Network Resiliency Against Earthquakes

Alessandro Valentini*, Balázs Vass†, Jorik Oostenbrink‡, Levente Csák†,
Fernando Kuipers‡, Bruno Pace*, David Hay§, János Tapolcai†
*DiSPUTer Department, University of Study "G.d'Annunzio" of Chieti-Pescara, Chieti, Italy,
{alessandro.valentini,bruno.pace}@unich.it
†MTA-BME Future Internet Research Group, High-Speed Networks Laboratory (HSN*Lab*),
Budapest University of Technology and Economics, {vb,tapolcai}@tmit.bme.hu, csaklevente@gmail.com
‡Delft University of Technology, Delft, the Netherlands, {J.Oostenbrink, F.A.Kuipers}@tudelft.nl
§School of Engineering and Computer Science, Hebrew University, Jerusalem, Israel, dhay@cs.huji.ac.il

***Abstract*—Guaranteeing a high availability of network services is a crucial part of network management. In this study, we show how to compute the availability of network services under earthquakes, by using empirical data. We take a multi-disciplinary approach and create an earthquake model based on seismological research and historical data. We then show how to integrate this empirical disaster model into existing network resiliency models to obtain the vulnerability and availability of a network under earthquakes.**

While previous studies have applied their models to ground shaking hazard models or earthquake scenarios, we compute (1) earthquake activity rates and (2) a relation between magnitude and disaster area, and use both as input data for our modeling. This approach is more in line with existing network resiliency models: it provides better information on the correlation between link failures than ground shaking hazard models and a more comprehensive view than a fixed set of scenarios.

I. Introduction

Service-level agreements between backbone network operators and clients define a minimum required availability of network services. A violation of this 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 taxonomy of Internet failures [1] has revealed that many core network outages last much longer, and are often caused by disasters that are beyond the protection schemes deployed, or due to not correctly taking into account the co-dependency and hence correlation in tightly-coupled systems. As traditional availability estimation approaches (wrongly) assume link-failure events to be independent, they overestimate the availability of network services. This can potentially cause severe penalty costs for network operators and significant degradation of the quality of experience to their customers.

Part of this work has been supported by COST Action CA15127 (RECODIS), the Hungarian Scientific Research Fund (grant No. OTKA K124171 and K128062), by the BME-Artificial Intelligence FIKP grant of EMMI (BME FIKP-MI/SC), and the HUJI Cyber Security Center in conjunction with the Israel National Cyber Directorate in the Prime Minister's Office.

One of the primary sources of link-failure correlation is the physical structure of the network; network components located in the same geographical area may fail simultaneously. Geographic failures could, for instance, be caused by natural disasters, such as earthquakes, hurricanes, or tsunamis [2], [3]. Such geographically correlated failure events are called *regional failures* and, due to their significant impact, are receiving increased attention [3]–[13].

This paper deals with *earthquakes*—one of the most common natural disasters to cause significant network outages. We take a multi-disciplinary approach, combining insights from both seismology and network resiliency to provide accurate assessments of the resiliency of communication networks to earthquakes in the studied regions. This is done in two steps: First, we create an earthquake probability model based on insights from past earthquakes (i.e. historical data). Then, we integrate this empirical disaster model into existing network resiliency models to obtain the vulnerability and availability of a network under earthquakes. Integrating empirical disaster models is an important step in determining the resilience of networks to natural disasters, but is often ignored in the existing network resiliency literature. We have applied our model to historical data from Italy, to indicate the vulnerabilities of an Italian communication network to earthquakes.

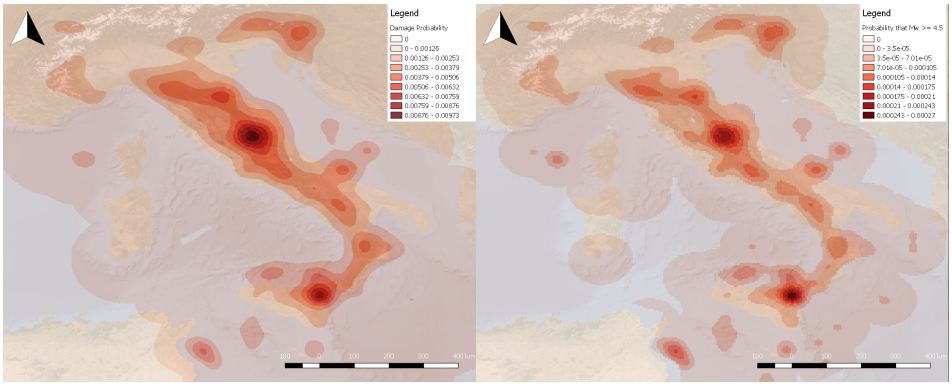
While some previous studies have applied their models on empirical earthquake data, these studies either made use of the ground shaking hazard model (see Fig. 1a) or a set of potential earthquake scenarios.

Ground shaking hazard models give the probability that, for a given site and within a time window, a given intensity level is exceeded. These models are created by aggregating over every potential future earthquake. However, by doing so for the analysis of network vulnerability, we lose information on the correlation between ground motions. Namely, we cannot distinguish between a situation where a single potential earthquake affects two cells simultaneously and a situation when each cell is affected by a different potential earthquake, and hence, not simultaneously. This makes ground shaking hazard models unsuitable for the analysis of network vulnerability, where the correlation between component failures is a crucial issue.

Considering specific earthquake scenarios, on the other

978-1-7281-4698-0/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: Hebrew University. Downloaded on June 02,2023 at 13:19:00 UTC from IEEE Xplore. Restrictions apply.



(a) Ground shaking hazard map. (b) Earthquake activity rates.

Fig. 1. An example of a ground-shaking hazard map (where the earthquake rates are converted in, each cell of the grid, the probability that the intensity level equal to VI is exceeded during the next earthquake of $M_w \geq 4.5$) and earthquake activity rates (which are converted into probabilities that the next earthquake, which has the epicenter in a given cell of the grid, has a magnitude $M_w \geq 4.5$). Intuitively speaking (a) is a blurred version of (b).

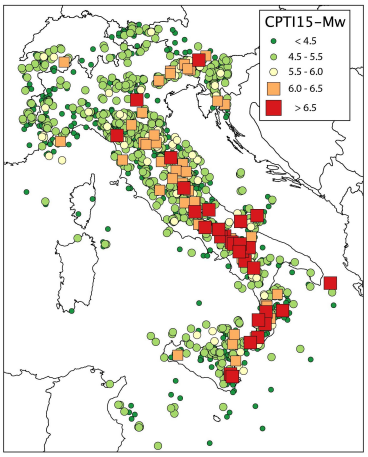


Fig. 2. Historical earthquakes from the most recent version of the historical parametric Italian catalogue [14]. M_w = moment magnitude.

hand, provides a detailed insight into the effects of these specific potential earthquakes (e.g., one can quickly determine which network components are likely to fail together). However, as the number of earthquake scenarios under consideration is limited, and no information is provided for other scenarios, this approach lacks the comprehensive view of probabilistic seismic hazard models that takes into account all potential earthquakes.

In this paper, we take a different approach by computing *earthquake activity rates* (Fig. 1b). Like in ground shaking hazard models, our novel data structure considers grid cells. However, for each grid cell C, and each moment magnitude range $(M_w - 0.1, M_w]$, the earthquake activity rate gives the rate, and so the occurrence probability, of earthquakes with magnitude within this range and epicenter in C. We consider all relevant magnitude ranges, where M_w is $4.6, 4.7, 4.8, \ldots, 8.1$, thus our collection of earthquake activity rates provides a comprehensive view of all earthquakes that are strong enough to damage communication equipment.

Given the magnitude of an earthquake, we attach to its earthquake distribution model a damaged zone radius $R(M_w)$, which is based on the observed structural damage after an earthquake of magnitude in $(M_w - 0.1, M_w]$. The collection of earthquake activity rates (along with the corresponding damaged zone radii), are then plugged into a network resiliency model. The models we use assume that all links that intersect a damaged zone (namely, within radius $R(M_w)$ of the epicenter of an earthquake of magnitude M_w) fail, and calculate the sets of links that can fail simultaneously (and with these sets, the connectivity and availability of the network and its services).

Our main contributions are as follows:

- We argue that commonly used probabilistic ground-shaking hazard models are not suitable for assessing network resiliency to earthquakes and propose to use earthquake activity rates instead. Sections III and V describe how to calculate these earthquake activity rates, in particular given the data available for Italy (namely, a catalog of earthquakes in the last 1000 years and intensity prediction equations, correlating the damage to the distance from the epicenter, which were based on observations within Italy).
- We show, in Section IV, how to integrate earthquake activity rates into two existing network disaster resiliency models: (1) the stochastic geographically correlated link failures model in [5], and (2) the disaster set and failure states model from [4]. Although these models take a different approach towards computing the availability of network services under natural disasters, both approaches give exactly the same results when applied to our empirical earthquake model.
- We apply these methods to generate the earthquake activity rates of Italy and investigate the network resiliency of an Italian network under the different models (Section V).
- We have created an open repository containing our input and output datasets[1].

To the best of our knowledge, we are the first to integrate seismological empirical data on earthquake activity rates into network resiliency models. We believe this multi-disciplinary approach is an important step towards assessing the *real* resiliency of communication networks to natural disasters, thus, based on our outputs, problems like resilient routing or resilient network design can be handled more efficiently.

II. Related work

Although considered as early as 1991 by D. Bienstock [15], research on geographically correlated failures, regional failures, or disasters has only really taken off in the last decade. A typical assumption is that the disaster area takes a fixed

[1]https://github.com/jtapolcai/regional-srlg/tree/master/earthquake

Authorized licensed use limited to: Hebrew University. Downloaded on June 02,2023 at 13:19:00 UTC from IEEE Xplore. Restrictions apply.

shape, such as a disk with fixed radius [16]. By making this assumption, we can compute the amount of disasters required to disconnect two nodes [17], the most vulnerable spot(s) of the network against this type of disaster [8], or the impact after a randomly placed disaster [18]. More often than not, the disaster is also assumed to be able to occur anywhere around the network with uniform probability and link failures are happening independently of each other in the disaster zone. In reality, disasters such as earthquakes are not spread out uniformly over the whole network area, and link failures are correlated in the disaster zone.

Few papers on the resilience of networks against disasters apply their model to empirical earthquake data. L. Ma et al. considered the problem of optimally placing data center networks and content [12]. In their experiments, L. Ma et al. set the parameters based on the peak ground acceleration with an exceedance probability of 2% in 50 years of the USGS 2014 National Seismic Hazard Map[2]. The occurrence probabilities were set randomly.

P. N. Tran and H. Saito proposed algorithms for minimizing the sum of all end-to-end disconnection probabilities under a set of earthquake scenarios by adding new links [10], and by finding geographical routes for network links [11]. They applied their model to 5 potential future earthquake scenarios from J-SHIS[3].

J. Oostenbrink and F. Kuipers proposed an efficient method to compute the distribution of a network performance measure, based on a finite set of disaster areas and occurrence probabilities [4]. To obtain these disaster areas, they made use of the same type of earthquake scenarios as in [10] and [11], but took a more deterministic approach by putting a threshold on the seismic intensity; links crossing any cell with an intensity above the threshold fail. By merging these cells, each scenario can be described by a single, potentially disjoint, disaster area.

J. Tapolcai et al. presented a tractable stochastic model for regional link failures taking into account the correlation between the failures [5]. In case of any link set S, the output of the model tells the probability that when the next disaster happens, the set of failed links will be exactly S.

III. On Seismic Hazard

A. Earthquake Activity Rates

Earthquake activity rates are the occurrence rates of earthquake events as a function of space, time, and magnitude. These rates are a primary component of any seismic hazard analysis, including our study of network earthquake resiliency. To obtain the earthquake activity rates, we first need to define an *earthquake source model*, an area or an active fault that could host earthquakes as testified by the recent and instrumental seismic activity, historical seismicity, geomorphological evidence, and regional tectonics.

In seismic hazard analysis, there are different approaches to defining earthquake source models, e.g.: *i)* the regionalization of the study region through an area source model and using seismotectonic provinces [19], *ii)* the combination of geological information from active faults with background seismicity [20], [21], and *iii)* smoothed seismicity approaches [22].

Earthquake source models commonly use the catalog of earthquakes as input, and the main differences among source models are in how they combine the earthquake catalog with knowledge of faults, geology, and active tectonics. The choice of the earthquake source model is strongly driven by the available knowledge of the area and by the scale of the problem. It may range from well-defined active faults, especially when one is working at local scale, to less understood and wider-scale seismotectonic provinces.

More in detail, a seismotectonic province, or seismogenic zone, is a geographic region where geological, seismological, and geophysical features are assumed to be uniform throughout the region. The use of seismogenic sources is useful in those regions where the instrumental and historical seismicity point out some seismic activity, but the identification of a causative source (active fault) is not an easy task (e.g. due to the presence of blind faults). In this case, it is assumed that an earthquake can occur randomly throughout the seismotectonic province, independently from what the historical and instrumental records show and even if they indicate a clustering at some preferred locations.

On the other hand, when the knowledge of the study area is quite large, the earthquake activity rates can be deduced by the geometry and kinematic of faults. To consider a fault as a seismic source, and treat it in the seismic hazard analysis, it must be active. An active fault may be defined as a fault that has evidence of recent activity, as testified by physiographic and/or paleoseismological evidence, and historical earthquake activity.

When the catalog of earthquakes covers a long period, it can be used to compute earthquake activity rates without any information of seismotectonic provinces and/or active faults, via a smoothed seismicity approach. As our catalog covers the last 1000 years, this is the approach we will take.

The standard methodology to estimate the density of seismicity in a grid is the one developed by [22]. The smoothed rate of events in each cell i is determined as follows:

$$Sr_i = \frac{\sum_j r_j \exp\left(\frac{-d^2(c_i,c_j)}{d_c^2}\right)}{\sum_j \exp\left(\frac{-d^2(c_i,c_j)}{d_c^2}\right)}, \tag{1}$$

where r_j is the cumulative rate of events with magnitudes greater than the completeness magnitude M_c in each cell i of the grid and computed from the historical catalogue of earthquakes, and $d(c_i, c_j)$ is the distance between the centers of grid cells c_i and c_j. The parameter d_c is the correlation distance. Then, the earthquake activity rates for each node of the grid are computed following the Truncated Gutenberg-Richter model [23]:

$$\lambda(M) = \lambda_0 \frac{\exp(-\beta M) - \exp(-\beta M_u)}{\exp(-\beta M_0) - \exp(-\beta M_u)} \tag{2}$$

[2]https://earthquake.usgs.gov/hazards/hazmaps/conterminous/index.php#2014

[3]http://www.j-shis.bosai.go.jp/en/

Authorized licensed use limited to: Hebrew University. Downloaded on June 02,2023 at 13:19:00 UTC from IEEE Xplore. Restrictions apply.

for all magnitudes M between M_0 (minimum magnitude) and M_u (upper or maximum magnitude); otherwise $\lambda(M)$ is 0. Additionally, λ_0 is the smoothed rate Sr_i of earthquakes at $M_w = 4.5$ and $\beta = b\ln(10)$, where b is the b-value of the magnitude-frequency distribution.

Once all possible earthquake source models that could affect the hazard of an investigated site are defined, it is possible to define the earthquake activity rates of each source by an earthquake probability distribution, known as magnitude-frequency distribution.

A magnitude-frequency distribution indicates the probability that an earthquake, of a size within the upper and lower bound of the distribution, may occur anywhere inside the source during a specified period of time. The upper and lower bound represent, respectively, the maximum magnitude, or the largest earthquake, considered for a particular source, and the minimum magnitude, or threshold value, below which there is no engineering interest or lack of data.

In this work, we focus on Italy (a European country with a high seismic hazard), and we use the historical catalog of earthquakes [14] depicted in Fig. 2 as input data to model the occurrence of moderate-to-large ($M_w > 4.5$) earthquakes for Italy via a standard smoothed approach [22] (for more details see Section V). While this does give us the rates for all combinations of epicenters and magnitudes, we still need the relation between magnitude and disaster area to be able to apply these rates to the network resiliency models.

B. The Radius of the Damaged Zone

The only earthquake effect that can be quantified at the scale of the whole country is ground shaking. Any other earthquake effect, such as induced landslides, the liquefaction potential, the coseismic fault displacement, or tsunamis, require a site investigation. Shaking intensity is localized and is generally diminishing with distance from the earthquake's epicenter. Although the intensity can be amplified in sedimentary basins, with certain kinds of unconsolidated soils, and due to particular topographic features, at the scale of the whole country we can assume conditions are relatively homogeneous, and the seismic intensity only depends on the distance from the earthquake epicenter.

In Italy, the Mercalli-Cancani-Sieberg (MCS) scale [24] is used to measure the intensity of shaking at any given location due to an earthquake. Intensity scales empirically categorize the intensity of shaking based on the effects reported by observers and are adapted for the effects that might be observed in a particular region. The MCS scale ranges from I to XII; the lower degrees of the MCS scale generally deal with how the earthquake is felt by people, while the higher numbers of the scale are based on observed structural damage. Structural damage starts being observed at an MCS intensity of VI. Thus, we assume all links (and nodes) inside the area with an MCS intensity of at least VI are damaged, while all components outside of this area remain functioning. Note that as we assume the intensity only depends on the distance from the epicenter, this disaster area takes the form of a disk.

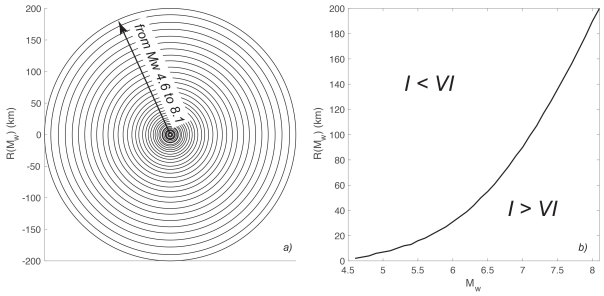


Fig. 3. a) Damaged zone ($I \geq 6$) for each 0.1 M_w (from 4.6 to 8.1). b) The black solid curve represents the variation of the distance $R(M_w)$ from the epicenter with the moment magnitude (M_w) for a VI degree of MCS scale.

Thus, to obtain all disaster areas, we now only need the disaster area radius for each magnitude $M_w \in \{4.6, 4.7, \ldots, 8.1\}$. For this purpose, we used an intensity prediction equation for Italy [25], where the expected intensity I at a site located at epicentral distance R is:

$$I = I_E - 0.0086(D - h) - 1.037(\ln D - \ln h), \qquad (3)$$

where $D = \sqrt{R^2 + h^2}$ is a sort of hypocentral distance, and $h = 3.91$km represents the hypocentral depth, which may be viewed as the average depth of the apparent radiating source [25]. I_E is the average expected intensity at the epicenter for a given earthquake and can be computed by using empirical relationships with the moment magnitude M_w. We computed the average expected intensity for a wide range of moment magnitudes (4.6 to 8.1), using the following equation taken from [25]:

$$I_E = 1.343 + 1.621 * M_w \qquad (4)$$

Then, for each I_E, we computed the expected intensity of I with Eq. (3) for an epicentral distance R that ranges from 2 to 200km. In this way, it is possible to compute for each I_E, and so M_w, the site-distance $R(M_w)$ from the epicenter of the desired intensity level (in our case equal to VI, Fig. 3).

IV. Model of Correlated Link Failures Caused by Earthquakes

The network is modeled as an undirected connected geometric graph $G = (V, E)$, with $n = |V|$ nodes and $m = |E|$ links embedded on the Earth's surface. The links can be either geodesics or a series of adjacent geodesics. 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 (see Fig. 4).

A. Stochastic Modeling of Earthquakes

To model regional failures caused by an earthquake, we apply the findings of [4] and [5]. As in the present setting, both approaches translate to the same model, we will only present the common way of modeling earthquakes. In this setting, we are investigating the failures caused by the next earthquake, thus we assume there is exactly one earthquake in the investigated time period. The earthquake is identified with its epicenter and moment magnitude:

Authorized licensed use limited to: Hebrew University. Downloaded on June 02,2023 at 13:19:00 UTC from IEEE Xplore. Restrictions apply.

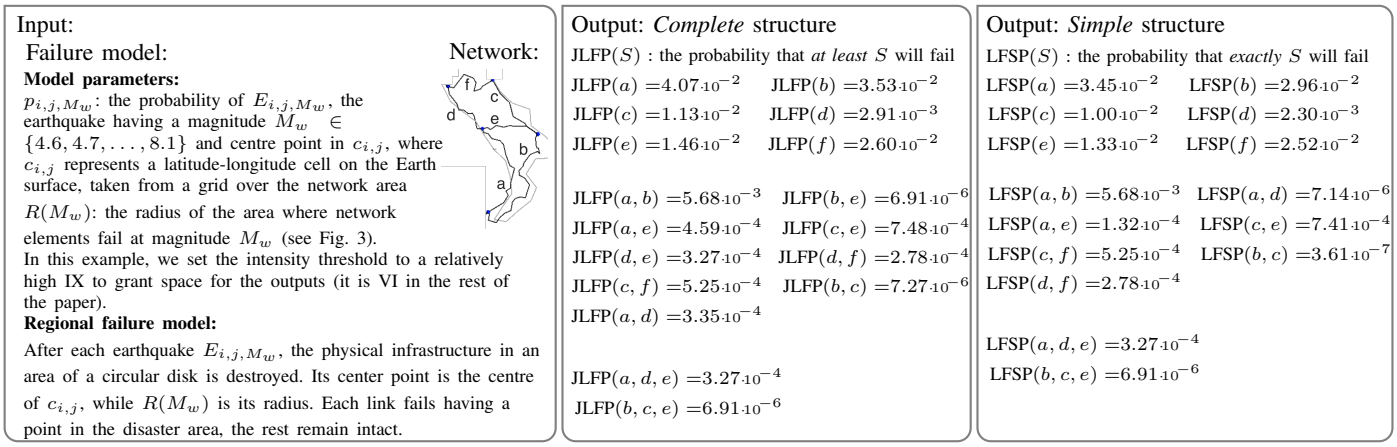


Fig. 4. An illustration of the problem inputs and outputs.

epicenter $c_{i,j}$ which represents a latitude-longitude cell on the Earth's surface, taken from a grid of cells over the network area.

moment magnitude $M_w \in \{4.6, 4.7, \ldots, 8.1\} =: \mathcal{M}$.

We index the cell grid such that $i \in \{1, \ldots, i_{max}\} =: \mathcal{I}_i, j \in \{1, \ldots, j_{max}\} =: \mathcal{I}_j$.

Let E_{i,j,M_w} denote the set of earthquakes with centre point in $c_{i,j}$ and magnitude in $(M_w - 0.1, M_w]$. As cells and magnitude intervals are considered small enough that the failures caused by each earthquake in E_{i,j,M_w} will often be the same, we will represent all E_{i,j,M_w} with a single earthquake having a center point in the center of $c_{i,j}$ and a magnitude of M_w. Let the probability that the next earthquake is in E_{i,j,M_w} be p_{i,j,M_w}. Note, that:

$$\sum_{i,j \in \mathcal{I}_i \times \mathcal{I}_j} \sum_{M_w \in \mathcal{M}} p_{i,j,M_w} = 1. \tag{5}$$

Our initial input are the activity rates r_{i,j,M_w} of earthquake types instead of the p_{i,j,M_w} values, so we have to translate these rates to probabilities. We claim that under the assumption that each kind of earthquake E_{i,j,M_w} is arriving according to independent Poisson arrival processes with parameters r_{i,j,M_w}, the rates of earthquakes E_{i,j,M_w} are straightforward to translate into the probabilities p_{i,j,M_w} of being the next earthquake to occur[4]:

$$p_{i,j,M_w} = r_{i,j,M_w} \Big/ \sum_{i,j \in \mathcal{I}_i \times \mathcal{I}_j} \sum_{M_w \in \mathcal{M}} r_{i,j,M_w}. \tag{6}$$

Although earthquakes can be clustered in time and space with their distribution that is over-dispersed if compared to the Poisson law [26], a common way to treat this problem (i.e. cluster in time and space) is to decluster the earthquake catalog, as we do and explain in Section V.

[4]Let $\mathcal{E}$ be the set of earthquake scenarios considered. At any given point in time, time $A(i,j,M_w)$ until the next $E_{i,j,M_w} \in \mathcal{E}$ has an exponential distribution with parameter r_{i,j,M_w}, and these distributions are independent of each other for all earthquakes from $\mathcal{E}$. Thus, the probability that $A(i,j,M_w) = \min(A(i,j,M_w)|(i,j) \in \mathcal{I}_i \times \mathcal{I}_j, M_w \in \mathcal{M})$ is $r_{i,j,M_w} \big/ \sum_{i,j \in \mathcal{I}_i \times \mathcal{I}_j} \sum_{M_w \in \mathcal{M}} r_{i,j,M_w}$. Since $A(i,j,M_w)$ being minimal means E_{i,j,M_w} is the next earthquake to occur among $\mathcal{E}$, Eq. (6) holds.

After each earthquake E_{i,j,M_w}, the physical infrastructure (such as optical fibers, amplifiers, routers, and switches) in an area *disk*$(c_{i,j}, R(M_w))$ of a circular disk is destroyed. The center point of *disk*$(c_{i,j}, R(M_w))$ is the center of $c_{i,j}$, while its radius $R(M_w)$ is a monotone increasing function of the magnitude M_w (see Fig. 3 for details.) As earthquakes can occur anywhere in the cell, we increase the radius by the distance between the center of the cell and its outer corners. This way, the disk is always an overestimate of the damaged area of any earthquake in cell $c_{i,j}$ with magnitude M_w.

Definition 1: We assume an earthquake E_{i,j,M_w} will result in a failure of every link of network G that has a point in *disk*$(c_{i,j}, R(M_w))$.

B. The Failure Probability of a Given Link Set

Let us denote the set of failed links by $F(i,j,M_w)$. Let $I_{i,j,M_w}(S)$ be the indicator variable of earthquake E_{i,j,M_w} hitting exactly link set S, i.e.:

$$I_{i,j,M_w}(S) = \begin{cases} 1 & \text{if } S = F(i,j,M_w) \\ 0 & \text{otherwise.} \end{cases}, \tag{7}$$

We will introduce two representations of the failure probabilities:

1) The Link Failure State Probability (LFSP)*:* of a link set S denotes the probability that the failed link set will be exactly S and let LFSP(S) denote this probability:

$$\text{LFSP}(S) = \sum_{i,j \in \mathcal{I}_i \times \mathcal{I}_j} \sum_{M_w \in \mathcal{M}} p_{i,j,M_w} I_{i,j,M_w}(S). \tag{8}$$

2) The Joint Link Failure Probability (JLFP)*:* of a link set S denotes the probability that the failed link set will be at least S and let JLFP(S) denote this probability:

$$\text{JLFP}(S) = \sum_{T \supseteq S, T \subseteq E} \text{LFSP}(T). \tag{9}$$

We will refer to the sets LFSPs and JLFPs describing the effects of the earthquakes as the *Simple* and *Complete* data structures, respectively.

Authorized licensed use limited to: Hebrew University. Downloaded on June 02,2023 at 13:19:00 UTC from IEEE Xplore. Restrictions apply.

Fig. 5. Interoute network topolgy with 25 nodes and 35 links.

Both representations have their advantages and drawbacks: while JLFP(S) can be queried directly to obtain the joint failure probability of a link-set, the number of LFSPs needed to describe the stochastic effect of the next earthquake can be significantly smaller.

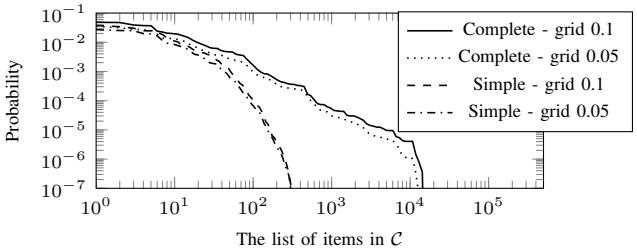
V. Simulation Results

We have used the backbone topology of Interoute's network, which is the largest privately owned Europe-wide IP cloud. They also act as an infrastructure provider, and the locations of their fibers are shown on their web page[5]. We created a copy of their network in the Italian region using the yEd graph drawing tool. We created at most 10 corner points of each fiber link such that the error in mapping is less than a few kilometers. Fig. 5 shows the network topology we created.

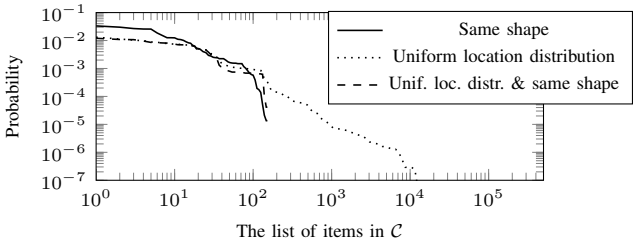
To compute the earthquake activity rates (r_{i,j,M_w}), we followed the smoothed approach described in Section III, and we used the Italian historical catalog, which covers approximately the last 1000 years (from 1 January 1005 to 28 December 2014) and consists of 4427 events. Before using the catalog, we removed all events not considered main shocks via a declustering filter [27]. This process resulted in a catalog composed of 1839 independent events. We adopted the completeness magnitude thresholds M_c over different periods given by [28], and, following [29], used a correlation distance d_c of 30km.

We took the maximum magnitude M_u assigned to each node of the grid from the European Seismic Hazard Model [30]. The ESHM13 project evaluated the maximum magnitudes of large areas of Europe based on a joint procedure involving historical observations and tectonic regionalization. We adopted the lowest value of the maximum magnitude distributions proposed by ESHM13. We calculated the b-value of the Gutenberg-Richter (GR) distribution on a regional basis using the maximum-likelihood method from [31], which allows for multiple periods with varying completeness levels to be combined. Following the approach proposed in [32] and based on [29], we used a penalized likelihood-based method for the spatial estimation of the GR b-values based on the

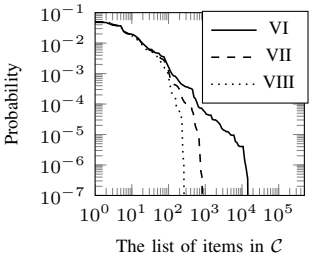
[5]https://network-map.interoute.com



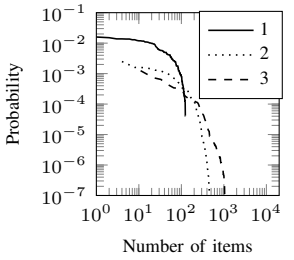
(a) Complete and Simple structures (intensity threshold VI)



(b) Complete structure with different scenarios in earthquake modeling



(c) Complete structures in case of different intensity thresholds (grid fineness 0.05)



(d) JLFPs with different cardinality in the Complete structure (intensity threshold VI, grid fineness 0.05)

Fig. 6. The size of the data structures representing regional failures caused by an earthquake.

Voronoi tessellation of space without tectonic dependency. Finally, we used Eq. 2 to compute the earthquake activity rates for each cell and each magnitude $M_w \in \{4.6, 4.7, \ldots, 8.1\}$.

We have repeated the simulation of [5] using the earthquake data we have discussed in this paper. Instead of evaluating the protection schemes we focus on probabilistic Shared Risk Link Groups (SRLGs) which are the key input to survivable routing, network planing algorithms and network service availability computations. We evaluated two data structures: Complete [5] and Simple [4]. In the Complete data structure, each set of links subject to joint failure because of a single earthquake is stored, while in Simple only the disaster set is stored [4]. Note that the Simple data structure is smaller, while it may take longer to compute the availability of an actual network service.

Fig. 6a corresponds to Fig. 6a of [5] and shows a distribution of the probabilities of stored items in the associative arrays. The number of link sets with probability more than 10^{-4} is around 1000, which was about 5000 in [5] when a uniform distribution was assumed for the epicenter of the natural disasters. We repeated the simulation with two grid parameters 0.1×0.1 decimal degree in latitude and longitude,

Authorized licensed use limited to: Hebrew University. Downloaded on June 02,2023 at 13:19:00 UTC from IEEE Xplore. Restrictions apply.

and 0.05×0.05. There was no significant difference between the two resolutions. To analyze the assumptions we used in earthquake modeling, we have computed the probabilistic SRLGs for three scenarios: *Uniform location distribution:* The probabilities of earthquakes are uniformly distributed over the grid cells. However, we still have multiple possible magnitudes at each location. Roughly speaking, we rescale their probability based on their rate compared to the total rate at the cell. *Same shape:* The epicenter distribution is not uniform. However, we only consider one event per cell: a circle with the (weighted) average of the radii of all possible/considered earthquakes. *Uniform location distribution & same shape:* Earthquakes are uniformly distributed over the grid cells, and each grid cell only has one event: a circle with the (weighted) average of the radii of all possible/considered earthquakes.

Comparing the graphs on Fig. 6b and 6a, we can deduce that the assumption of the uniform location distribution has little effect, since we keep the same link sets of LFSPs, and only their probabilities change. In contrast to this, fixing the disaster shape drastically reduces the number of LFSPs.

Fig. 6c shows the data set if we change the MCS intensity threshold representing the resistance of the network infrastructure against earthquakes. In other words, MCS VIII means the network is built with more earthquake-proof buildings and infrastructure. Fig. 6d 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 $|S|$.

VI. Conclusion

In this study, we focused on computing the vulnerability of networks to earthquakes, using empirical data. To capture the inter-correlation of link failures, traditional ground-shaking hazard maps are insufficient. We have thus created an earthquake model – applied to Italy in this paper – based on an historical earthquake catalogue over the last 1000 years. Based on this catalogue, we generated earthquake activity rates and investigated how the ground-shaking intensity diminishes with the distance from the epicenter and identified the network links that were subject to structural damage in case of possible earthquakes. Combining these, we analyzed the vulnerabilities of the Italian portion of Interoute's fiber network.

This paper is a step towards a deeper understanding of the risk of failures that backbone networks are subject to by bringing two fields together: Seismology and Telecommunications.

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*. USC Tech. Rep., 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] J. Tapolcai, B. Vass, Z. Heszberger, J. Biró, D. Hay *et al.*, "A tractable stochastic model of correlated link failures caused by disasters," in *Proc. IEEE INFOCOM*, Honolulu, USA, Apr. 2018.

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

[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. Netw.*, vol. 21, no. 5, 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. Netw.*, 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," *J. Lightw. Technol.*, vol. 30, no. 16, pp. 2563–2573, 2012.

[10] P. N. Tran and H. Saito, "Enhancing physical network robustness against earthquake disasters with additional links," *Journal of Lightwave Technology*, vol. 34, no. 22, pp. 5226–5238, 2016.

[11] ——, "Geographical route design of physical networks using earthquake risk information," *IEEE Communications Magazine*, vol. 54, no. 7, 2016.

[12] L. Ma, X. Jiang, B. Wu, A. Pattavina, and N. Shiratori, "Probabilistic region failure-aware data center network and content placement," *Computer Networks*, vol. 103, pp. 56–66, 2016.

[13] T. Gomes, J. Tapolcai, C. Esposito, D. Hutchison, F. Kuipers *et al.*, "A survey of strategies for communication networks to protect against large-scale natural disasters," in *RNDM*, 2016.

[14] A. Rovida, M. Locati, R. Camassi, B. Lolli, and P. Gasperini, "Cpti15, the 2015 version of the parametric catalogue of italian earthquakes," *Istituto Nazionale di Geofisica e Vulcanologia*, 2016.

[15] D. Bienstock, "Some generalized max-flow min-cut problems in the plane," *Mathematics of Operations Research*, vol. 16, no. 2, 1991.

[16] 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.

[17] S. Neumayer, A. Efrat, and E. Modiano, "Geographic max-flow and min-cut under a circular disk failure model," *Computer Networks*, 2015.

[18] X. Wang, X. Jiang, and A. Pattavina, "Assessing network vulnerability under probabilistic region failure model," in *IEEE High Performance Switching and Routing (HPSR)*, 2011, pp. 164–170.

[19] C. Cornell, "Engineering seismic risk analysis," *Bullettin of the Seismological Society of America*, vol. 58, no. 5, pp. 1583–1606, 1968.

[20] E. H. Field, G. P. Biasi, P. Bird, T. E. Dawson, K. R. Felzer *et al.*, "Long-term time-dependent probabilities for the third uniform california earthquake rupture forecast (ucerf3)," *Bullettin of the Seismological Society of America*, vol. 105, pp. 511–543, 2015.

[21] A. Valentini, B. Pace, P. Boncio, F. Visini, A. Pagliaroli, and F. Pergalani, "Definition of seismic input from fault-based psha: Remarks after the 2016 central italy earthquake sequence," *Tectonics*, vol. 38, 2019.

[22] A. Frankel, "Simulating strong motions of large earthquakes using recordings of small earthquakes: the loma prieta mainshock as a test case," *Bulletin of the Seismological Society of America*, 1995.

[23] Y. Y. Kagan, "Seismic moment distribution revisited: I. statistical results," *Geophysical Journal International*, vol. 148, no. 3, 2002.

[24] A. Sieberg, "Erdebeben," *Handbuch der Geophysic*, vol. 4, 1931.

[25] P. C., D. Albarello, P. Gasperini, V. DAmico, and B. Lolli, "The attenuation of seismic intensity in Italy, part ii: modeling and validation," *Bulletin of the Seismological Society of America*, vol. 98, no. 2, 2008.

[26] Y. Y. Kagan, "Statistical distributions of earthquake numbers: consequence of branching process," *Geophysical Journal Int.*, vol. 180, 2010.

[27] J. K. Gardner and L. Knopoff, "Is the sequence of earthquakes in southern california, with aftershocks removed, poissonian?" *Bulletin of the Seismological Society of America*, vol. 64, pp. 1363–1367, 1974.

[28] M. Stucchi, C. Meletti, V. Montaldo, H. Crowley, G. M. Calvi, and E. Boschi, "Seismic hazard assessment (2003-2009) for the Italian building code," *Bulletin of the Seismological Society of America*, 2011.

[29] A. Valentini, F. Visini, and B. Pace, "Integrating faults and past earthquakes into a probabilistic seismic hazard model for peninsular italy," *Natural Hazards and Earth System Sciences*, vol. 17, 2017.

[30] J. Woessner, D. Laurentiu, D. Giardini, H. Crowley, F. Cotton *et al.*, "The 2013 european seismic hazard model: key components and results," *Bulletin of Earthquake Engineering*, vol. 13, no. 12, 2015.

[31] D. H. Weichert, "Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes," *Bulletin of the Seismological Society of America*, vol. 70, no. 4, 1980.

[32] Y. Kamer and S. Hiemer, "Data-driven spatial b value estimation with applications to california seismicity: To b or not to b," *Journal of Geophysical Research: Solid Earth*, vol. 120, no. 7, 2015.

Authorized licensed use limited to: Hebrew University. Downloaded on June 02,2023 at 13:19:00 UTC from IEEE Xplore. Restrictions apply.