The internet topology zoo. *Selected Areas*

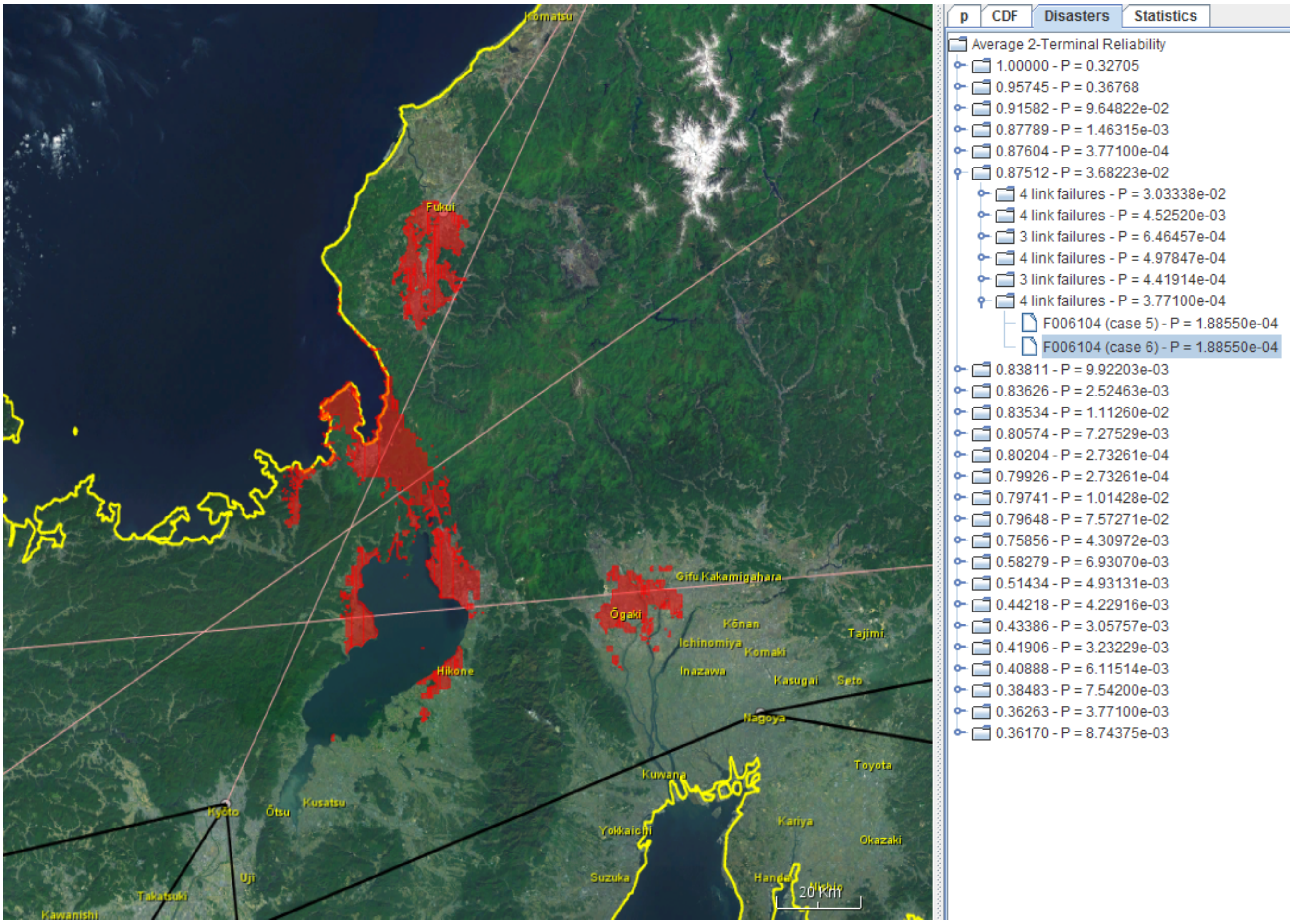

Figure 1: Visualization of distributions and disasters. Red: disaster area, pink: affected links.

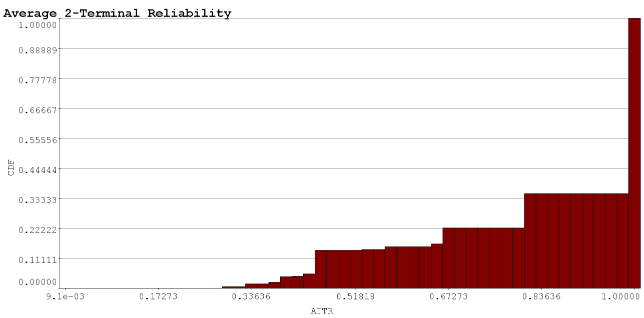

Figure 2: ATTR distribution of JGN2plus-Japan.

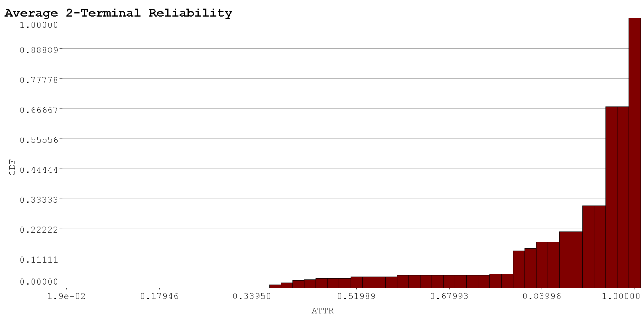

Figure 3: ATTR distribution of Sinet.

in Communications, IEEE Journal on, 29(9):1765 –1775, october 2011.

[5] S. C. Liew and K. W. Lu. A framework for characterizing disaster-based network survivability. *IEEE Journal on Selected Areas in Communications*, 12(1):52–58, 1994.

[6] S. Neumayer, A. Efrat, and E. Modiano. Geographic max-flow and min-cut under a circular disk failure model. *Computer Networks*, 77:117 – 127, 2015.

[7] S. Trajanovski, F. A. Kuipers, A. Ilić, J. Crowcroft, and P. Van Mieghem. Finding critical regions and region-disjoint paths in a network. *IEEE/ACM Transactions on Networking (TON)*, 23(3):908–921, 2015.

[8] X. Wang, X. Jiang, and A. Pattavina. Assessing network vulnerability under probabilistic region failure model. In *High Performance Switching and Routing (HPSR), 2011 IEEE 12th International Conference on*, pages 164–170. IEEE, 2011.