How to Model and Enumerate Geographically Correlated Failure Events in Communication Networks

Vass, Balazs; Tapolcai, János; Hay, David; Oostenbrink, Jorik; Kuipers, Fernando

DOI
10.1007/978-3-030-44685-7_4

Publication date
2020

Document Version
Accepted author manuscript

Published in
Guide to Disaster-Resilient Communication Networks

Citation (APA)
Vass, B., Tapolcai, J., Hay, D., Oostenbrink, J., & Kuipers, F. (2020). How to Model and Enumerate Geographically Correlated Failure Events in Communication Networks. In J. Rak, & D. Hutchison (Eds.), *Guide to Disaster-Resilient Communication Networks: Springer International Publishing* (pp. 87-115) https://doi.org/10.1007/978-3-030-44685-7_4

How to Model and Enumerate Geographically Correlated Failure Events in Communication Networks

Balázs Vass, János Tapolcai, David Hay, Jorik Oostenbrink, Fernando Kuipers

Abstract Several works shed light on the vulnerability of networks against regional failures, which are failures of multiple pieces of equipment in a geographical region as a result of a natural or human-made disaster. This chapter overviews how this information can be added to existing network protocols through defining *Shared Risk Link Groups (SRLGs)* and *Probabilistic SRLGs (PSRLGs)*. The output of this chapter can be the inputs of later chapters to design and operate the networks to enhance the preparedness against disasters and regional failures in general. In particular, we are focusing on the state-of-the-art algorithmic approaches for generating lists of (P)SRLGs of the communication networks protecting different sets of disasters.

1 Introduction

The Internet is a critical infrastructure. Due to the importance of telecommunication services, improving the preparedness of networks to regional failures is becoming a key issue [5, 6, 9, 11, 13, 14, 10, 21, 22, 23, 28, 41]. The majority of severe network outages happen because of a disaster (such as an earthquake, hurricane, tsunami, tornado, etc.) taking down many or all equipment in a given geographical area. Such failures are called *regional failures*. Many studies touched the problem

Balázs Vass, János Tapolcai
MTA-BME Future Internet Research Group, High-Speed Networks Laboratory (HSN*Lab*) and Budapest University of Technology and Economics, Budapest, Hungary e-mail: {vb, tapolcai}@tmit.bme.hu

David Hay
School of Engineering and Computer Science, Hebrew University, Jerusalem, Israel, e-mail: dhay@cs.huji.ac.il

Jorik Oostenbrink, Fernando Kuipers
Delft University of Technology, Delft, the Netherlands e-mail: {J.Oostenbrink, F.A.Kuipers}@tudelft.nl

of how to prepare networks to survive regional failures. The topic was started as a subproblem in the related articles solved with simple solutions, such as assuming that fibres in the same duct or the 50km neighbourhood of every network node is subject to regional failure [19, 43]. Next, they were improved by examining the historical data of different type of disasters (e.g. seismic hazard maps for earthquakes) and identify the hotspots of the disasters [6, 11, 13, 21, 22, 28]. The weak point of these approaches is that during installing a network equipment many risks are also considered and compensated. For example, earthquake-proof infrastructure is used in areas with larger seismic intensity. Therefore, based on historical data, it is far not obvious to reduce the problem space, and it may be more realistic to assume that any set of equipment physically close to each other has a higher chance to fail together. More recent studies are purely devoted to this particular problem and adapt combinatorial geometric based approaches to capture all of the regional failures and represent them in a compact way [3, 8, 23, 31, 32, 34, 37]. Here the challenge is that these regional failures can have arbitrary locations, shapes, sizes, effects, etc. This section is devoted to overview of the state of the art and suggests unified definitions, notions and terminology.

The output of the approaches discussed in this section can serve as the input of the network design and management tools. Currently, network recovery mechanisms are implemented to protect a small set of pre-defined failure scenarios. Each recovery plan corresponds to the failure of some equipment. Informally speaking when a link fails, the network has a ready-to-use plan on how to recover itself. Technically a set of so-called Shared Risk Link Groups (SRLGs) are defined by the network operators, where each SRLG is a set of links whose joint failure the recovery mechanism should be prepared for. In this chapter, we are purely focusing on how to define SRLGs that cover all types of disasters, and we omit the problem of how to implement a recovery mechanism for a specific SRLG. We will also address refinements of the SRLG model defined in the next section.

2 Notions Related to Vulnerable Regions

When several network elements may fail together as the result of some event, they are often characterized by *Shared Risk Group*s (SRGs). Each SRG has a corresponding failure event (or events); when such an event occurs, all elements in the SRG fail together. Specifically, the communication network is modelled as a graph $G = \langle V, E \rangle$, whose vertices are routers, PoPs, optical cross-connects (OXC), and users, while the edges are communication links (mostly optical fibres). SRGs are then defined as subgraphs $\langle V', E' \rangle$, where $V' \subseteq V$ and $E' \subseteq E'$.

In many cases, it is sufficient to consider only links in SRGs, and in this case, these groups are called *Shared Risk Link Groups* (SRLGs). For example, an SRLG may contain one edge (to capture single-link failures) or all edges that touch one vertex (to capture single-node failures). In this chapter, most of the SRLGs are more complex and represent the simultaneous failures of multiple network elements. In particular, we are focusing on regional failures in which all links within a specific geographical region fail.

A set $\mathscr{S}$ of SRLGs can be used as an input to network design and network recovery/protection mechanisms to ensure these mechanisms withstand the failures corresponding to these SRLGs. For example, to ensure connectivity between two nodes, protection mechanisms may use two edge-disjoint paths when $\mathscr{S} = \{\{e\}|e \in E\}$, two node-disjoint paths when $\mathscr{S} = \{\{(u,v) \in E\}|v \in V\}$, and two paths that do not traverse the same geographical region when $\mathscr{S}$ corresponds to all sets of links that are close-by and can fail simultaneously.

The following definition captures the notion of SRLG introduced by regional failures, such as a natural disaster or an attack. For ease of presentation, we will call these failure events *disasters*, regardless of their cause.

Definition 1 (SRLG). A set of links $S \subseteq E$ is an SRLG if we may assume there will be a disaster that can cause all edges in S to fail together. If the disaster can be characterized by a bounded geographical area in the two-dimensional plane $D \subset \mathbb{R}^2$, and S is the set of edges that intersect with D, then S is called the *regional SRLG that represents* D, and is denoted by $S = \text{SRLG}(D)$. If D is a circular disk, we call $\text{SRLG}(D)$ a *circular* SRLG.

Circular SRLGs, which are the most common in literature, can also be characterized by the failure epicenter $p \in \mathbb{R}^2$ and the failure radius $r \in \mathbb{R}$; in this case $S = \{e \in E|d(e,p) \leq r\}$, where $d(e,p)$ is the Euclidean distance between edge e and point p.

Below we give two best practices for reducing the number of listed regional SRLGs:

As the size of $\mathscr{S}$ determines the run-time and complexity of the mechanisms that use it, an important goal is to keep $\mathscr{S}$ as small as possible. For example, when two sets S_1, S_2 are in $\mathscr{S}$ and $S_1 \subseteq S_2$, it is sufficient to include only S_2 in $\mathscr{S}$; omitting S_1 from $\mathscr{S}$ does not affect the outcome of the underlying mechanisms. Moreover, some works use *over-approximation* to reduce the size of $\mathscr{S}$: instead of including two sets S_1, S_2 one can include a single set $S_1 \cup S_2$ (this is especially appealing if $S_1 \cap S_2$ is of non-negligible size); such over-approximation, however, can degrade the outcome of the underlying mechanisms. For regional SRLG, over-approximation is achieved by taking a larger failure region. The most common practice is to take a simpler shape (e.g., circular disk) that completely contains the original failure region:

Assumption 1 *The failed region has a circular shape.*

Dealing with circular SRLGs are in fact over-approximations of regional SRLGs. Notice for example, that one can over-approximate failures with a certain radius r, with failures of radius $r' > r$ (namely, assuming all disasters cause larger damage). If such an over-approximation is plugged into network protection mechanism (e.g., one that computes secondary paths that are SRLG-disjoint from the primary paths), this will cause a performance degradation (namely, longer secondary paths).

Another very common practice is to assume that there cannot be multiple disasters:

Assumption 2 *In the investigated time period, there will be at most one disaster.*

The likelihood of a disaster to occur is not the same at all points of the plane. For example, earthquakes are more likely to occur in rapture zones than in other places, and regions with lower altitude are more likely to suffer from floods. Thus, the probability of an event to occur is important. This probability is sometimes given in the form of a *epicenter distribution map*, which gives for each location $p \in \mathbb{R}^2$, the probability that a disaster happened with epicentre p. Moreover, the size (or radius) of the disaster can also be a random variable (e.g., earthquakes with a larger magnitude are less likely to happen than earthquakes with smaller magnitude, even if their epicentres are the same). Thus, it is customary to consider a set $\mathscr{D}$ of disasters $D \subseteq \mathbb{R}^2$ (that can be of infinite, size) and attach a probabilistic measure to this set. For simplicity, let's assume that $\mathscr{D}$ is finite, and let $p_D = \Pr[disaster\, D \in \mathscr{D} \text{ occurs}]$[1] We note that an SRLG S can represent more than one disaster in $\mathscr{D}$; thus, we denote by the support$(S) = \{D \in \mathscr{D} | S = SRLG(D)\}$.

Definitions 2-4. capture the probabilistic nature of disasters and their effect on SRLGs. A FP (Def. 2) tells the probalility that the failed link set will be *exactly S*, while a CFP (Def. 3) tells the probability that *at least S* will fail:

Definition 2 (FP). Given a set $\mathscr{D}$ of disasters $D \subseteq \mathbb{R}^2$ and a probability p_D for each disaster in $\mathscr{D}$, a *Link Failure State Probability (FP)* (S, p_S) is a set of links $S \subseteq E$ and their failure probability: $p_S = \sum_{D \in \text{support}(S)} p_D$. We note that if a disaster in support(S) actually occurs, then all links in S fail (with probability 1).

Definition 3 (CFP). Given a set $\mathscr{D}$ of disasters $D \subseteq \mathbb{R}^2$ and a probability p_D for each disaster in $\mathscr{D}$, a *Cumulative Link Failure Probability (CFP)* (S, P_S) is a set of links $S \subseteq E$ and their failure probability: $P_S = \sum_{T \supseteq S} \sum_{D \in \text{support}(T)} p_D$. We note that if a disaster in $\bigcup_{T \supseteq S}$ support(T) actually occurs, then all links in S fail (with probability 1).

In a sense, FPs are like probability density functions (PDFs), while CFPs are like their cumulative distribution functions (CDFs).

In case of two-stage PSRLGs (Def. 4), links are not necessarily destroyed if a disaster hits them:

Definition 4 (two-stage PSRLG). Let be given 1) a set $\mathscr{D}$ of disasters $D \subseteq \mathbb{R}^2$, 2) for each disaster in $\mathscr{D}$, a probability p_D of occuring, and 3) if D occurs, for each link $l \in SRLG(D)$, a probability $p_{D,l}$ of failing (independently of the other links). A *two-stage Probabilistic SRLG*, $(S, p_S; p_1, \ldots, p_{|S|})$ is a set of links $S = \{l_1, \ldots, l_{|S|}\} \subseteq E$, the failure probability of the set: $p_S = \sum_{D \in \text{support}(S)} p_D$, and failure probabilities of links l_i: $p_i = \frac{1}{p_S} \sum_{D \in \text{support}(S)} p_D p_{D,l_i}$, where if a disaster in support(S) actually occurs, then link $l_i \in S$ fails with probability p_i, independently of the other links.

[1] For infinite sets, one can use discretization and consider only finite number of sets, albeit with a small error.

If in case of each two-stage PSRLG S, links being part of S fail with the same probability, S is called a *homogeneous two-stage PSRLG*, or else it is a *heterogeneous two-stage PSRLG*.

Collectively, we call FPs, CFPs and two-stage PSRLGs as *Probabilistic SRLGs (PSRLGs)*. Fig. 1 depicts the connections between these notions.

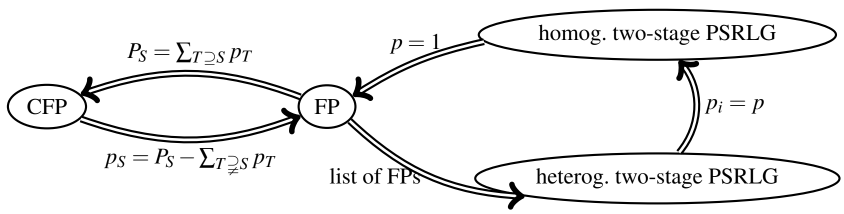


Fig. 1: Relation between Probabilistic SRLGS (PSRLGs) (two-srage PSRLGs, FPs, and CFPs): an FPis a homogeneous two-stage PSRLG with $p = 1$, which is a heterogeneous two-stage PSRLG with $p_i = p$. In addition, a heterogeneous two-stage PSRLG can be represented by a list of FPs, and lists CFPsand FPsare also easily interchangeable by definition.

We can convert a list of heterogeneous two-stage PSRLG into a list of FPs as follows. Take a heterogeneous PSRLG $(S, p_S; p_1, \dots, p_{|S|})$; the probability p_P of failing exactly a nonempty set $P \subseteq S$ is $p_S \prod_{e \in P} p_i$, thus one can store sets P with probabilities p_P in a list of FPs. A list of heterogenous PSRLG can be transformed similarly.

Tables 1 and 2 give an overview of the works presented in this chapter. Papers offering lists of SRLGs and PSRLGs translate the composed geometric problem of protecting telecommunication networks against regional failures to purely combinatorial and probabilistic problems, respectively[2] [3].

In the followings, we will present works showing that the composed geometric problem of protecting telecommunication networks against regional failures translates to combinatorial problems via generating (P)SRLGs. Then one can use a variety of known tools to handle the translated combinatorial problems.

3 Calculating Lists of SRLGs

Beforehand we note that although many of the presented methods are designed to handle links which are considered as line segments (or geodesics) between their endpoints, these results can be extended to a more general setting, where the links

[2] Some papers like [7, 21, 40, 42, 43] are loosely related to the regional (P)SRLG generating problem, however our goal is presenting the most relevant works in this field.

[3] A list of (P)SRLGs can be used as a pre-computed input for various problems [4, 12, 15, 18].

Paper	Current chapter	Goal	Geometric info.		Assumptions		Algorithms		
			Physical network	Planar/ Spherical	Disaster shape	Single disaster	Precise/ approximate	Poly- nomial	Paramet- rized
Tapolcai et al. [30, 31]	Sssec. 3.1.3	SRLG list	good	plane	circular	✓	precise	✓	✓
Vass et al. [37, 38]	Sssec. 3.1.2	SRLG list	poor	plane	circular	✓	precise	✓	✓
-	Sssec. 3.1.1	SRLG list	no	plane	-	✓	precise	✓	✗
Vass et al. [39]	3.1.4, 3.2	SRLG list	any	plane+ shpere	bounded by segments+arcs	✓	precise+ approximate	✓	✓
Iqbal et al. [17]	Ssec. 3.3.1	SRLG list	good	plane	-	✓	precise	✗	✗
Neumayer et al. [23]	Sssec. 3.3.2	most vulnerable point	good	plane	circular or line segment	✓	precise	✗	✗
Pašić et al. [26, 27]	Sec. 5	SRLG list	good	plane	any	✓	approximate	✓	✗

Table 1: Papers enumerating regional SRLGs. While the rest of the papers consider deterministic disaster scenarios, in [26, 27] SRLGs are obtained from PSRLG lists.

Paper	Current chapter	Goal	Correlated link failures inside the disaster	Natural disaster / attack
Agarwal et al. [2, 3]	Subsec. 4.2	most vulnerable point	✗	attack
Oostenbrink et al. [24]	Sssec. 4.1.1	FP list	(✓)	-
Tapolcai et al. [32]	Sssec. 4.1.2	CFP list	✓	natural disaster
Valentini et al. [36]	Sssec. 4.1.3	FP list + CFP list	✓	natural disaster (earthquake)

Table 2: Papers enumerating regional PSRLGs

are polygonal chains (or series of geodesics) at the price of a polynomial increase in runtime[4].

3.1 *Precise Polynomial Algorithms Enumerating Lists of SRLGs*

In case of no geometric information:

3.1.1 SRLG Lists Induced by Hop Count

The current best practice to increase the resilience of the networks against disasters is to ensure that the primary and backup paths assigned to a connection are node disjoint. Compared to edge-disjointness, in this way operators ensure that the distance between the nodes of the primary and backup paths (except at the terminal nodes) are at least 1-hop-distance from each other. The in-

[4] Polygonal chains can be dismantled to a set of line segments, the method can be applied, then the sets of line segments can be joined.

tuitive reasoning is that a link in a backbone network is typically a few hundred kilometres long, while natural disasters are never larger than a few hundred kilometres.

Let M_h denote the set of link sets ensuring a distance of h hops. Trivially, $M_{h=1} = E$, i.e. it is the set of single link failures, while protection against single node failures ensures 2 hops distance, i.e. $M_{h=2} = \cup_{v \in V} \{u, v\} \in E$. To ensure an even number of hops, for every node v, $M_{h=2k}$ contains the edges of the tree of the shortest paths to v from the nodes not further from v than k hops.. Similarly, for every link $e = \{u, v\}$, $M_{h=2k+1}$ contains the edges of the tree of the shortest paths to e from the nodes not further from u or v than k hops. We can conclude that the number of SRLGs in M_h is low $|M_{h=2k}|$ being $|V|$, and $|M_{h=2k+1}|$ being $|E|$.

Fig. 2 depicts average number of links contained by SRLGs in M_h. Clearly, this average is 1 for $h = 1$, and is equal for the average nodal degree for $h = 2$. For bigger values of h the average seems to grow slightly superlinearly before the growth slows down to platoe at $|V| - 1$.

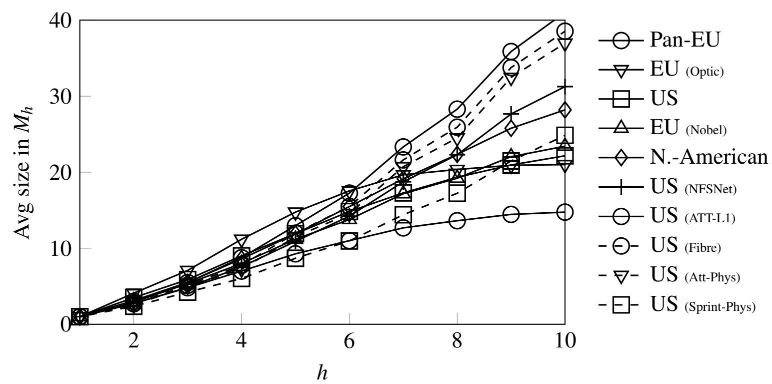


Fig. 2: The average number of links contained in the SRLGs of M_h in case of physical backbone topologies of [25].

Clearly, M_h can be computed in low polynomial time of n. To generate Fig. 2, we generated M_h for $h \in \{1, \ldots, 12\}$ for all the networks on the figure in less then 17 seconds on a commodity laptop, using a code written in Python3, not optimized for speed.

In case of poor quality geometric information:

3.1.2 SRLG Lists of Disasters Being Circular Disks Containing k Nodes

As mentioned before, the current best practice is to ensure that the primary and backup paths assigned to a connection are node disjoint. Compared to edge-disjointness, in this way operators ensure that the distance between the nodes of the primary and backup paths (except at the terminal nodes) are at least 1-hop-distance from each other. The intuitive reasoning is that a link in a backbone network is typically a few hundred kilometres long, while natural disasters are never larger than a few hundred kilometres. The root of the outages is usually because: I) *close nodes* when two nodes are placed close to each other; for example, in highly populated areas. II) *parallel links* when two links are placed close to each other because of some geographic reasons.

Unfortunately, handling the geometric information with the network topology is not part of the current best practice. Furthermore, the Internet Service Provides usually hire the links as a service from an independent company, called the Physical Infrastructure Provider, and thus, operators have no information about the route of the links, or the physical coordinates of the intermediate routing nodes.

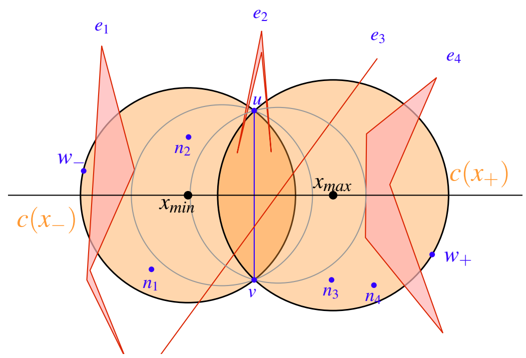


Fig. 3: Illustration of an apple with $k = 2$. Apple $A_k^{u,v}$ consists of specific ordered lists of links and nodes which can be hit by a disk from $\mathscr{C}_k^{u,v}$.

In [38], a limited geometric information failure model is defined, which is based on the following assumptions:

1. The network is a geometric graph $G(V,E)$ embedded in a 2D plane.
2. The exact route of the conduits of the network links are not known, but contained by a polygonal region.
3. The shape of the regional failure is assumed to be a circular disk with an arbitrary radius and centre position.
4. It focuses on *regional link k-node failures*, failures that hit k nodes for $k \in \{0, |V|-2\}$.

[38] presents a low-polynomial algorithm for determining the set M_k of maximal regional link k-node failures. The proposed method is based on a set of (computational) geometric considerations. The key observation is that for any element of M_k there exists a circular disk-shaped disaster having k nodes in the interior which has 1) 2 nodes on its boundary, or else 2) only 1 node u on its boundary and having an infinite radius. This allows us to enumerate all possible maximal failures using a sweep surface method as follows.

Let $u\{u,v\} \subseteq V$ be two nodes for which the set $\mathscr{C}_k^{u,v}$ if circles which have k nodes in its interior and u and v on its boundary is not empty. These $\{u,v\}$ pairs are part of the set E_k of *k-Delaunay* edges, and their set can be determined in low-polynomial time [29, Thm. 2.4]. In [38], data struc-

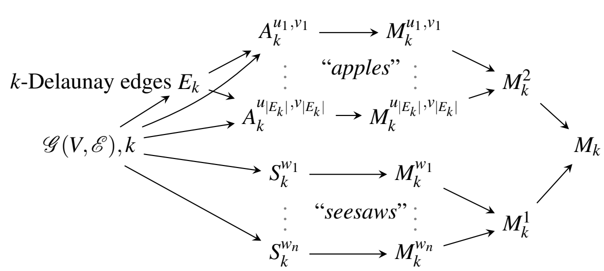
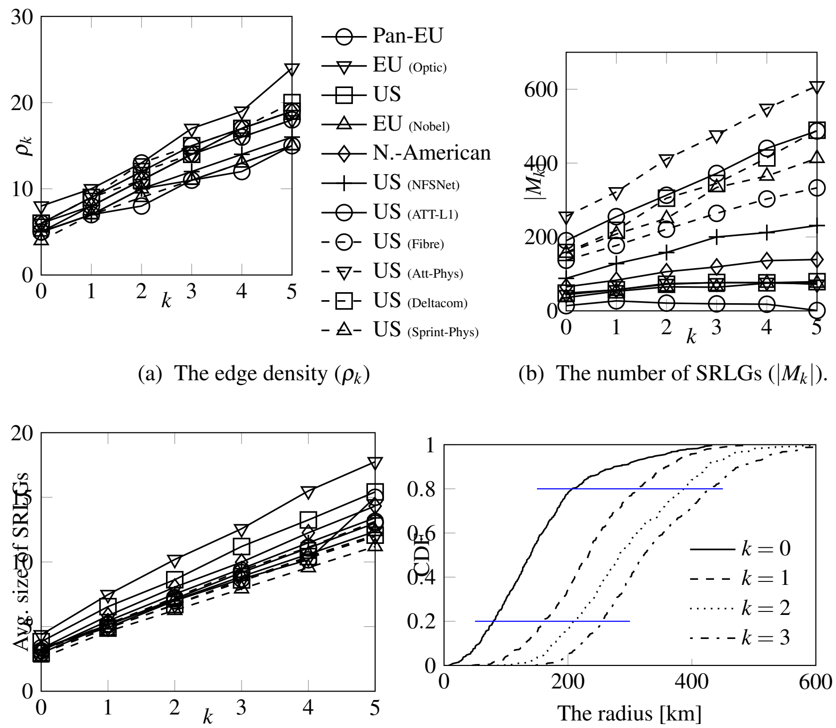


Fig. 4: Sketch of algorithm for enumerating set M_k of maximal link sets which can be hit by a circular disk hitting k nodes.

ture *apple*[5] $A_k^{u,v}$ is defined, which contains ordered lists of links and nodes which can be hit by a circle from $\mathscr{C}_k^{u,v}$. Suppose u and v are positioned as in 3. With the help of $A_k^{u,v}$, one can sweep through circles of $\mathscr{C}_k^{u,v}$ ordered by the abscissas of their centre points allowing to collect the set $M_k^{u,v}$ of maximal hit link sets by disks from $\mathscr{C}_k^{u,v}$. Then the globally maximal elements of all lists $M_k^{u,v}$ are collected in M_k^2.

In the second case, the set of maximal failures M_1 from M_k for which exist a half plane going through a node and hitting them can be calculated similarly via turning a half plane around every node while checking the set of hit links and the number of hit nodes. Finally, M_k can be obtained via collecting the maximal elements of M_k^1 and M_k^2.

The process is sketched in Fig. 4. The complexity of the algorithm is low-polynomial and squared in the number of nodes n [38, Thm. 3, Cor. 25]. Besides theoretical upper bounds, simulation results show that the number of maximal failures is approximately $1.2n$ and $2.2n$ for $k = 0$ and $k = 1$, respectively (Fig. 5).



(a) The edge density (ρ_k)

(b) The number of SRLGs ($|M_k|$).

(c) The average size of SRLGs

(d) The cumulative distribution function of the radius of the disk having a given k number of nodes interior.

Fig. 5: The edge density, number and size of SRLGs for each network and $k = \{0,\ldots,5\}$ in case of polygonal chain links.

[5] For more details, please check [38].

In case of good quality geometric information:

3.1.3 Circular SRLG Lists of Disasters with Radius r

If the physical positions of the network elements are known, a fast systematic approach to generate the list M_r of maximal SRLGs that represent circular disks of a given radius r is clearly desired[6].

Paper [31] presents a low-polynomial algorithm for computing M_r, when links are considered as line segments (and the network is embedded in the plane). It shows that the number of elements of M_r is linear in the number of nodes in the network n, and its calculation can be done in a squared complexity of n (Theorem 6 of [31]). Simulations indicate that this list has a size of $\approx 1.2n$ in practice.

To be more precise with the theoretical results, Corollary 4 and Theorem 6 of [31], respectively tell that the number of SRLGs in M_r is at most proportional to both the number of nodes n plus the number of link intersections x, and in the cardinality ρ_r of the biggest link set contained. The computing time needed is $O\left((n+x)^2\rho_r^5\right)$. We note that x is 0 or a small number, and according to simulation results, ρ_r increases linearly with r, suggesting an $O(n^2 r^5)$ runtime for $r > 0$.

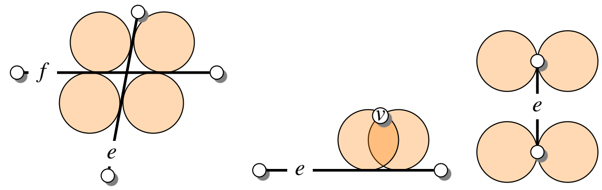
Algorithm 1: Sketch of algorithm proposed in [31]

Input: graph $G(V,E)$ embedded in plane, radius r
Output: List M_r of maximal SRLGs of disasters being circular disks with radius r
begin
    $M_r' := \emptyset$
    Calculate $X :=$ {`points of edge crossings`}
    for $w \in V \cup X$ **do**
        Determine $E_w :=$ {`edges not further from` w `than` $3r$}
        for $e_1, e_2 \in E_w$ **do**
            Calculate circles c_i described in Fig. 6/(a)
        for $e \in E_w$ **do**
            Calculate circles c_j described in Fig. 6/(b) with w as point
            Calculate circles c_k described in Fig. 6/(c)
        Refresh[7] M_r with link sets hit by circles c_i, c_j, c_k (1 circle at a time)
    return M_r

In the following, we give an overview of the proposed algorithm (Alg. 1), which relies on a series of geometric considerations. The most important one is Theorem 1 of [31], which leverages that the link sets possibly hit by any of the infinite number of possible disaster locations can be determined via checking the effect of a quadratic number of disks on the network edges. In particular, for a positive real r, and a nonempty set of edges H which is hit by a circular disk of radius r, there exists a disk c of radius r which hits the edges of H such that at least one of the following holds (see Fig. 6 for illustrations): (a) There are two non-parallel intervals in H such that c intersects both of them in a single point. These two points are different. (b) There are two intervals in H such that

[6] [19] offers a mistaken heuristic for computing M_r. It claims the disc failures having nodes of the network as their centre point represent the worst-case of failures of radius r, which is clearly not the case. Consider e.g. a network being an equilateral triangle with side length 3, and $r = 1$; here M_r consists of a single SRLG containing all the 3 links instead of the 3 link-pairs claimed by [19].

[7] This means that M' is the set of maximal failures among which are already checked, and if f is maximal amongst them, it is added to M' and all f's subsets are eliminated from M'; or if f is not maximal in M', nothing happens.

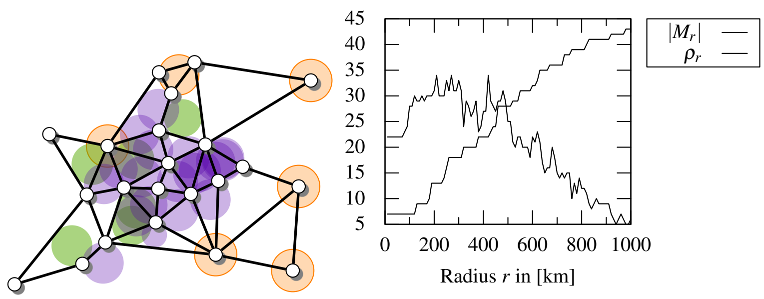


(a) $\forall e \in E$ and $\forall f \in E$ (b) $\forall e \in E$ and $\forall v \in V$ (c) $\forall e \in E$

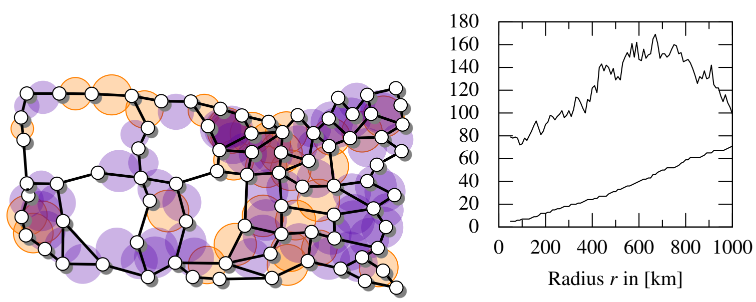
Fig. 6: The disk failures examined.

c intersects both of them in a single point. These two points are different, and one of them is an endpoint of its interval. (c) Disk c touches the line of an interval $e \in H$ at an endpoint of e.

Intuitively, there is no reason for checking for the circles described in Fig. 6 in case of two network elements which are much further apart than the disaster radius r. Indeed, one can build up the solution of the global problem based on some local calculations, as follows. Let X be the set of link intersection points. After determining X, one has to collect edges not further from w than $3r$ into a set E_w, for all $w \in V \cup X$, then determine the maximal failures of sets E_w, and finally get the result by collecting the maximal elements of the resulting lists.



(a) 22-node EU network (Nobel)



(b) 79-node US network (NSFNet)

Fig. 7: Simulation results for determining list M_r of maximal SRLGs of disasters being circular disks with radius r.

With the help of some additional computational geometric ideas, for determining M_r, one could achieve an almost linear computing complexity in the number of nodes n [30][8].

Precise representations:

3.1.4 Circular SRLG Lists of Disasters with Radius r on a Sphere

The Earth is not flat, as its shape (geoid) is much more like a sphere. With this in mind we can deduce that when studying spread-out networks (e.g. the optic fibre network of the US), in order to reach a higher precision, one should consider that networks are embedded on a spherical surface instead of the much more widespread planar embedding. Note that [39] found that M_r^s and $M_r^p = M_r$ can be different even in the case of a network having a geographical extension of 100`km`.

More precisely, [38] took network `AboveNet` [1], and its shrunk instances, where `AboveNet/c` means that `AboveNet` was rotated such that the average lat and lon coordinates to be both 0, then each coordinate was divided by c. As a similarity measure, $\mathscr{M}(r) := |M_r^p \triangle M_r^s|/(|M_r^p| + |M_r^s|) \in [0,1]$ (the ratio of SRLGs, which are present in only one of M_r^p and M_r^s) was used : if $\mathscr{M}(r)$ is close to 1, it means the two lists are very different, while if it is close to 0, it means there are few differences. Radius $r = 8$ was set to be a bit larger than the half of the diameter of the current network, $r = 0$ was set to be a small radius, the rest of the r values were linearly interpolated.

Fig. 8 shows that, while in case of `AboveNet`, M_r^p and M_r^s are almost entirely different for many values of r, the tendency is that $\mathscr{M}(r)$ decreases as the physical size of the network decreases, which nicely fits the intuition. Surprisingly, $\mathscr{M}(r)$ is not 0 for every range r even for `AboveNet/300`, which equals to the case when the approximative network diameter is 104`km`, `AboveNet/400` (having a diameter of approx. 74`km`) being the most spread out instance where M_r^p and M_r^s are the same for all investigated r ranges.

As a rule of thumb, it can be said that the difference between the planar and spherical representation of the network can result in different SRLG lists even in case of networks having a geographic extension as small as 100`km`.

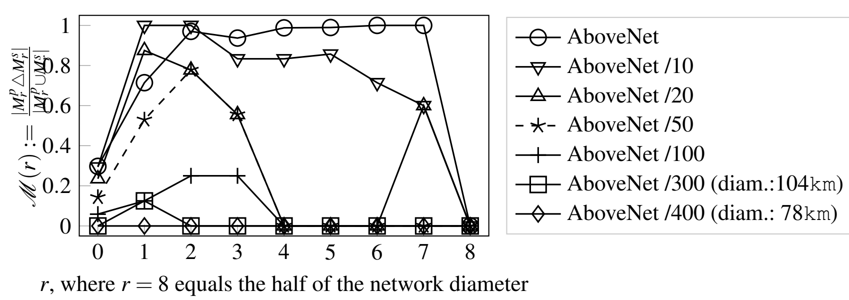


Fig. 8: The ratio of those SRLGs which are different in M_r^p and M_r^s, i.e. $|M_r^p \triangle M_r^s|/|M_r^p \cup M_r^s|$.

Regarding to calculating list M_r^s of maximal SRLGs of disasters represented as circular discs with radius r on a sphere, basically the same ideas could be repeated on the sphere as we have seen earlier. While considering a more general model, where links of networks are represented as polygonal chains consisting of at most γ line segments between their endpoints (where γ is a parameter), paper [39] presents an approach similar to the one seen in Subsubsec. 3.1.3 for determining M_r^s. However, as it only aims to present that 'planar' approaches can be repeated on

[8] Under certain conditions, the complexity of [30] is optimal.

the sphere, it provides a moderately sophisticated algorithm and complexity bound on determining M_r^s. For the sake of complexity analyis, an additional parameter λ is defined, which is the maximal length of the list of suspected maximal failures M while collecting the maximal failures.

According to Cor. 9 of [39], if both x and λ is $O(n)$, and γ is bounded by a constant, the list of maximal link sets which can be hit by a circular disk on the sphere M_r^s can be computed in $O(n^4\rho_r)$. Simulation results show that ρ_r is proportional to $\frac{2r}{diam}$ in the interval $(0, diam/2]$, where *diam* is the geometric diameter of the network. This means an $O(n^4\frac{r}{\mathrm{diam}})$ total running time in practice.

3.2 Approximate Polynomial Algorithms Listing SRLGs

We could see in Subsubsec. 3.1.3 a part of the sophisticated theory and relatively complex algorithms which have to be built in order to be able to provide an algorithm for determining just a single kind of regional SRLG list. This raises the question if one could approach the problem better, or at least more general. As we will see in this Subsection, the answer is yes. In a sense, one of the aims of paper [39] is to show that while there is a struggle for fast algorithms determining basically any kind of SRLG list *precisely*, with relatively low effort one can design discretized approaches which can make small mistakes, but which might be permissible given the uncertainty in the failure modelling and the network data. We note that in [38], links are represented as polygonal chains (or chains of geodesics) between their endpoints, allowing to represent real topologies accurately.

For a point P (in the plane or on the sphere) and node $v \in V$, let the node-distance couple be $[v, d(v,P)]$, where $d(v,P)$ is the distance of v to P. Let $v(P)$ be the list of node-distance pairs of all nodes $v \in V$. We define $e(P)$ to be the list of edge-distance pairs defined similarly. It can be proved that for a given point P, $v(P)$ can be computed in $O(n)$, and $e(P)$ in $O(n+x)$ (where x is the number of edge crossings).

The plan is to determine these lists for enough points which are also placed well enough to be able to determine the maximal SRLG lists based on these node-distance and edge-distance lists. Let $\mathscr{P}$ denote the set of points P for which we want to construct $v(P)$ and $e(P)$.

Let us restrict ourselves to planar geometry for a moment. Intuitively, we can calculate M_r by including the grid points of a sufficiently fine grid (let's say containing 1 km $\times$ 1 km squares) in $\mathscr{P}$. On a sphere, we should choose a similar nice covering.

Algorithm 2 is an example discretized algorithm for determining M_r^g (where 'g' stands for *geometry type*, M_r^g being M_r or M_r^s, if the geometrical representation is planar or spherical, respectively) for circular disks, which has a complexity of $O(|\mathscr{P}|n\frac{r}{\mathrm{diam}})$ under some practical assumptions, and it has a low-polynomial complexity indifferent of the nature of the problem input [39, Thm. 11, Cor. 13]. We can see that although Alg. 2 is much simpler to implement, it is competitive with much more complex precise Alg. 1 in terms of asymptotical runtime.

Regarding to its accuarcy, 1) let $d_{\mathscr{P}}$ be the maximal distance of any geometric location from the (closed) convex hull of the geometric embedding of graph G to the closest point of set $\mathscr{P}$, i.e. $d_{\mathscr{P}} := \max_{t\in\mathrm{conv}(G)} \min_{p\in\mathscr{P}} \mathrm{dist}(p,t)$, and 2) let us denote the relationship of two (link) sets E_1 and E_2 by $E_1 \sqsupseteq E_2$ if and only if for all $e_2 \in E_2$ there exists an $e_1 \in E_1$, such that $e_1 \supseteq e_2$. Using these notations, $M_r^g \sqsupseteq H_r^g \sqsupseteq M_{r-d_P}^g$, where H_r^g is the output of Alg. 2 [38, Thm. 11]. Based on this, if one wants to protect disasters caused by disks with radius r, it is only needed to run Alg. 2 initializing the radius as $r + d_{\mathscr{P}}$. Furthermore, by choosing $\mathscr{P}$ such that $d_{\mathscr{P}}$ to be small, one can avoid enumerating overprotective SRLGs, more precisely, $\lim_{d_{\mathscr{P}}\to 0} H_r^g = M_r^g$, for any fixed network.

---

[9] Similarly to the precise algorithms, this means that M_r^g is the set of maximal failures among which are already checked, and if $e(P)_{\mathrm{hit}}$ is maximal amongst them, it is added to M_r^g and all $e(P)_{\mathrm{hit}}$'s subsets are eliminated from M_r^g; or if $e(P)_{\mathrm{hit}}$ is not maximal in M', nothing happens.

Algorithm 2: Approximate algorithm for determining the maximal r-range SRLG lists

Input: $\mathscr{G}(V,\mathscr{E})$, r, $\mathscr{P}$, geometry type g, coordinates of nodes and polylines of edges
Output: M_r^g
begin
  for $P \in \mathscr{P}$ **do**
    determine $e(P)_{\text{hit}}$
    if $e(P)_{hit} \neq \emptyset$ **then**
      refresh[9] M_r^g with $e(P)_{\text{hit}}$
  return M_r^g

Regarding non-circular SRLGS, engineering fast precise algorithms for determining SRLG lists for arbitrary disaster shape instead of a circle is not trivial[10], but approximate algorithms similar to the one described for determining M_r^g can be easily designed. In short, while the disk is invariant to rotation, the only additional hardness here is that the different orientations of the fixed shape should be also considered. In other words, one should check the links hit by the shape at every centre point *and* every orientation. Discretizing the possible orientations of the shape can be handled just as discretizing the places of centre points. Based on this idea, [39, Alg. 4.] approximately calculates the list $M_{\texttt{shape}}$ of maximal failures which can be caused by a disaster shape in $O(a|\mathscr{P}|n\frac{r}{\text{diam}})$ under some practical assumptions, where a is the number of orientations of the shape which are considered. Its complexity is low-polynomial indifferent of the problem input, and, in limit, the output is precisely $M_{\texttt{shape}}$, both in the plane and on the sphere [39, Thm. 15, Cor. 17].

3.3 More SRLG Determining Approaches

Almost circular SRLGs:

3.3.1 SRLGs of Spatially-Close Fibers

F. Iqbal et al. [17] proposed to call a pair of fibre *spatially-close* if their distance is at most r', i.e. they can be covered with a circular disk of radius $\frac{r'}{2}$. They propose to define SRLG as a set of fibres where any pair of fibres are spatially-close, in other words, any pair of fibres can be covered with a circular disk of radius $\frac{r'}{2}$, see Fig. 9 as an example of the two SRLG models. The intuition is that r' is a smaller parameter than r, representing those fibres that are close together have a higher probability of failing simultaneously due to regional failures. The high-level idea is to provide an approach that generates SRLGs, not considering failure shapes, but simply considers a threshold r': any fibre pairs with a separation distance smaller or equal to r' are considered spatially close.

In [17], three spatially-close fibre problems are considered: (1) finding all pairs of spatially-close fibre segments, (2) finding all spatially-close intervals of fibre to a set of other fibres, and (3) grouping spatially-close fibres into SRLGs.

Fibres are modelled as non-straight concatenations of fibre segments of irregular lengths. Each of these fibre segments is a straight line connecting two fibre-points of known geodetic coordinates (latitude and longitude). For easier calculations, the coordinates are projected to two-dimensional Cartesian coordinates, embedding the fibres into the 2-D plane.

[10] [35] tackles a similar problem.

(a) There are 3 SRLGs each are a pair of links. (b) There is a single SRLG with three links

Fig. 9: Example of SRLG according to (a) Def. 1, and (b) Sec. 3.3.1.

Problem 1 (Detection of Spatially-Close Fiber Segments (DSCFS) [17]). Given a set E of fibers and a distance r'. Each fiber $e \in E$ is associated with a set $\mathscr{T}_e$ of T_e fiber segments. Each fiber segment $t \in \mathscr{T}_e$ is associated with two fiber points (u_{t1}, v_{t1}) and (u_{t2}, v_{t2}).

Find all fibre segment pairs of different fibres that have a minimum separation distance of at most r'.

Clearly, the DSCFS problem is solvable in polynomial time, as the naive approach (computing the separation distance of all fibre segments) has a time complexity of $O(|E|^2T^2)$, where T is the maximum number of fibre segments per fibre. In practice, the runtime can be reduced significantly by first storing each segment in an R-tree[11] [17].

The probability that two spatially-close fibres fail simultaneously depends on the length of the interval(s) of the fibres that are close together.

Problem 2 (Intervals to a Set of Spatially-Close Fibers (ISSCF) Problem [17]). Given a fiber e_i, a set $\mathscr{Y}$ of Y fibers, and a distance r'. Each fiber e_i or $e_j \in \mathscr{Y}$ is associated with a set $\mathscr{T}_i/\mathscr{T}_j$ of T_i/T_j fiber segments. Each fiber segment $t \in \mathscr{T}_i/\mathscr{T}_j$ is associated with two fiber points (u_{t1}, v_{t1}) and (u_{t2}, v_{t2}).

Find the intervals of fiber e_i that have a minimum separation distance of at most r' to any fiber $e_j \in \mathscr{Y}$.

This problem can be solved in $O(YT^2)$ time, where T is the maximum number of fibre segments per fibre, by first finding all fibre segments of Y that are spatially-close to e_i and then computing the spatially-close intervals to these segments by solving sets of quadratic equations (see Alg. 3 in [17]).

Finally, if a set of fibres are grouped into an SRLG if every pair of fibres are spatially-close to each other:

Problem 3 (Grouping of Spatially-Close Fibers (GSCF) Problem [17]). Given a set $\mathscr{F}$ of F spatially-close fibre pairs. Group all fibres that are spatially close to each other, *such that the number of distinct SRLGs is minimized*.

In other words, we want to find all maximal SRLGs, where a maximal SRLG is a set of fibres that are spatially close to every other fibre in the set, and which is not a subset of any other SRLG.

In [17] a heuristic algorithm was given that first transforms it to the Maximal Clique Enumeration (MCE) problem. Second, a variant of the Bron-Kerbosch algorithm [33] to find all maximal cliques is used to find all maximal SRLGs. Note that MCE is an NP-hard problem in general.

[11] An R-tree is an efficient tree data-structure for storing spatial objects. Objects are grouped based on their minimum bounding rectangle.

3.3.2 A Single Worst SRLG in case of a Fixed Disaster Shape and Metric

[23] presents three flavors of problems for finding a most vulnerable place of the network in case of multiple network vulnerability measures[12]. The first problem assumes that the network is bipartite in the topological and geographic sense and that the cuts are vertical line segments. In the latter two problems network, links can be in almost arbitrary locations on the plane. In one of the problems, the disaster shapes are line segments in any direction. In the other, the disasters are circular disks with a given range. To solve the problem instances, in [23], both MILP formulations and polynomial algorithms are given.

We note that in the natural condition when, for any link set $E_1 \subseteq E_2$, the failure of E_2 is *worse* than the failure of E_1 according to the vulnerability measure, the worst SRLG will be part of the set of exclusion-wise SRLGs generated by disaster shape $\mathscr{F}$, which can be determined techniques depicted in Subsec. 3.1 and 3.2. Thus the worst case cut can be found via calculating the measure value to all the maximal SRLGs.

4 Algorithms Determining Lists of PSRLGs

4.1 Computing Lists of FPs and CFPs

4.1.1 Computing FPs From Disaster Sets

Most algorithms for analyzing the vulnerability of networks to disasters, or for creating regional SLRG or FP lists, assume the regional failure takes a fixed shape everywhere in the network area. In reality, the affected region greatly depends on the properties of the disaster, as well as those of the surrounding area. For example, the region affected by an earthquake depends on the earthquake's magnitude, as well as the properties of the rocks and sediments that the earthquake waves travel through. Thus, it makes sense to base SLRG groups on a variety of possible failure shapes.

J. Oostenbrink and F. Kuipers [24] proposed computing the vulnerability of a network to a set of representative disasters $\mathscr{D}$ (each of any shape), instead of to a fixed disaster shape. Each disaster $D \in \mathscr{D}$ is assigned a disaster area $D \subseteq \mathbf{R}^2$, and an occurrence probability p_D. As the probability of simultaneous disasters is low (ignoring strongly correlated events such as aftershocks, which can be combined into a single composite disaster), it is assumed exactly one disaster will occur, i.e. $\sum_{D \in \mathscr{D}} p_D = 1$. Furthermore, it is assumed that if disaster D occurs, all links intersecting its disaster area will fail. A disaster D can take any shape and does not have to be connected, as long as it is possible to compute if a line segment intersects it.

A representative disaster set $\mathscr{D}$ can be obtained in a variety of ways, preferably in collaboration with experts (e.g. seismologists). For example, one can use a tool to randomly generate sets of possible disasters, use the last N historical disasters, or construct a set of custom disaster-scenarios. The concept of a set of representative disasters is similar to that of the Stochastic Event Set used in Cat-modeling, for which a large number of models exist. In all cases, it is necessary to convert hazard intensity values such as ground motions to a binary area D using some threshold function. In [24], J. Oostenbrink and F. Kuipers give an example of converting earthquake scenarios to a disaster set $\mathscr{D}$.

---

[12] The investigated measures are: 1) the total expected capacity of the intersected links, 2) the fraction of pairs of nodes that remain connected, 3) the maximum flow between a given pair of nodes, 4) the average value of maximum flow between all pairs of nodes.

Note that if the disaster locations and shapes are both finite discrete random values (e.g. a division of the plane into grid points), we can generate a finite disaster set $\mathscr{D}$ of all possible disasters by simply adding each possible combination of disaster location and shape to $\mathscr{D}$.

Let a failure state s be defined as a set $s \subseteq E$, where $e \in s$ if and only if e has failed. Now, the failure set $S(D)$ of links that are affected by disaster D is the set of links that intersect D. Note that this definition of $S(D)$ is equivalent to that of a regional SLRG, $SRLG(D)$.

Let $\mathscr{S}$ be the random value indicating the failure state after the disaster. Given a disaster set $\mathscr{D}$, we can obtain the distribution of $\mathscr{S}$ as follows [24]:

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

We now have each possible failures state, as well at its occurrence probability. That is, we have the complete list of FPs, based on the disaster set $\mathscr{D}$.

4.1.2 Calculating Lists of CFPs Based on Correlated Link Failures

A study dealing with the probabilities of correlated link failures is [32], which models the regional failures as having a random epicentre, and a random size (described with a size parameter s in $[0,1]$). Two assumptions are made: 1) in the investigated time period there is at most one disaster and 2) for every possible failure epicentre and failure sizes $s_1 < s_2$ the region destroyed by disaster with size s_1 is totally contained by the region hit by the one with size s_2.

Fig. 10 briefly depicts the model. It shows an example network and the corresponding failure probabilities. Suppose we need to establish a high-availability connection from the top node through the 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.

Now an implementation strategy follows which 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 $\times$ 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"[13] for every p inside each grid cell c. This way the whole integration problem translates 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 neighbouring c.

Let r denote the absolute maximum range of a disaster in km.

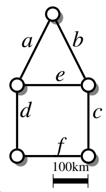
Let (x_{min}, y_{min}) be the bottom left corner and (x_{max}, y_{max}) the top right corner of a rectangular area in which the network lies. It is sufficient to process each c in the rectangle of bottom left corner $(x_{min}-r, y_{min}-r)$ and top right corner $(x_{max}+r, y_{max}+r)$, and we denote 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. $h_{i,j} = \int_{p \in c_{i,j}} h(p) \mathrm{d}p$.

The query time of sets can be reduced to a constant with very high probability (with the help of hashing) if we store all CFPs.

[13] 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 $f(e,c)$.

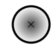
Input:

Network:



Failure model: Type: tornado | earthquake | EMP | ...

Model parameters:
p_d: the probability of a disaster of type d in the given area and time period;
$h(p)$: quantitative hazard map of the area, that is the probability density function of the location of the disaster epicenter (e.g. uniform distribution on a bounded area on $\mathbb{R}^2$);
$r(p,s)$: the shape function of the disaster depending on epicenter p and size s returning the damaged zone of the disaster (e.g. a circular disk centered on p with radius s);
Regional failure model:
Hazard epicenter: random variable on $\mathbb{R}^2$ with probability density $h()$.
Relative size: random variable uniform distribution on $[0,1]$. Each link fails having a point in the disaster area defined by shape function $r()$, the rest remain intact.



Output:

$P(a)=.0055$ $P(b)=.0055$ $P(c)=.005$ $P(d)=.005$ $P(e)=.005$ $P(f)=.005$
$P(a,b)=.00068$ $P(b,e)=.00064$ $P(a,e)=.00064$ $P(c,e)=.00056$
$P(d,e)=.00056$ $P(d,f)=.00056$ $P(c,f)=.00056$ $P(c,b)=.00052$
$P(a,d)=.00052$ $P(a,e,d)=.00031$ $P(b,e,c)=.00031$ $P(a,b,e)=0$

Fig. 10: An illustration of the CFP problem inputs and outputs.

Using self-balancing binary trees, its worst-case query time is always $O(\rho\log((n+x)\rho))$.

To build up the list of CFPs, we use an associative array $\mathscr{J}$, which can be addressed by an (unordered) set of links $\{e_1,e_2,\dots,e_k\}$ 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 $\mathscr{S}_{i,j}=(e_1,e_2,\dots,e_k)$, and increment the $\mathscr{J}$ values accordingly, i.e. $\mathscr{J}[\{e_1\}]+=h_{i,j}\cdot f(e_1,c_{i,j})$, $\mathscr{J}[\{e_2\}]+=h_{i,j}\cdot f(e_2,c_{i,j})$, $\mathscr{J}[\{e_1,e_2\}]+=h_{i,j}\cdot f(e_2,c_{i,j})$, etc.

For the probability p_S of failing exactly the set S of links, we need to look up S in $\mathscr{J}$. If not found, then $P_S=0$.

The drawback of the CFP list is that it has an $\Omega(2^\rho)$ space complexity, which makes it inefficient for bigger network densities. With this in mind, one can build up also a list of FPs representing the same disasters, which will be significantly shorter, but some precomputations will be needed to determine P_S.

4.1.3 CFPs and FPs from Historical Earthquake Data

Intuitively, the models presented in Subsubsec. 4.1.1 and 4.1.2 are related. In fact, both models could be used for computing lists of FPs and CFPs. In [36], an approach for determining the list of CFPs and FPs based on the available historical Earthquake data is presented. This approach can be viewed as a special case of both models presented in Subsubsec. 4.1.1 and 4.1.2.

Namely, the next possible earthquake has a random epicentre taken from a set of grid points over the evaluated area, and the disaster area has also a random range (which is a function of the earthquake moment magnitude) taken from a discretized scale. This results in a set of possible disaster scenarios with some probabilities as in Subsubsec. 4.1.1, and also can be viewed as a discrete version of model in Subsubsec. 4.1.2.

4.2 Probabilistic Modelling of the Worst Place of a Disaster

Similarly to [23] (in Subsubsec. 3.3.2), [3] aims to find the single worst place of a disaster under a certain metric. However, while the first models the disaster effect to be deterministic (every network element which has an intersection with the disaster area fails with probability 1), in the latter every link has a probability $\in [0,1]$ of failing in case of each disaster place. However, an incompleteness of the paper is that, in case of a fixed disaster, it considers that the affected links fail independently of each other.

To be more precise, the model of [3] is the following. They define a failure probability distribution function $f : Q \times \mathbb{R}^2 \to R \geq 0$. Given a disaster location $P \in \mathbb{R}^2$ and $q \in Q$, $f(q,p) = f_q(p)$ is the probability that q is affected by the disaster at p. For a compound component π composed of a sequence of simple components $q_1,\dots,q_r$, the probability of being damaged by a disaster at a location p is denoted as $f_\pi(p)$ and being defined to be the probability that at least one of its simple component if damaged, i.e., $f_\pi(p) = 1 - \prod_{q\in\pi}(1 - f_q(p))$.

For finding the most vulnerable point according to various metrics (expected component damage, average two-terminal availability, and expected maximum post-attack flow), [3] presents Las Vegas and Monte Carlo algorithms. It also offers approximate solutions to the problem of finding the worst arrangement of k simulaneous disasters (attacks), which is a generalization of the NP-hard maximum set cover problem[16].

4.3 On Two-Stage PSRLGs and Denomination Issues

The first paper considering probabilistic SRLGs was [20]. There the structure which in this chapter is called 'two-stage PSRLG' is named simply as 'Probabilistic SRLG' (PSRLG). Since we felt that FPs and CFPs deserve to be called probabilistic SRLGs at least as much as the structure defined in [20], we decided to call these three structures collectively as PSRLGs, and name the [20]-PSRLG as 'two-stage PSRLG'.

Due to this historical reason, we believe it worth presenting its model even though it does not tackle the question of calculating PSRLGs (it only uses PSRLGs as inputs for a diverse routing problem). [20] defines the two-stage PSRLGs as follows. There is a set R of SRLG events that can incur link failures. Each SRLG event $r \in R$ occurs with probability π_r, and once an SRLG event r occurs, link (i,j) will fail with probability $p^r_{i,j} \in [0,1]$. Link (i,j) is part of the resulting (two-stage) PSRLG if $p^r_{i,j} > 0$. This definition from [20] is slightly generalized in Def. 4 while keeping the form of the data structure.

As Fig. 1 also suggests, using lists of two-stage PSRLGs one could store the same information more compactly as in lists of FPs or CFPs. However, there are numerous open questions related to this field, as to the best of our knowledge, no paper investigated how to enumerate in an efficient way lists of two-stage PSRLGs.

5 SRLG Lists Obtained from PSRLG Lists

It is a natural idea to list the (maximal) link sets which have a probability of failing together higher than a given threshold T (like in [27]). Obviously, for this, as an intermediate step, one has to generate a set of probabilistic SRLGs. More precisely, CFP is the most useful structure in this context, since, by definition, for a link set S, the Cumulative Failure Probability P_S is the probability that at least the links of S will fail. The advantage of this approach is that SRLG lists can be generated based on sophisticated objectives. Alg. 3 sketches this framework.

Algorithm 3: SRLG list obtained from CFP list

Input: graph $G(V,E)$, threshold $T \in [0,1]$, CFP model C (e.g. as in one of [24, 26, 27, 32]), additional parameters needed for C

Output: List M_T of maximal link sets having a CFP at least T

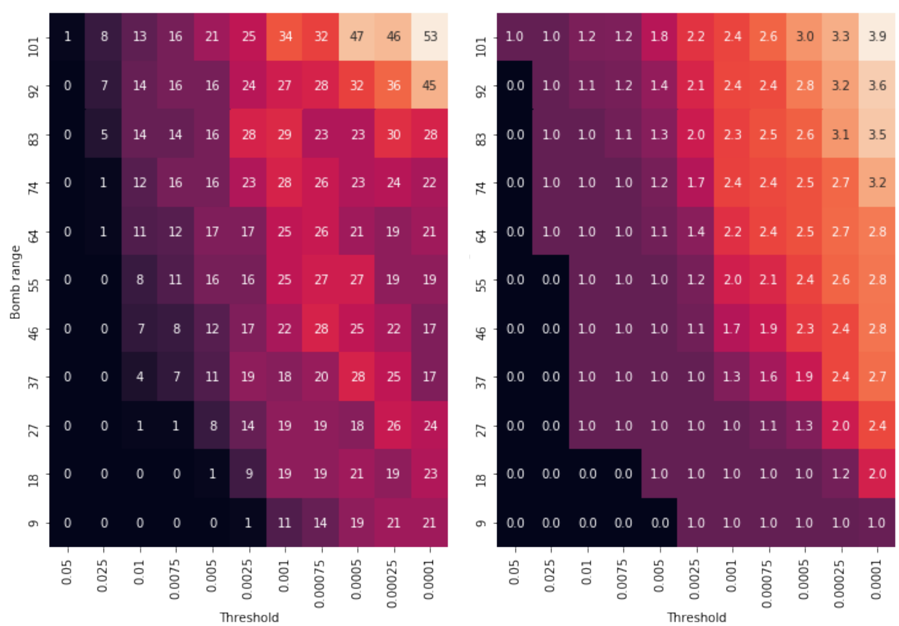
Calculate list L of CFPs according to C

Collect CFPs of L with probability $\geq T$ in list F_T

return list M_T of maximal elements of F_T

As a concrete example, [27] modifies the CFP enumerating model presented in [32] (and Subsubsec. 4.1.2) in order to take in count also the availabilities of the links. Compared to [32], a link e with low availability makes CFPs it is involved in to have higher probabilities, while reliable links decrease these probabilities. Fig. 11 shows the cardinality of $|M_T|$ and avg. length of SRLGs in M_T in function of threshold T and maximal disaster area R in case of backbone topology `16_optic_pan_eu`[25]. Note that the unit of R is not a `km`, as the Euclidean distances are altered during the CFP enumerating process in function of the availabilities.

The first observation made in Fig. 11 is that a radius $R \geq 80$ (which roughly corresponds to the 20% of the network diameter) or larger combined with a threshold $T \leq 0.001$ yields a high number of maximal probable failures. This translates to the fact that a bigger disaster possibly hits a larger number of edges, and the failures above the small threshold cannot be dominated by only a few sets



(a) $|M_T|$ (b) Average SRLG length in M_T

Fig. 11: Cardinality of $|M_T|$ and avg. length of SRLGs in M_T in function of T and R in case of backbone topology `16_optic_pan_eu` [25]

from M_T. Of course, in a non-practical extreme case of R being greater than half of the network diameter, it is possible that $M_T = \{E\}$, meaning $|M_T| = 1$.

Further observations of [27] are that: (i) if $R \in [0,80]$, M_T is likely to contain only a handful of most probable SRLGs; (ii) similar $R \cdot T$ value indicates similar cardinality of $\mathscr{M}_T$. Hence, for reasonable disaster sizes, M_T has a manageable size, with its cardinality being comparable with the number of network elements. In addition, one can observe that the average size of the SRLGs scales with the disaster radius.

6 Conclusion

In this chapter, we overviewed the state of the art algorithms for enumerating *regional Shared Risk Link Groups (SRLGs)* and *regional Probabilistic SRLGs (PSRLGs)*, which structures are key in translating the composed geometric problem of protecting telecommunication networks against regional failures to purely combinatorial and probabilistic problems, respectively. We showed that the best technique to choose for enumerating the vulnerable regions varies on 1) the available geometric information on the network topology, 2) (probabilistic) information on the effects of possible disasters in the network area, and 3) the desired output structure (SRLG/PSRLG). In the chapter, first we presented a range of deterministic approaches for enumerating maximal regional SRLGs under various conditions. Then, for regional PSRLG enumeration, we visited some models, which are easily tunable to the available knowledge on the network topology and the disasters. Lastly, as an advanced technique, we described an SRLG enumerating approach, which uses a probabilistic model in an intermediate step.

Acknowledgements This chapter is based on work from COST Action CA15127 ("Resilient communication services protecting end-user applications from disaster-based failures - RECODIS") supported by COST (European Cooperation in Science and Technology). Part of this work has been supported by 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.

References

1. The internet topology zoo. URL http://www.topology-zoo.org/dataset.html
2. Agarwal, P.K., Efrat, A., Ganjugunte, S.K., Hay, D., Sankararaman, S., Zussman, G.: Network vulnerability to single, multiple, and probabilistic physical attacks. In: Military Communications Conference (MILCOM), pp. 1824–1829. IEEE (2010)
3. Agarwal, P.K., Efrat, A., Ganjugunte, S.K., Hay, D., Sankararaman, S., Zussman, G.: The resilience of WDM networks to probabilistic geographical failures. IEEE/ACM Transactions on Networking (TON) **21**(5), 1525–1538 (2013)
4. Colman-Meixner, C., Dikbiyik, F., Habib, M.F., Tornatore, M., Chuah, C.N., Mukherjee, B.: Disaster-survivable cloud-network mapping. Photonic Network Communications **27**(3), 141–153 (2014)
5. de Sousa, A., Santos, D.: The minimum cost d-geodiverse anycast routing with optimal selection of anycast nodes. In: 2019 15th International Conference on the Design of Reliable Communication Networks (DRCN), pp. 21–28 (2019). DOI 10.1109/DRCN.2019.8713729
6. Dikbiyik, F., Tornatore, M., Mukherjee, B.: Minimizing the risk from disaster failures in optical backbone networks. Journal of Lightwave Technology **32**(18), 3175–3183 (2014)

7. Dinh, T.N., Thai, M.T.: Network under joint node and link attacks: Vulnerability assessment methods and analysis. IEEE/ACM Transactions on Networking **23**(3), 1001–1011 (2015)
8. Gardner, M.T., Beard, C.: Evaluating geographic vulnerabilities in networks. In: IEEE Int. Workshop Technical Committee on Communications Quality and Reliability (CQR), pp. 1–6 (2011)
9. Gerstel, O., Jinno, M., Lord, A., Yoo, S.B.: Elastic optical networking: A new dawn for the optical layer? Communications Magazine, IEEE **50**(2), s12–s20 (2012)
10. Gomes, T., Tapolcai, J., Esposito, C., Hutchison, D., Kuipers, F., Rak, J., de Sousa, A., Iossifides, A., Travanca, R., André, J., Jorge, L., Martins, L., Ugalde, P.O., Pašić, A., Pezaros, D., Jouet, S., Secci, S., Tornatore, M.: A survey of strategies for communication networks to protect against large-scale natural disasters. In: 8th International Workshop on Resilient Networks Design and Modeling (RNDM), pp. 11–22 (2016). DOI 10.1109/RNDM.2016.7608263
11. Habib, M.F., Tornatore, M., De Leenheer, M., Dikbiyik, F., Mukherjee, B.: Design of disaster-resilient optical datacenter networks. Journal of Lightwave Technology **30**(16), 2563–2573 (2012)
12. Habib, M.F., Tornatore, M., Mukherjee, B.: Fault-tolerant virtual network mapping to provide content connectivity in optical networks. In: Optical Fiber Communication Conference, pp. OTh3E–4. OSA (2013)
13. Harter, I.B.B., Schupke, D., Hoffmann, M., Carle, G., et al.: Network virtualization for disaster resilience of cloud services. Communications Magazine, IEEE **52**(12), 88–95 (2014)
14. Heidemann, J., Quan, L., Pradkin, Y.: A preliminary analysis of network outages during hurricane Sandy. University of Southern California, Information Sciences Institute (2012)
15. Hmaity, A., Musumeci, F., Tornatore, M.: Survivable virtual network mapping to provide content connectivity against double-link failures. In: Design of Reliable Communication Networks (DRCN), pp. 160–166. IEEE (2016)
16. Hochbaum, D., Pathria, A.: Analysis of the greedy approach in problems of maximum k-coverage. In: Naval Research Logistics (NRL) (1998)
17. Iqbal, F., Trajanovski, S., Kuipers, F.: Detection of spatially-close fiber segments in optical networks. In: 2016 12th International Conference on the Design of Reliable Communication Networks (DRCN), Drcn, pp. 95–102. IEEE (2016). DOI 10.1109/DRCN.2016.7470840. URL http://ieeexplore.ieee.org/document/7470840/
18. Jaumard, B., Hoang, H.A.: Design and dimensioning of logical survivable topologies against multiple failures. Journal of Optical Communications and Networking **5**(1), 23–36 (2013)
19. Kushwaha, A., Kapadia, D., Gumaste, A., Somani, A.: Designing multi-layer provider networks for circular disc failures. In: International Conference on Optical Network Design and Modeling (ONDM). Dublin, Ireland (2018)
20. Lee, H.W., Modiano, E., Lee, K.: Diverse routing in networks with probabilistic failures. IEEE/ACM Transactions on Networking **18**(6), 1895–1907 (2010)
21. Long, X., Tipper, D., Gomes, T.: Measuring the survivability of networks to geographic correlated failures. Optical Switching and Networking **14**, 117–133 (2014)
22. Mukherjee, B., Habib, M., Dikbiyik, F.: Network adaptability from disaster disruptions and cascading failures. Communications Magazine **52**(5), 230–238 (2014)
23. Neumayer, S., Zussman, G., Cohen, R., Modiano, E.: Assessing the vulnerability of the fiber infrastructure to disasters. IEEE/ACM Transactions on Networking (TON) **19**(6), 1610–1623 (2011)
24. Oostenbrink, J., Kuipers, F.: Computing the impact of disasters on networks. ACM SIGMETRICS Performance Evaluation Review **45**(2), 107–110 (2017)
25. Orlowski, S., Wessäly, R., Pióro, M., Tomaszewski, A.: Sndlib 1.0: survivable network design library. Networks **55**(3), 276–286 (2010)
26. Pašić, A., Girao-Silva, R., Vass, B., Gomes, T., Babarczi, P.: FRADIR: A novel framework for disaster resilience. In: Int. Workshop on Resilient Networks Design and Modeling (RNDM). Longyearbyen (Svalbard), Norway (2018)
27. Pašić, A., Girao-Silva, R., Vass, B., Gomes, T., Mogyorósi, F., Babarczi, P., Tapolcai, J.: FRADIR-II: An improved framework for disaster resilience. In: Int. Workshop on Resilient Networks Design and Modeling (RNDM). Nicosia, Cyprus (2019)

28. Souza Couto, R., Secci, S., Mitre Campista, M., Costa, K., Maciel, L.: Network design requirements for disaster resilience in iaas clouds. Communications Magazine, IEEE **52**(10), 52–58 (2014)
29. Su, T.H., Chang, R.C.: The *k*-Gabriel graphs and their applications. In: Algorithms. Lecture Notes in Computer Science, vol.450. Berlin, Heidelberg (1990)
30. Tapolcai, J., Rónyai, L., Vass, B., Gyimóthi, L.: Fast enumeration of regional link failures caused by disasters with limited size. In: Technical report. URL http://lendulet.tmit.bme.hu/lendulet_website/wp-content/papercite-data/pdf/tapolcai2019fast.pdf
31. Tapolcai, J., Rónyai, L., Vass, B., Gyimóthi, L.: List of shared risk link groups representing regional failures with limited size. In: Proc. IEEE INFOCOM. Atlanta, USA (2017)
32. Tapolcai, J., Vass, B., Heszberger, Z., Biró, J., Hay, D., Kuipers, F.A., Rónyai, L.: A tractable stochastic model of correlated link failures caused by disasters. In: Proc. IEEE INFOCOM. Honolulu, USA (2018)
33. Tomita, E., Tanaka, A., Takahashi, H.: The worst-case time complexity for generating all maximal cliques and computational experiments. Theoretical Computer Science **363**(1), 28–42 (2006)
34. Trajanovski, S., Kuipers, F., Van Mieghem, P., et al.: Finding critical regions in a network. In: IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 223–228. IEEE (2013)
35. Trajanovski, S., Kuipers, F.A., Ilić, A., Crowcroft, J., Van Mieghem, P.: Finding critical regions and region-disjoint paths in a network. IEEE/ACM Transactions on Networking (TON) **23**(3), 908–921 (2015)
36. Valentini, A., Vass, B., Oostenbrink, J., Csák, L., Kuipers, F.A., Pace, B., Hay, D., Tapolcai, J.: Network resiliency against earthquakes. In: Int. Workshop on Resilient Networks Design and Modeling (RNDM). Nicosia, Cyprus (2019)
37. Vass, B., Bérczi-Kovács, E., Tapolcai, J.: Enumerating shared risk link groups of circular disk failures hitting k nodes. In: Proc. International Workshop on Design Of Reliable Communication Networks (DRCN). Munich, Germany (2017)
38. Vass, B., Kovács, E., Tapolcai, J.: Enumerating maximal shared risk link groups of circular disk failures hitting k nodes. In: Technical report. URL http://lendulet.tmit.bme.hu/lendulet_website/wp-content/papercite-data/pdf/vass2019enumerating.pdf
39. Vass, B., Németh, L., Tapolcai, J.: The Earth is nearly flat: Precise and approximate algorithms for detecting vulnerable regions of networks in plane and on sphere. In: Networks, Wiley (2019)
40. Veremyev, A., Boginski, V., Pasiliao, E.L.: Exact identification of critical nodes in sparse networks via new compact formulations. Optimization Letters **8**(4), 1245–1259 (2014)
41. Xie, A., Wang, X., Lu, S.: Risk minimization routing against geographically correlated failures. IEEE Access **7**, 62,920–62,929 (2019). DOI 10.1109/ACCESS.2019.2916834
42. Yallouz, J., Rottenstreich, O., Orda, A.: Tunable survivable spanning trees. ACM SIGMETRICS Performance Evaluation Review **42**(1), 315–327 (2014)
43. Zang, H., Ou, C., Mukherjee, B.: Path-protection routing and wavelength assignment (RWA) in WDM mesh networks under duct-layer constraints. IEEE/ACM Transactions on Networking (TON) **11**(2), 248–258 (2003)