A Tractable Stochastic Model of Correlated Link Failures Caused by Disasters

János Tapolcai*, Balázs Vass*, Zalán Heszberger*, József Bíró*, David Hay†, Fernando A. Kuipers‡, Lajos Rónyai§
*MTA-BME Lendület Future Internet Research Group and MTA-BME Information Systems Research Group
Budapest University of Technology and Economics (BME), {tapolcai, balazs.vass, heszi, biro}@tmit.bme.hu
†School of Engineering and Computer Science, Hebrew University, Jerusalem, Israel, dhay@cs.huji.ac.il
‡Delft University of Technology, The Netherlands, F.A.Kuipers@tudelft.nl
§Computer and Automation Research Institute Hungarian Academy of Sciences and BME, ronyai@sztaki.hu

***Abstract*—In order to evaluate the expected availability of a service, a network administrator should consider all possible failure scenarios under the specific service availability model stipulated in the corresponding service-level agreement. Given the increase in natural disasters and malicious attacks with geographically extensive impact, considering only independent single link failures is often insufficient. In this paper, we build a stochastic model of geographically correlated link failures caused by disasters, in order to estimate the hazards a network may be prone to, and to understand the complex correlation between possible link failures. With such a model, one can quickly extract information, such as the probability of an arbitrary set of links to fail simultaneously, the probability of two nodes to be disconnected, the probability of a path to survive a failure, etc. Furthermore, we introduce a pre-computation process, which enables us to succinctly represent the joint probability distribution of link failures. In particular, we generate, in polynomial time, a quasilinear-sized data structure, with which the joint failure probability of any set of links can be computed efficiently.**

I. Introduction

Being able to guarantee high availability of network services is a crucial part of network management. The required level of service availability is usually explicitly defined in a contract between the service provider and the client, called service-level agreement (SLA). A violation of the agreed-upon service availability may lead to a financial penalty for the network operator, hence, network operators must carefully (under-)estimate the availability of their services and, if necessary, reserve protection resources and implement recovery schemes to meet the availability demands. A typical availability value is "five nines" (99.999%), which translates to an average of at most 5.26 minutes downtime per year. However, a recent taxonomy of Internet failures [1] has revealed that big network outages last much longer, and are often caused by disasters that are beyond the protection schemes deployed, or due to not properly taking into account the co-dependency and hence correlation in tightly-coupled systems. Unfortunately, traditional availability estimation approaches, (wrongly) assume link-failure events to be independent.

Part of this work has been supported by COST Action CA15127 (RECODIS), the Hungarian Scientific Research Fund (grant No. OTKA K124171 and K115288), and the HUJI Cyber Security Center in conjunction with the Israel National Cyber Directorate in the Prime Minister's Office.

The problem of correlated links failures has become more severe in the last decades, due to the increased use of virtual environments, whose physical structure is typically hidden from the user. Nevertheless, networks are built on physical infrastructure and comprise elements such as switches, routers, and optical links, which are prone to physical failures. While some of these failures are isolated, in many cases several nodes and links located in a geographic area fail simultaneously. Geographic failures could for instance be caused by natural disasters, such as earthquakes, hurricanes, or tsunamis [2], [3]. A recent example is the outage due to Cyclone Vardah in India on December 12th, 2016, when Autonomous System AS15169 (operated by Google) and dependent Internet services were severely affected for many hours and slightly degraded for several months more. Such geographically correlated failure events are called *regional failures* and, due to their significant impact, are receiving increased attention [3]–[21]. Unfortunately, in addition to natural disasters, network operators also need to prepare more for destructive human activities, such as terrorist attacks.

A. Related work

Computing availability in the presence of independent single-point failures is a well-investigated topic (cf. [22] and references therein). Also dealing with correlated failures has a long history in the form of Shared Risk Link Groups (SRLG) (e.g., [23]–[27]). An SRLG typically comprises few network components (links or nodes) with considerable risk of failing together. There have been some efforts to attach probability values to an SRLG, called Probabilistic SRLG (PSRLG) [28], [29]. A natural approach is to select a set of disaster scenarios as input [4], e.g. based on historical data. Mostly, it is assumed that the risk groups are given, after which, for example, a pair of risk-disjoint paths needs to be found. There has been some work, e.g. [20], [30], where the risk groups are based on the proximity of links to each other, which may be considered a primitive form of geographically correlated failures.

Much of the work on regional failures has assumed a given disaster shape (often circular disk, or even line segments) and, under that particular model, has addressed specific sub-problems in network planning, like finding the most vulnerable

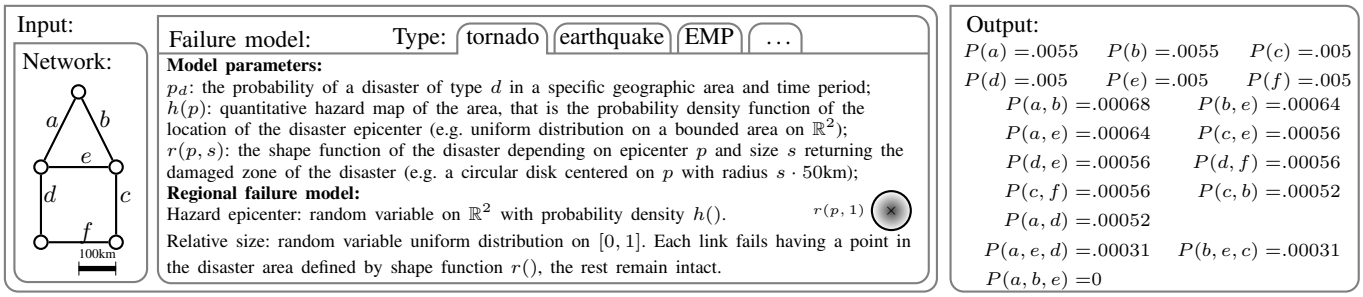


Fig. 1. An illustration of the problem inputs and outputs.

part(s) of the network [5], [6], [8], [12], studying the impact on the network of a randomly placed disaster [16]–[18], designing a network and its services with disaster resiliency in mind [9], [11], [13], [14], and (re)routing of connections to minimize service impact due to a disaster [10], [19]. Some work has considered probabilities, either in the context of a disaster having a certain probability to disconnect a link, e.g. [7], or in the context of only having partial (probabilistic) information on the geographical layout of a network, e.g. [15].

While the above-mentioned papers considered geographically correlated failures, a common property of the targeted sub-problems is to search for the location(s) where a disaster will cause the maximum *expected* damage to the network. In particular, the result is the expected value of a sum of random variables. This is a crude averaging process which is unable to exhibit correlations among many important failure events. The problem of precisely and quickly calculating the correlations between link failures, to conduct a more thorough network vulnerability assessment, has insufficiently been addressed.

B. Availability of correlated links

To evaluate the availability of network services, we need the following two primitives: (i) the probability of a failure of a single element within a given set S and (ii) the probability of simultaneous failure of all elements in a given set S. All the other use-cases can be computed by iteratively calling the above two primitives.

Fig. 1 shows an example network and the corresponding failure probabilities on the right. Suppose we need to establish a high-availability connection from the top node through working path of link b and protection path $a-e$. The unavailability of the working path can be computed as $P(b) = 0.0055$, and for the protection path it is $P(a)+P(e)-P(a,e) = 0.00986$. In the traditional approach, the two paths are assumed to fail independently; thus, the total connection availability is estimated as $1-0.0055 \cdot 0.00986 = 0.999945$, i.e. four nines. However, considering the joint failure probabilities of the links (provided in the example), the total connection availability should be $1 - P(a,b) - P(b,e) + P(a,b,e) = 0.9987$, i.e. not even three nines, which is a significant difference.

Unfortunately, (correlated) network failures are hard to compute and predict. Nonetheless, in order to evaluate the expected availability of a service, a network administrator should consider all possible failure scenarios under the specific service availability model stipulated in the corresponding service-level agreement.

C. Main contributions

The main contributions of this paper are the following:

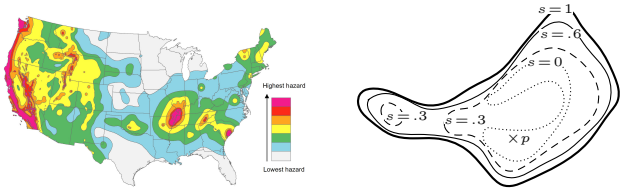
- To our knowledge, this is the first study developing a general stochastic model of disasters in order to explicitly capture the correlations between link failures, as a result of regional failures.
- We devise a pre-computation process to perform the necessary numerical integration off-line. In terms of the network size, there may be exponentially many joint failure events. However, we construct a succinct representation of the joint probability distribution of link failures, which under some practical assumptions has space complexity $O((n+x)\rho^3)$, where n is the number of nodes, x is the number of link crossings (in practice $x \ll n$), and ρ represents a density of the topology, which is independent of the network size.
- We provide a proof-of-concept implementation and simulations to demonstrate how the above-mentioned stochastic model can be efficiently computed, even on commodity computers. This facilitates comprehensive service availability analyses considering disaster failures.

This paper is organized as follows: Sec. II explains the stochastic model we use to represent regional failures. Sec. III proposes an off-line pre-computation process with performance guarantees. Sec. IV demonstrates how the pre-computation and the query of the data structure can be computed efficiently. Sec. V provides a numerical evaluation of the proposed schemes and we conclude in Sec. VI.

II. The Network and Regional Failure Model

To compute the availability of a path composed of a set of network elements (links and nodes) S, we need to compute the probability that any item of a set of network elements fails. The availability of the path is at least $1 - \sum_{e \in S} P(e)$, where $P(e)$ denotes the failure probability of network element e. In case of independent failures or even under light correlation, if $P(e) \ll 1$, this bound gives a good estimate on the availability.

To compute the probability that a set of links (usually forming a cut) fails, we need to answer the question: **what is the**



(a) Hazard map $h(p)$ for earthquakes as function of epicenter p. [32]

(b) Shape of the regional failure $r(p, s)$ for epicenter p and different sizes $s = 0, 0.3, 0.6, 1$.

Fig. 2. Example of real-world inputs.

probability that a set of links S fails simultaneously? Let us denote the above probability by $P(S)$, which typically has a more complex relation with the correlation structure of the link failures: both exclusivity of the events or strong correlation can affect the result substantially. To answer the question, we propose a general stochastic model of the possible network failure events. This, after some pre-computation, will allow us to build a succinct representation of the joint probability distribution of link failures. In our model, the failures are considered to come solely from disasters affecting a bounded geographical area. We focus only on the failed links, where a node failure is manifested as the joint failure of the set of all links adjacent to the node.

While traditional approaches focus on single-point failures, which represent hardware/node failures, cable/link cuts, etc., we adopt a model for regional failures and focus on computing the conditional probability $P_d(S)$ that, in a given time period, a set of links S fail together under a disaster of type d (e.g., a tornado, earthquake, Electromagnetic Pulse (EMP), etc.).

Assumption 1: We assume that, in the investigated time period, there will be at most one disaster.[1]

In such a case, to obtain the availability values, we need to build a model for each disaster type, and the resulting availability of S can be expressed as $1 - \sum_{d \in D} p_d \cdot P_d(S)$, where D denotes the set of failure types modeled and p_d is the probability of disaster d. From now on, for ease of notation, we will consider a fixed failure type d, and therefore, the subscript d is omitted hereafter.

A. Stochastic Modeling of Regional Failures

The network is modeled as an undirected connected geometric graph $G = (V, E)$, with $n = |V|$ nodes and $m = |E|$ links embedded in $\mathbb{R}^2$. The links can be either line segments or polylines built up from adjacent line segments. Note that our algorithms are mostly linear in the network size, thus a link represented by adjacent line segments, can be modeled as a series of 2-dimensional points.

We model regional failures caused by a disaster with the following parameters with randomly chosen values:

epicenter p, which is a point in the plane $\mathbb{R}^2$,
shape (and size) s, which is a real value in $[0, 1]$.

[1]The case when more disasters are expected to happen simultaneously can be handled by defining a new mixed disaster type, see also [31].

Each point $p \in \mathbb{R}^2$ is assigned with a **hazard** $h(p)$ representing the probability that p becomes the epicenter of the next disaster (see Fig. 2a). Specifically, $h(p)$ is a probability density function on the area $\mathbb{R}^2$, and therefore,

$$\int_{p \in \mathbb{R}^2} h(p) \mathrm{d}p = 1 \ . \tag{1}$$

After a regional failure of the examined type (e.g. EMP attack, natural disasters, such as solar flares, earthquakes, hurricanes, and floods) the physical infrastructure (such as optical fibers, amplifiers, routers, and switches) in some area is destroyed. The possible shapes for this area are defined by a set $r(p, s)$ that represents a closed region on the plane (the actual shape of the area in which every communication link is destroyed) as a function of epicenter p and the shape/size parameter s. This is a general disaster model, where several possible damage areas can be defined by $r(p, s)$.

Definition 1: We assume a *regional failure* of epicenter p and shape/size s will result in a failure of every link of network G that has a point in $r(p, s)$. A *disk failure* is a special type of regional failures where $r(p, s)$ are always circular disks.

We assume that $r(p, s)$ is monotone increasing in s (see Fig. 2b for an example)[2], or more formally we assume that

Assumption 2:

$$r(p, s) \subseteq r(p, s') \text{ if } s < s' \quad \forall p \in \mathbb{R}^2, 0 \leq s, s' \leq 1 \ , \tag{2}$$

$r(p, s)$ for a given p is a result of uniform sampling of damage areas. Namely, for a given p the probability of the failure to be of size smaller than s is exactly s. Thus, s is called *relative size* in the remainder of the paper.

It is important to notice that given the disaster epicenter and relative size, the outcome of the attack is deterministic. In other words, any link e within $r(p, s)$ fails with probability 1, if a failure event with parameters p and s occurs. Let us denote the set of failed links by $R(p, s)$. Assumption 1 implies that, given a point p, $R(p, s) \subseteq R(p, s')$ if $s \leq s'$. Let $s(p, e)$ denote the corresponding smallest size s for which a failure at point p can cover link e. Furthermore, we denote by ρ the maximum number of links that can be affected by a single failure (of maximum size $s = 1$):

$$\rho = \max_{p \in \mathbb{R}^2} R(p, 1) \ . \tag{3}$$

B. The Probability of Multiple Failing Links

First, we will explain how to compute the probability that a set of links $S \subseteq E$ fail simultaneously in the next disaster.

Let $f(e, p)$ be the **probability** that link e fails if a disaster with epicenter p happens. Note that $f(e, p) > 0$ can occur iff $e \in R(p, 1)$. $f(e, p)$ can be computed from $R(p, s)$, where s is in the range $[0, 1]$. Hence,

$$f(e, p) = \int_{s=0}^{1} I_{R(p,s)}(e) \mathrm{d}s \ , \tag{4}$$

[2]Various failure shapes were studied so far [3], [5]–[20], mainly in the form of circular regional failures or line-segment failures, but in some cases also for arbitrary geometric objects [7], [8]. All of these models meet Assumption 2. Note that we do not require the regions to be connected, but can be the union of multiple disjoint sets as a part of the same disaster (e.g. $s = .3$ on Fig. 2b).

where the indicator function $I_{R(p,s)}(e)$ indicates whether $e \in R(p,s)$. Thus,

$$I_{R(p,s)}(e) = \begin{cases} 1 & \text{if } e \in R(p,s) \ , \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

If $I_{R(p,s)}(e) = 1$, then $I_{R(p,s')}(e) = 1$, for $s \leq s'$.

We now extend our notation to capture the probability of the failure of link e in the next disaster:

$$P(e) := \int_{p \in \mathbb{R}^2} h(p) f(e,p) \mathrm{d}p. \tag{6}$$

We denote the probability that a set of links $S \subseteq E$ fail simultaneously, given that the disaster epicenter is $p \in \mathbb{R}^2$:

$$f(S,p) := \int_{s=0}^{1} \prod_{e \in S} I_{R(p,s)}(e) \mathrm{d}s \ . \tag{7}$$

In other words, if the sequence of links is $S = (e_1, e_2, \ldots, e_{|S|}) \subseteq R(p,1)$ and $s(p,e_1) \leq s(p,e_2) \leq \cdots \leq s(p,e_{|S|})$, then $\prod_{e \in S} I_{R(p,s)}(e) = 1$ iff $s \geq s(p,e_{|S|})$, otherwise the product is 0. This implies that

$$f(S,p) = f(e_{|S|},p) = \min_{e \in S} f(e,p) \ . \tag{8}$$

Finally, $P(S)$ denotes the probability that all links of a given set S fail simultaneously. Using the above results:

$$P(S) = \int_{p \in \mathbb{R}^2} h(p) f(S,p) \mathrm{d}p = \int_{p \in \mathbb{R}^2} h(p) \min_{e \in S} f(e,p) \mathrm{d}p \ . \tag{9}$$

For example, on the right of Fig. 1, the results of applying the formula to the 5-node network are shown for all the non-zero joint link failure probabilities. In this example, $r(p,s)$ is always a circular disk of radius $s \cdot 50$km. Potentially there are exponentially many joint failure events in terms of the network size; however, links far from each other have zero probability to fail jointly because of a single disaster. This holds, for example, for links f and e, whose smallest distance is 200km.

Former works (e.g., [7, in the proof of Lemma 8]) expressed the joint failure probability of a set S by multiplying the failure probabilities of the links in S, thus implicitly assuming these failures are independent. Unlike [7], our model assumes deterministic failure outcome (once its epicenter and shape are set). This implies that, in our model, failures are dependent. For example, two lines in the same location (e.g., within the same conduit) always fail together.

C. Example of the Geographical Correlation of Failures

In this section, we first consider a simple linear and discrete model for some of the ideas presented so far. We assume that the ground set of our simplified world is the set of 1000 integer points of a line with coordinates between $z_{min} = -499$, $z_{max} = 500$ and we have two links e_0 and e_z, which themselves are integer points from the interval $[-499, 500]$, e_0 is at position 0, and e_z is at position z. Let the probability that i is the location of a disaster be $h_i = 10^{-3}$ for $i = -499, \ldots 500$ so that $\sum_{i=-499}^{500} h_i = 1$. According to Eq. (9), the probability of the failure of link e_0 is

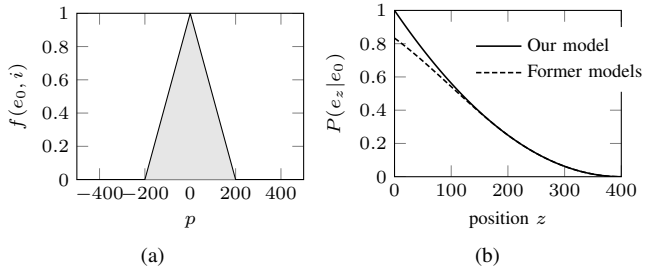


Fig. 3. An example of $f_i(0)$ at different i positions and the corresponding $P(e_z|e_0)$ depending on z. Former models assumed the link failures are independent given an epicenter of the disaster.

$$P(e_0) := \sum_{i=-499}^{500} h_i f(e_0,i) \ , \tag{10}$$

where $f(e_0,i)$ is the conditional probability that link e_0 fails if the failure is at position i. According to Eq. (9), the joint probability of the failure of both links e_0 and e_z is

$$P(\{e_0,e_z\}) := \sum_{i=-499}^{500} h_i \min(f(e_0,i), f(e_z,i)) \ . \tag{11}$$

Let $P(e_z|e_0)$ denote the conditional probability that e_z fails, on the condition that e_0 fails. By definition we have

$$P(e_z|e_0) := \frac{P(\{e_0,e_z\})}{P(e_0)} \ . \tag{12}$$

This is a function of z in our setting. Intuitively, $P(e_z|e_0)$ is close to 1 if the two links are exactly in the same location (i.e. $z = 0$). Besides, $P(e_z|e_0)$ should be a decreasing function of z in the range of $[0, 500]$. See Fig. 3 for an example of $f(e_0,i)$ values and the corresponding $P(e_z|e_0)$.

III. Pre-Computation to Speed up Queries

In the previous section, we have described a model that generates a regional failure according to a hazard density $h(p)$ and a failure shape function $r(p,s)$. Recall that our task is to return $P(S)$ for a set of links $S \subseteq E$, which is the probability that links S fail together in case of disaster d.

Unfortunately, the calculation of integrals (9) can be a computationally intensive process. Thus, our aim is to do some preprocessing in advance, so that when a query comes on the failure probability of an arbitrary set of links S, then the task would be to sum up some of the pre-computed values. The *space complexity* of the proposed data structure is related to the number of pre-computed values. In [20], it was shown that the number of joint link failure events with non-zero probability is $O((n+x)2^{\rho})$, where n is the number of nodes, x is the number of link crossings, and ρ represents a density of the topology (see Eq. (3)). To have a scalable approach, instead of storing every non-negative joint probability, we will introduce a data

structure with $O((n+x)\rho^3)$ items, from which every joint probability can be derived by summing up few items.

In the remainder of this section, we make the following assumptions to be able to apply some computational geometry results.

1) The shapes $r(p,s)$ are limited to circular disks centered at p. This corresponds to the case where the failure of a link e depends on the Euclidean distance $dist(p,e)$ of e to the epicenter of the disaster p. In this case, instead of $r(p,s)$, the input is given by d as a function of s. The maximum radius r is the same for every point, i.e. $r(p,1)$ is a circular disks with radius r and center p for $\forall p \in \mathbb{R}^2$.
2) The relative size s is a *uniformly Lipschitz continuous function* of radius d. That is, there exists a positive number K such that for every point p in the plane, if we have neighborhoods $r(p,s')$ and $r(p,s)$ with respective radii d' and d, then $|s'-s| \leq K|d'-d|$ holds.
3) In our geometric reasoning, we will transform the links of the graph into line segments by slightly shortening them to ensure that no two segments share a common endpoint (see the details of the transformation in Appendix of [33]). We also assume no more than two links intersect in the same point, and no more than two end points lie on the same line.

For ease of presentation, we slightly reduce the domain we are integrating over. Let $\mathcal{P}$ denote the set of points p of the plane such that $dist(p,e) \neq dist(p,e')$ whenever e and e' are different segments from E. We have that $\mathbb{R}^2 \setminus \mathcal{P}$ is of measure zero, hence in our considerations integrating over the plane $\mathbb{R}^2$ can be replaced by integrating over $\mathcal{P}$.

Inspired by (8), now we can define the sequence of possible link failures (see Fig. 4a), when the epicenter of the attack is at p:

Definition 2: The *sequence of link failures* for epicenter $p \in \mathcal{P}$ is an ordered set of links $\mathcal{S}(p) = (e_1, e_2, \ldots, e_l)$, such that $s(p,e_1) < s(p,e_2) < \cdots < s(p,e_l)$, where $l = |R(p,1)|$. Let $\mathcal{S}^j(p)$ denote the first j links of $\mathcal{S}(p)$, i.e. $\mathcal{S}^j(p) = (e_1, e_2, \ldots, e_j)$.

Furthermore, the ordinal number of a set S within $\mathcal{S}(p)$ is defined as follows:

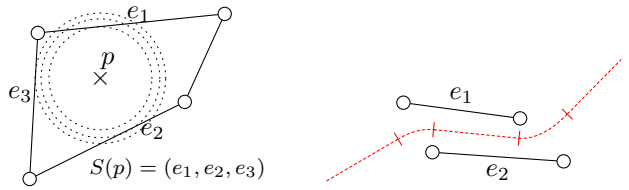
Definition 3:

$$j(S, \mathcal{S}(p)) = \begin{cases} j, & \text{if } S \not\subset \mathcal{S}^{j-1}(p) \text{ and } S \subseteq \mathcal{S}^j(p) \\ 0, & \text{otherwise.} \end{cases}$$

Due to Assumption 2 and using also (9), if there is a disaster at point p, the conditional probability of a set of links $S \subseteq \mathcal{S}(p)$ failing together is

$$f(S,p) = f(\mathcal{S}^{j(S,\mathcal{S}(p))}(p), p) = f(e_{j(S,\mathcal{S}(p))}, p) \ . \quad (13)$$

Finally, we use two practical input parameters, x and ρ, in estimating the space complexity of our approaches. Let x be the number of link crossings in the network G. For backbone networks, x is a small number, as typically a switch is also installed on each link crossing [34]. The second parameter



(a) The sequence of link failures for epicenter p.

(b) The bisector curve of two line segments e_1 and e_2.

Fig. 4. Illustration of definitions in Sec. III

is ρ, the *link density* of the network, which is defined as the maximal number of links that could fail together (i.e. could be covered by a circle of radius r). The link density ρ, practically, should not depend on the network size. Moreover, ρ is at least the maximal nodal degree in the graph.

A. Basic Upper Bound on the Number of Regions

To derive a basic upper bound on the number of regions, we could consider dividing the plane $\mathcal{P}$ into disjoint regions, where every point in each region has the same sequence of link failures. Let $\mathcal{R}_1, \ldots, \mathcal{R}_k$ denote such a set of mutually disjoint regions, where k denotes the total number of possible sequences of link failures with respect to any point p in the plane as epicenter (see Fig. 5a). Let $\mathcal{S}_i$ denote the sequence of link failures corresponding to any point $p \in \mathcal{R}_i$, i.e. $\mathcal{S}(p) \equiv \mathcal{S}_i$, for $i = 1, \ldots, k$.

Based on the observation of (13), it is sufficient to precompute and store the following integrals:

$$P^{i,j} = \int_{p \in \mathcal{R}_i} h(p) f(e_{i,j}, p) \mathrm{d}p \quad i = 1, \ldots, k, \ j = 1, \ldots, |\mathcal{S}_i|, \quad (14)$$

where $e_{i,j}$ denotes the j-th link in $\mathcal{S}_i$.

Finally, since the regions are mutually disjoint as a subset of $\mathcal{P}$ and cover it entirely, equation (9) can be written as a sum and, according to (13), the failure probability of any link set $S \subseteq E$ can be evaluated as

$$P(S) := \sum_{i=1}^{k} \int_{p \in \mathcal{R}_i} h(p) f(S,p) \mathrm{d}p = \sum_{i=1}^{k} P^{i,j(S,\mathcal{S}_i)} \ , \quad (15)$$

where we define $P^{i,0} := 0$ for every $i = 1, \ldots, k$.

Lemma 1: If the $r(p,s)$ are circular disks, then the number of regions k with different failure sequences is $O(m^4)$, where m is the number of links.

Proof: Recall that our line segments are either disjoint or intersect in points that are not endpoints of the respective segments. Let e_1 and e_2 be two such segments. Then $\mathcal{P}$ can be divided into two disjoint (not necessarily connected) domains D_1 and D_2. The points $p \in D_1$ are closer to e_1 than to e_2, and the points $p \in D_2$ are closer to e_2 than to e_1. These domains are bordered by the bisector curve L, which is composed of a finite B number of line segments and parabola arcs. The bisector L of two regions line segments is a simple curve that disconnects the plane into two domains and that can be split into at most seven [35, Lemma 19.2.3] such arcs (see Fig. 4b).

In general, for intersecting segments, L may be the union of two simple curves and we obtain a larger bound B. Note that, if e_1 and e_2 intersect each other, then D_1 and D_2 composed of $2+2$ regions and L is composed of two simple curves intersecting each other.

To complete the proof, we show that the bisectors of the $b = \frac{m(m-1)}{2}$ link pairs divide the set $\mathcal{P}$ into at most $O(m^4)$ regions. In fact, these bisectors together are a union of at most $2b$ simple curves C, where any such simple curve is composed of at most B parts, each part being a segment of a parabola or a line. Note that any pair of such segments in general position intersects an other in at most four points. Moreover, any two such simple curves C intersect each otherin at most $4B^2$ points. Thus, the curves C have at most $2b \cdot 4B^2$ intersection points, dividing up the curves into at most $2b \cdot (4B^2+1) = O(m^2)$ curve segments without division points. The total number of such curve segments is at most $2b \cdot O(m^2) = O(m^4)$. But this gives a bound on the number k of regions as well, since each such curve segment bounds at most two regions and every region has a bounding curve segment. ∎

Next we show that the above bound is tight up to a constant factor. We will define a network by specifying coordinates for the links in $\mathbb{R}^2$. For a positive integer m, the network will have $2m$ links. The first m are vertical line segments, where the i-th line segment connects the point $(2^i, 0)$ to $(2^i, 2^m)$ for $i = 1, \ldots, m$. The bisector curve between the i-th and j-th link is the vertical line whose equation is $x = \frac{2^i+2^j}{2}$. Clearly, x is always an integer, and all the $\binom{m}{2}$ bisector lines correspond to different integers. Overall, they define $\binom{m}{2}+1$ regions, each with different sequence of link failures among the first m links. The second collection of m links is placed in a vertical direction in the same way: $(0, 2^j)$ is connected to $(2^m, 2^j)$, for $j = 1, \ldots, m$. Now the $2m$ segments will define $\left(\binom{m}{2}+1\right)^2 = \Omega(m^4)$ regions with different orders in the sequence of link failures.

Theorem 1: If $r(p, s)$ are closed circular disks, then the space complexity of the basic data structure defined by (14) is $O((n+x)\rho^6)$.

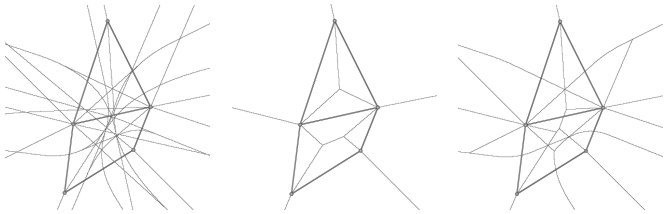
Proof: According to [20, Cor. 4][3], there are at most $O((n+x)\rho)$ SRLGs, such that each SRLG is a set of at most ρ links that can be covered by a circular disk of radius r. Note that every sequence of link failures is a subset of an SRLG.

To complete the proof, for each SRLG S we will count the possible number of sequences of link failures that can be composed from S. Recall that, according to Lemma 1, ρ links may define $O(\rho^4)$ failure sequences. This adds up to $O((n+x)\rho^5)$ for k in (15). Finally, $j(S, \mathcal{S}_i) \leq \rho$, thus each failure sequence consists of at most ρ items in (15).

∎

B. Improved Upper Bound on the Number of Regions

To achieve an improved bound, we will take advantage of recent results of higher-order Voronoi regions. The high-level idea is to merge some of the regions defined by the sequence of link failures. The key observation is that in the integral of (9), for a given set of links S, only the link e with the largest relative distance s matters and the order of the links $S \setminus e$ is irrelevant; thus, the corresponding regions can be merged. To do so, we will generalize the higher-order Voronoi regions, and introduce the $(k, 1)$-Voronoi diagram for line segments, where each region (a.k.a. cell) has the same set of k nearest neighbor line segments and the same line segment as the $(k+1)^{\text{th}}$ nearest neighbor (see Fig. 5b and 5c).[4]



(a) Regions with same sequence of link failures. (b) Nearest Neighbor Voronoi Diagram (c) $(1,1)$-Voronoi diagram

Fig. 5. Different partitions of the plane into regions in the pre-computation process.

Let $H \subset E$ be a set of line segments and $e \in E \setminus H$ be a line segment. The *Voronoi region* $\mathcal{R}(H, e)$ of the pair (H, e) is the set of points $p \in \mathcal{P}$ for which the $|H|$ closest segments from E are exactly the elements of H (in an arbitrary order) and the $|H|+1$-th is exactly e. We shall apply a result from [36]. For this reason we assume that $x = o(n^2)$.

Lemma 2: The number of nonempty Voronoi regions $\mathcal{R}(H, e)$, with $|H| = k-1$, is at most $O(k^2(n-k)+kx))$.

Proof: $\mathcal{R}(H, e)$ is contained in exactly one order-k Voronoi region [36], the one belonging to the set $Y = H \cup \{e\}$. Conversely, for any set of links Y with $|Y| = k$ the order-k Voronoi cell of Y contains at most k nonempty sets of the form $\mathcal{R}(H, e)$. These are the sets $\mathcal{R}(Y \setminus \{e\}, e)$ with $e \in Y$. We know from [36, Thm. 5] that the number of k-Voronoi cells is $O(k(n-k)+x)$. ∎

Next we pre-compute the integrals

$$P^{(H,e)} := \int_{p \in \mathcal{R}(H,e)} h(p) f(e,p) \mathrm{d}p \quad \forall \mathcal{R}(H, e) \ , \qquad (16)$$

for each of the $O(k^2(n-k)+kx)$ Voronoi regions for $k = 1, \ldots, \rho$, where (H, e) ranges over the pairs, such that H is a subset of E, $e \in E \setminus H$ and $|H|+1 \leq \rho$, and $\mathcal{R}(H, e)$ is nonempty.

Let $S \subseteq E$ be a set of links, then

$$P(S) = \sum P^{(H,e)} \ , \qquad (17)$$

where summation is for the pairs (H, e) with $|H| < \rho$, $S \subseteq H \cup \{e\}$, $e \in S \setminus H$. Note that for two such pairs (H, e) and (H', e') the corresponding regions $\mathcal{R}(H, e)$ and $\mathcal{R}(H', e')$ are disjoint: If p is in both of them, then the farthest segment of S from p is e and e', hence $e = e'$. If e is the k-th farthest

[3]Note that, ρ is denoted by σ_r in [20]. Here we use the assumption that the radius of $r(p, 1)$ is equal to r for every point p in the plane.

[4]Note that the $(0, 1)$-Voronoi diagram is equivalent to the nearest neighbor Voronoi diagram.

segment of E from p, then the $k-1$ preceding segments form $H = H'$. Moreover, any point $p \in \mathcal{P}$, for which $f(e,p) \neq 0$ and e is the farthest segment of S from p, does belong to one of the Voronoi regions $\mathcal{R}(H,e)$ listed on the right hand side. Indeed, let H be the set of segments $e' \in E$ closer to p than e. On a region $\mathcal{R}(H,e)$, as above, we have by (9)

$$\int\limits_{p\in\mathcal{R}(H,e)} h(p)f(S,p)\mathrm{d}p = \int\limits_{p\in\mathcal{R}(H,e)} h(p)f(e,p)\mathrm{d}p \ .$$

Theorem 2: If $r(p,s)$ are circular disks, the space complexity of the improved data structure is $O((n+x)\rho^3)$.

Proof: The data structure stores $P^{(H,e)}$ for every possible (H,e) pair that could appear in the formula (17). Note that in the formula (17) the value of $|H|$ is $0,\ldots,\rho-1$. Finally, Lemma 2 gives an upper bound on the number of Voronoi regions $\mathcal{R}(H,e)$ as $O(k^2(n-k)+kx)) \leq O((n+x)k^2)$ with $k = 1,\ldots,\rho$. ∎

IV. Implementation Issues

In Section III, we provided approaches with performance guarantees under the assumption that the shape of the regional failure is always a circular disk. In this section, we suggest implementing another approach with the following features: (1) it can accommodate any shape for the disasters, (2) it is easy to implement as it does not require any advanced geometric algorithms, while the time of pre-computation is rarely a concern, and (3) it nicely processes digital inputs as it uses discrete functions instead of continuous ones.

We discretize the problem by defining a sufficiently fine resolution, say 1 km, and place a grid of 1 km × 1 km squares over the plane to assume that the disaster regions $r(p,s)$ and hit link sets $R(p,s)$ are "almost identical"[5] for every p inside each grid cell c. This way the whole integration problem boils down to a summation. We will define the inputs over the grid, and consider $\mathbb{R}^2$ as a Cartesian coordinate system. We will define $r(p,s)$ over the Cartesian coordinate system, so that for each c we will define an s value for the neighboring c. Let r denote the absolute maximum range of a disaster in km. Let (x_{min}, y_{min}) be the bottom left corner and (x_{max}, y_{max}) the top right corner of a rectangular area in which the network lies. It is sufficient to process each c in the rectangle of bottom left corner $(x_{min}-r, y_{min}-r)$ and top right corner $(x_{max}+r, y_{max}+r)$, and we denote by $c_{i,j}$ the grid cell in the i-th column and j-th row. In this range, for each $c_{i,j}$, we will consider the probability $h_{i,j}$ of the next disaster having epicenter p in the cell $c_{i,j}$, i.e. $h_{i,j} = \int_{p\in c_{i,j}} h(p)\mathrm{d}p$.

For each c, we will compute the sequence of link failures and store the link sets and corresponding s values as follows. We implemented three types of data structures to store the failure probability of link sets. The first and second types are described in Subsec. III-A and III-B, and their length can be upper bounded in case of failures corresponding to Sec. III, let us call them *Basic* and *Improved*, respectively. While both of them have similar computation time and a size linear in $(n+x)$, Improved appears to use significantly smaller space than Basic (see Fig. 7), thus we omit discussing the implementation issues of the Basic structure. To return the failure probability of any link set, in case of both structures, one must go through this list and sum up the corresponding elements.

The query time of sets can be reduced to a constant with very high probability (with the help of hashing) if the data structure stores every possibly failing link set, which is the third data structure we have, called *Complete*. Using self-balancing binary trees, its worst-case query time is always $O(\rho\log((n+x)\rho))$, which is still very impressive. The drawback of structure Complete is that it has an $\Omega(2^\rho)$ space complexity, which makes it very inefficient for bigger network densities.

1) Improved structure: We use an associative array $\mathcal{I}$, which can be addressed by a pair of an (unordered) set of links H and a link e, with $e \notin H$, and returns a series of probability values p, one for each index entry. In the pre-computation process, we build up $\mathcal{I}$ as described in Algorithm 1. We take every little square $c_{i,j}$ and determine the sequence of link failures belonging to epicenters $p \in c_{i,j}$ (see Def. 2), denoted by $\mathcal{S}_{i,j} = (e_1,\ldots,e_k)$, where $k = |\mathcal{S}_{i,j}| \leq \rho$. We take the first l links of $\mathcal{S}_{i,j}$ for $l = 1,\ldots,k$ and check if the pair of link set $\{e_1,\ldots,e_{l-1}\}$ and e_l is in $\mathcal{I}$ (note that for $l=1$ the link set is an empty set). If it is not, we add it to $\mathcal{I}$ with probability value $h_{i,j}\cdot f(e_l, c_{i,j})$; otherwise, we add the previous probability value to its stored version (i.e. to $\mathcal{I}\big[\big(\{e_1,\ldots,e_{l-1}\},e_l\big)\big]$).

Algorithm 1: Building up associative array $\mathcal{I}$

```
for i = x_min−r, ..., x_max+r, j = y_min−r, ..., y_max+r do
    Determine S_{i,j} = (e_1, e_2, ..., e_k).
    for l=1,..., k do
        I[({e_1, ..., e_{l-1}}, e_l)] += h_{i,j} · f(e_l, c_{i,j})
```

Note that for circular disk failures Thm. 2 provides an upper bound of $O((n+x)\rho^3)$ on the number of items in $\mathcal{I}$.

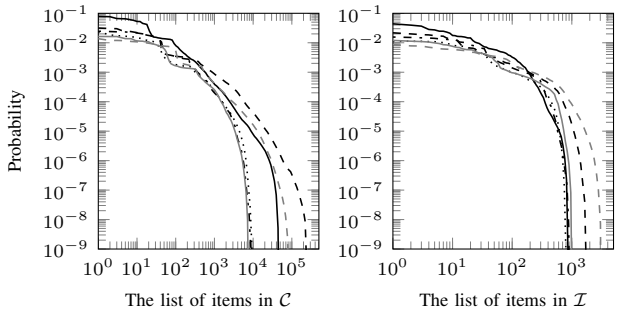
To compute the probability that a set of links S fail simultaneously, we need to examine every item in the associative array as described in Algorithm 2.

Algorithm 2: Computing $P(S)$ with the Improved structure

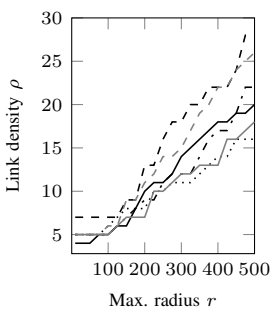
```
p = 0 ;
forall [({e_1, ..., e_{l-1}}, e_l) → p_i] ∈ I do
    if S ⊆ {e_1, e_2, ..., e_{l-1}} ∪ e_l and e_l ∈ S then
        p = p + p_i;
return p
```

2) Complete structure: We use an associative array $\mathcal{C}$, which can be addressed by an (unordered) set of links $\{e_1,e_2,\ldots,e_k\}$ and returns its joint probability value. In this case, in the pre-computation process, we have to extract the contribution of $c_{i,j}$ to the failure probability of every subset S of links. We do this by working with the list $\mathcal{S}_{i,j} = (e_1,e_2,\ldots,e_k)$, and increment the $\mathcal{C}$ values accordingly, i.e. $\mathcal{C}[\{e_1\}] += h_{i,j}\cdot f(e_1,c_{i,j})$, $\mathcal{C}[\{e_2\}] += h_{i,j}\cdot f(e_2,c_{i,j})$,

[5] In particular, we may assume that $f(e,p)$ is independent of p as long as it is in c. We denote this common value by $f(e,c)$.



(a) Complete structure, $r = 200$ (b) Improved structure, $r = 200$



(c) Link density ρ

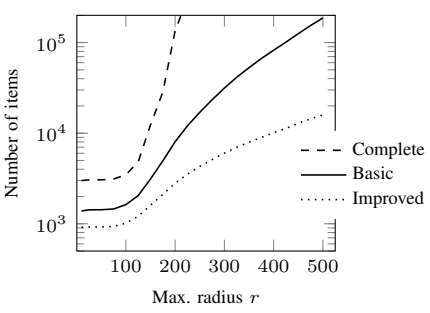
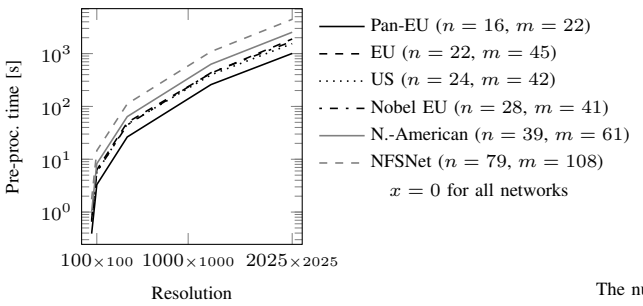
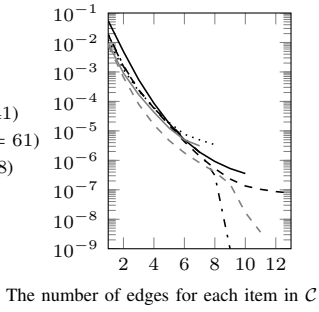


Fig. 7. Average space complexity



(d) The running time, $r = 200$



(e) The joint failure probability $P(S)$ vs. $|S|$

Fig. 6. The space and time complexity of the data structures for the examined network topologies.

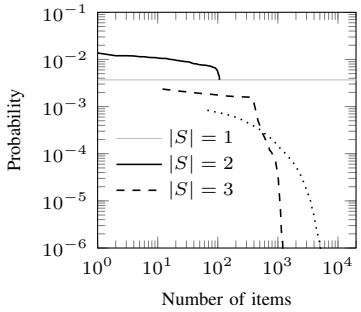


Fig. 8. The probabilities of the items in $\mathcal{C}$ for NFSNet and $r = 200$ km.

$\mathcal{C}[\{e_1, e_2\}] += h_{i,j} \cdot f(e_2, c_{i,j})$, etc. For $P(S)$ we need to look up S in $\mathcal{C}$. If not found, then $P(S) = 0$.

V. Simulation Results

In this section, we present numerical results that validate our model and demonstrate the use of the proposed algorithms on some real backbone networks. The algorithms were implemented in Python 2.7.12., using its various libraries[6]. Runtimes were measured on a commodity laptop with core i5 CPU at 2.3 GHz with 8 GiB of RAM.

All three structures were generated for networks from [33] for maximal radii $r = 10, 25, 50, 75, \ldots, 500$ km. The hazard map $h(p)$ is a uniform distribution in the bounding rectangle described in Sec. IV. The shape $r(p, s)$ is a circular disk with centre p and radius $r \cdot s$. I.e., we select circular disk failures having a range from a uniform distribution over $[0, r]$.

Since the running time of the pre-computation is dominated by the construction of sequence of link failures for all c, there is no significant difference between the construction time of Basic and Improved structures for the same environment. See the average and maximum time in Fig. 6d with respect to the resolution. The solution quality very moderately improves for resolutions higher than 1000×1000.

For smaller network densities, the running time of Complete is similar too, but at a given point ($\rho \simeq 14$) the exponential factor 2^ρ at the space (and therefore time) requirement requirement starts to kick in. See also Fig. 6c for how ρ increases with r. In other words, due to its very short query time, Complete outperforms the other structures for small maximum radii (< 150km in case of the studied networks), but becomes unacceptably large and slow to compute at larger maximum radii (> 200km). Improved is more compact than Basic, and in the area of larger radii this difference is significant (see Fig. 7). We can conclude that, in backbone networks, Complete is the best choice for disasters ranging to at most 200km radius of the destroyed area, but if $r > 200$km, we should stick to Improved.

Fig. 6a and 6b show a distribution of the probabilities of stored items in the associative arrays. For Complete, it is the distribution of the failure probabilities of the set of links. Here $f(x)$ equals to the failure probability of the x^{th} most probable link set. We can witness the emergence of a power-law distribution with exponential cutoff. In practice, the service availability is computed with some precision, and the link sets with very small probability to fail are not stored and used in the evaluation. For example, if we ignore the link sets with probability less than 10^{-4}, it is sufficient to deal with at most 5000 items. Roughly speaking, the power-law distribution means, if we want to increase the precision with one more nines, we need to quadruple the size. The same idea can be used for Improved (and Basic), see Fig. 6b.

We have also investigated the average probability of a set of links with given cardinality. Fig. 6e shows the average failure probability with respect to the number of links failing together. Single links have an average probability of 0.015 to fail, the

[6]The simulation data can be dowloaded from [33].

listed double links 0.0014, the triple 0.00022, which meets our expectation that the correlation between link failures is significant. Fig. 8 further investigates the dependency between the failure probability of a set of links and the set cardinality. We grouped the elements S of Complete by their size $|S|$: there are 108 single link failures in NSFNet whose failure probabilities range between $[0.0037, 0.014]$, there are 1245 dual link failures with non-zero probabilities between $[10^{-6}, 0.003]$, there are 6189 triple link failures with non-zero probabilities between $[10^{-6}, 0.0014]$.

VI. Conclusion

In this paper, we have proposed a general stochastic model of regional failures of a physical network. In particular, we have evaluated the joint failure probability of a set of links. The evaluation is composed of the pre-computation and query phases. The pre-computation is performed off-line during the network planning, which requires to compute numerical integrals using hazard maps and information about the network equipment. As a result of pre-computation, all the probabilities of link sets with positive joint failure probability are stored, if feasible (for disaster ranges $<\sim 150$km in our experience); or else a space-efficient data structure is formed that enables to quickly compute the joint failure probability of an arbitrary set of links. We have proved that the latter data structure stores $O((n + x)\rho^3)$ items, if the failure of a link depends only on the distance to the epicenter of the disaster, where n is the number of nodes, x is the number of link crossings (in practice $x \ll n$), and ρ is the maximal number of links subject to a disaster failure. Our approach facilitates a comprehensive service availability analysis, and can be used to answer related questions, such as where to place VMs in order to guarantee a certain SLA.

References

[1] C. Doerr and F. Kuipers, "All quiet on the internet front?" *IEEE Commun. Mag.*, vol. 52, no. 10, pp. 46–51, 2014.

[2] Y. Nemoto and K. Hamaguchi, "Resilient ICT research based on lessons learned from the Great East Japan Earthquake," *IEEE Commun. Mag.*, vol. 52, no. 3, pp. 38–43, 2014.

[3] J. Heidemann, L. Quan, and Y. Pradkin, *A preliminary analysis of network outages during Hurricane Sandy*. University of Southern California, Information Sciences Institute, 2012.

[4] J. Oostenbrink and F. Kuipers, "Computing the impact of disasters on networks," *ACM SIGMETRICS Performance Evaluation Review*, vol. 45, no. 2, pp. 107–110, 2017.

[5] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, "Assessing the vulnerability of the fiber infrastructure to disasters," *IEEE/ACM Trans. Networking*, vol. 19, no. 6, pp. 1610–1623, 2011.

[6] M. T. Gardner and C. Beard, "Evaluating geographic vulnerabilities in networks," in *IEEE Int. Communications Quality and Reliability Workshop (CQR)*, 2011, pp. 1–6.

[7] P. K. Agarwal, A. Efrat, S. K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, "The resilience of WDM networks to probabilistic geographical failures," *IEEE/ACM Trans. Networking*, vol. 21, no. 5, pp. 1525–1538, 2013.

[8] 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 Trans. Networking*, vol. 23, no. 3, pp. 908–921, 2015.

[9] M. F. Habib, M. Tornatore, M. De Leenheer, F. Dikbiyik, and B. Mukherjee, "Design of disaster-resilient optical datacenter networks," *IEEE/OSA J. Lightwave Technol.*, vol. 30, no. 16, pp. 2563–2573, 2012.

[10] F. Dikbiyik, M. Tornatore, and B. Mukherjee, "Minimizing the risk from disaster failures in optical backbone networks," *IEEE/OSA J. Lightwave Technol.*, vol. 32, no. 18, pp. 3175–3183, 2014.

[11] I. B. B. Harter, D. Schupke, M. Hoffmann, G. Carle *et al.*, "Network virtualization for disaster resilience of cloud services," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 88–95, 2014.

[12] X. Long, D. Tipper, and T. Gomes, "Measuring the survivability of networks to geographic correlated failures," *Optical Switching and Networking*, vol. 14, pp. 117–133, 2014.

[13] B. Mukherjee, M. Habib, and F. Dikbiyik, "Network adaptability from disaster disruptions and cascading failures," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 230–238, 2014.

[14] R. Souza Couto, S. Secci, M. Mitre Campista, K. Costa, and L. Maciel, "Network design requirements for disaster resilience in IaaS clouds," *IEEE Commun. Mag.*, vol. 52, no. 10, pp. 52–58, 2014.

[15] O. Gold and R. Cohen, "Coping with physical attacks on random network structures," in *IEEE ICC*, 2014, pp. 1166–1172.

[16] X. Wang, X. Jiang, A. Pattavina, and S. Lu, "Assessing physical network vulnerability under random line-segment failure model," in *IEEE High Performance Switching and Routing (HPSR)*, 2012, pp. 121–126.

[17] H. Saito, "Analysis of geometric disaster evaluation model for physical networks," *IEEE/ACM Trans. Networking*, 23(6), pp. 1777–1789, 2015.

[18] ——, "Spatial design of physical network robust against earthquakes," *IEEE/OSA J. Lightwave Technol.*, vol. 33, no. 2, pp. 443–458, 2015.

[19] F. Iqbal and F. Kuipers, "Spatiotemporal risk-averse routing," in *IEEE INFOCOM Workshop on Cross-Layer Cyber Physical Systems Security (CPSS)*, 2016.

[20] J. Tapolcai, L. Rónyai, B. Vass, and L. Gyimóthi, "List of shared risk link groups representing regional failures with limited size," in *IEEE INFOCOM*, Atlanta, USA, May 2017.

[21] T. Gomes, J. Tapolcai *et al.*, "A survey of strategies for communication networks to protect against large-scale natural disasters," in *Int. Workshop on Reliable Networks Design and Modeling (RNDM)*, 2016.

[22] S. Verbrugge, D. Colle, P. Demeester, R. Huelsermann, and M. Jaeger, "General availability model for multilayer transport networks," in *Proc. DRCN*, Lacco Ameno, Italy, Oct. 16-19, 2005.

[23] D. Zhou and S. Subramaniam, "Survivability in optical networks," *IEEE network*, vol. 14, no. 6, pp. 16–23, 2000.

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

[25] C. S. Ou and B. Mukherjee, *Survivable Optical WDM Networks*. Springer Science & Business Media, 2005.

[26] A. Somani, *Survivability and traffic grooming in WDM optical networks*. Cambridge University Press, 2006.

[27] S. Yang, S. Trajanovski, and F. Kuipers, "Availability-based path selection and network vulnerability assessment," *Wiley Networks*, vol. 66, no. 4, pp. 306–319, 2015.

[28] H.-W. Lee, E. Modiano, and K. Lee, "Diverse routing in networks with probabilistic failures," *IEEE/ACM Trans. Networking*, vol. 18, no. 6, pp. 1895–1907, 2010.

[29] J. Liu, J. Zhang *et al.*, "Differentiated quality-of-protection provisioning with probabilistic SRLG in flexi-grid optical networks," in *OSA Asia Communications and Photonics Conference*, 2013, pp. AF2G–8.

[30] F. Iqbal, S. Trajanovski, and F. Kuipers, "Detection of spatially-close fiber segments in optical networks," in *Proc. DRCN*, 2016, pp. 95–102.

[31] M. Rahnamay-Naeini, J. E. Pezoa, G. Azar, N. Ghani, and M. M. Hayat, "Modeling stochastic correlated failures and their effects on network reliability," in *IEEE Int. Conf. on Comp. Comm. and Networks (ICCCN)*, 2011, pp. 1–6.

[32] US National Seismic Hazard Maps. [Online]. Available: https://earthquake.usgs.gov/hazards/hazmaps/conterminous/

[33] Network library. [Online]. Available: https://github.com/jtapolcai/regional-srlg/tree/master/psrl

[34] D. Eppstein, M. T. Goodrich, and D. Strash, "Linear-time algorithms for geometric graphs with sublinearly many edge crossings," *SIAM Journal on Computing*, vol. 39, no. 8, pp. 3814–3829, 2010.

[35] J.-D. Boissonnat and M. Yvinec, *Algorithmic geometry*. Cambridge university press, 1998.

[36] E. Papadopoulou and M. Zavershynskyi, "The higher-order Voronoi diagram of line segments," *Algorithmica*, 74(1), pp. 415–439, 2016.