Evaluating Network Reduction Strategies for Consistent Risk Assessment of Critical Infrastructures

Mauricio Monsalve

School of Engineering, Pontificia Universidad Católica de Chile, and Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, Chile. E-mail: mauricio.monsalve@cigiden.cl

Elisa Ferrario

School of Engineering, Pontificia Universidad Católica de Chile, and Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, Chile. E-mail: elisa.ferrario@cigiden.cl

Yolanda Alberto

Department of Civil Engineering, University of Chile, and Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, Chile. E-mail: yalberto@uchile.cl

Felipe Arróspide

Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, Chile. E-mail: faarrosp@uc.cl

Sebastián Castro

Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, Chile. E-mail: srcastro@uc.cl

Alan Poulos

Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, Chile. E-mail: alan.poulos@cigiden.cl

Juan Carlos de la Llera

Department of Structural and Geotechnical Engineering, Pontificia Universidad Católica de Chile, and Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, Chile. E-mail: jcllera@ing.puc.cl

Critical infrastructure networks are continuously growing, gaining complexity with each urban sprawl, conurbation, technological change, and regulatory update. Consequently, their detailed risk analysis demands large amounts of data, computational resources (required by simulations, optimization, flow equilibria, etc.), and dealing with complex interpretations of the results. This comes with several drawbacks: scarcity of adequately curated data, which instead are usually incomplete and sometimes even incorrect, algorithmic runtime that impairs the full use of Monte Carlo simulations, errors that may propagate extensively, and results that cannot be generalized and extended to other cases. Therefore, researchers have also resorted to analyzing simplified versions of these infrastructure systems. This work evaluates three algorithms for reducing the complexity of infrastructure networks while keeping reasonable accuracy for statistical interpretation. These algorithms transform a detailed graph into a more compact representation, where risk assessments can be performed more easily. The strategies used herein are based on the detection of important edges (backbone detection) and the merging or lumping similar or proximate elements (clustering, contractions). The different complexity reduction algorithms are evaluated on three infrastructure networks, namely: the electric transmission network of Chile, the electric distribution network of the Greater Valparaíso and the drinking water distribution network of the Greater Valparaíso. The experiments show that two of the three graph reduction criteria proposed in this work yield good approximations of the connectivity of the original graphs, when these are reduced to 25% of their size.

Keywords: Critical infrastructure, network reduction, sparsification, coarsening, backbone extraction, clustering.

Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference
Edited by Piero Baraldi, Francesco Di Maio and Enrico Zio

ISBN: 978-981-14-8593-0; doi:10.3850/978-981-14-8593-0_5115-cd

1. Introduction

One of the main themes of research in critical infrastructure systems has been the identification of critical elements and localized attacks (Dunn and Wilkinson, 2013; Cuadra et al., 2015; Marlim et al., 2019). In this theme of research, it is assumed that a malicious entity seeks to cause a large affectation in a system. A typical outcome of this research is the identification of network elements that require additional protection and monitoring. However, this is not the only setting to which infrastructure networks are threatened.

Spatially distributed hazards, such as earthquakes, volcanic eruptions and severe weather, impact networks beyond just their critical elements, by instead "targeting" large portions of their structure. Therefore, critical element analysis needs to be supplemented by other techniques. In this line, probabilistic risk analysis (PRA) appears as a suitable complement to critical elements (Yehia and Swei, 2020; Hernandez-Fajardo and Dueñas-Osorio, 2013).

In PRA, the typical way to proceed is to simulate the behavior of a system over a large number of stochastically-generated scenarios. In each scenario, it is assessed whether system components experience loads beyond their tolerance, which causes these components to underperform or fail, inducing an overall decrease in system performance. This degradation in performance is quantified and serves to characterize risk.

The computational time required by PRA may be considerable, especially when simulating large systems and systems whose faulty behavior is seen only in extreme cases. This is especially problematic when assessments need to be done quickly (e.g. early warning) and when PRA is part of an optimization loop (e.g. risk mitigation investment decisions). Thus, analysts have resorted to techniques such as importance sampling (Jayaram and Baker, 2010), subset simulation (Wang et al., 2019; Zuev et al., 2015) and surrogate models (Xiao et al., 2018) in order to decrease the computational time required to draw statistically significant inferences from PRA.

This work explores the use of graph reduction techniques to speed-up PRA on networked infrastructures. In particular, it proposes and evaluates three graph reductions on the electricity and drinking water infrastructures from central Chile, under stochastically-generated seismic hazard. Furthermore, it proposes a calculation method to estimate the original system affectation using information from the reduced networks.

2. Background

2.1. *Probabilistic seismic risk*

Herein, probabilistic seismic risk is assessed using a Monte Carlo simulation scheme for estimating system affectation under stochastically-generated seismicity.The first step consists in sampling earthquakes according to recurrence relations which regulate the frequency of earthquakes according to their magnitude and spatial locations (Poulos et al., 2018).

The second step consists in estimating intensity measures on site for these earthquakes; this is done by using ground motion prediction equations, which associate earthquake location and magnitude, as well as site parameters (geological and geotechnical), with intensity measures of interest (Montalva et al., 2017; Stewart et al., 2015). Commonly used intensity measures include peak ground acceleration, peak ground velocity, permanent ground deformation, etc.

The third step consists in assessing the physical inventory (roads, buildings, pipes, etc.) after being exposed to these intensity measures, by using fragility functions (Lallemant et al., 2015; Pitilakis et al., 2014). These functions indicate the probability a physical element suffers damage (or experiences a reduction in functionality) given the intensity measures.

The fourth step consists in evaluating the performance of the remaining physical inventory after each seismic event. For lifelines and network infrastructure, this implies evaluating a number of performance measures (Han and Davidson, 2012; Franchin and Cavalieri, 2015; Salman and Li, 2018). A consequential step may also involve post-earthquake service restoration, recovery and reconstruction (Salman and Li, 2018).

The computational burden of this framework resides in the fourth step, in which system simulation is performed. Thus, reducing the computational resources required to assess infrastructure system performance is relevant, especially in the case of large scale systems.

2.2. *Graph reduction*

Graph reductions have been used to analyze very large graphs in affordable computational time and space (memory). Several approaches have been used to simplify graphs. Edge contraction, for instance, consists in eliminating an edge (link) from a graph and then merging the vertices connected by the removed edge (Van et al., 2004; Salzman et al., 2014). Different criteria have been devised to choose which edges are to be contracted. According to the criterion used, it is possible to arrive at a clustering procedure (Gdalyahu et al., 1999; Delling et al., 2009; Meyerhenke et al., 2014).

Clustering or community detection itself has also been used to simplify graphs (Wu et al., 2004; Qiu and Hancock, 2007; Liu et al., 2015). In this case, identified clusters or communities are replaced by vertices in the graph.

Graph reduction can also be achieved by the removal of edges without merging vertices. In this case, it is called graph sparsification, because a graph of lesser density is obtained from this process (Satuluri et al., 2011; Spielman and Srivastava, 2011; Ahn et al., 2012). Particularly popular are the edge filtering procedures, in which "low importance" edges are removed from the graph (Serrano et al., 2009; Hamann et al., 2016; Feng, 2018). Various scores have been devised to achieve "good quality" sparsified graphs according to different criteria.

A task related to graph sparsification is that of backbone extraction. The backbone of a network is the subset of edges that best represents the connections of the network (Glattfelder and Battiston, 2009; Serrano et al., 2009; Foti et al., 2011). Backbones can be identified by filtering edges or by using spanning trees, i.e., subgraphs of the network that cover all vertices but have no loops (Scellato et al., 2006; Du et al., 2007).

A typical application of graph reduction takes place within the graph coarsening procedures. Graph coarsening is a strategy used for simplifying computation with large graphs by using smaller surrogate graphs obtained from graph reduction strategies (Hendrickson and Leland, 1995; Gupta, 1997; Ron et al., 2011). Computation with the reduced graph is exact. Then, detail is progressively added back to the reduced graph and the computed solution is refined in the presence of this new detail.

Some works have already applied graph reduction methods in critical infrastructure research. Backbone extraction has been applied to transportation networks, in particular in the identification of main routes (Scellato et al., 2006; Jiang, 2007; Dai et al., 2018). Graph coarsening and clustering are often used in the context of network partitioning or zonation of water and electricity supply networks (Di Nardo et al., 2011; Perelman and Ostfeld, 2012; Miao et al., 2013; Di Nardo et al., 2014; Ding et al., 2015; Buluç et al., 2016; Ferrario et al., 2016; Zhang et al., 2017). The choice of an adequate zonation facilitates efficient management and maintenance of infrastructures.

3. Problem statement

The objective of this work is to evaluate the use of graph reduction methods in the estimation of infrastructure affectation, especially in the context of probabilistic seismic risk assessment. More specifically, this work is concerned with the estimation of the number of elements connected to the main portion of the infrastructure network.

The graph reduction algorithms considered herein combine the ideas of edge contraction and edge filtering in their design. The general idea is to simplify the network by contracting edges in the graphs repeatedly. This results in *thick* vertices and edges that represent several elements. Then, network affectation is computed in the reduced graphs. Fig. 1 shows an overview of this process.

3.1. *Definitions*

Some notation is needed to introduce the algorithms. Let $G = (V, E)$ describe a graph of vertex set V and edge set E. Edges have no *a priori* orientation in this work; therefore, an edge $e \in E$ is described as a set of two vertices, i.e., $e = \{u, v\}$, $u, v \in V$, with $u \neq v$.

An important notion for this work is that of *neighborhood*. Let $N : V \to \mathcal{P}(V)$ be the vertex neighborhood function, where $\mathcal{P}$ stands for the powerset. For vertex $v \in V$, the neighborhood of v, denoted by $N(v)$, is defined as the set of vertices connected with v by an edge, $N(v) = \{u \in V : \{u, v\} \in E\}$.

Related to the neighborhood is the concept of *degree*, which is the number of neighbors a vertex has, i.e., $\delta(v) = |N(v)|$, where $\delta(v)$ stands for the degree of vertex $v \in V$.

The next relevant concept is that of *connectivity*. A *path* is a sequence of adjacent edges, e.g., $\{w_1, w_2\}$, $\{w_2, w_3\}$, ..., $\{w_{l-1}, w_l\} \in E$. Two vertices $u, v \in V$ are said to be *connected* if there is a *path* between them (e.g., $w_1 = u$ and $w_l = v$). Finally, a *connected component*, or simple a *component*, is a collection of connected vertices, and all the edges among them, that is maximal, that is, a component is not contained within another component.

Reduced graph. In this work, the notion of reduced graph refers to the elements of one graph mapped into another graph of fewer vertices and edges. More formally, graph $G' = (V', E')$ is said to be a *κ-reduced graph* of $G = (V, E)$, if $V' = \{\kappa(v) : v \in V\}$ and $E' = \{\{\kappa(u), \kappa(v)\} : \{u, v\} \in E, \kappa(u) \neq \kappa(v)\}$.

Function $\kappa : V \to V'$ maps the elements of the original graph G into the corresponding elements of the reduced graph G'. Similarly, function κ^{-1} can be defined as the preimage of κ; for $v \in V'$, $\kappa^{-1}(v) = \{u \in V : \kappa(u) = v\}$. The algorithms proposed in this work are different ways for defining κ and, thus, inducing κ-reduced graphs.

3.2. *π-connectivity*

When exposed to seismic events, the infrastructure systems considered may lose the functionality of some of its elements, leading to a reduction in total functionality and network connectivity.

Let π be a random function such that $\pi(v)$ indicates whether vertex $v \in V$ is functional and $\pi(e)$ indicates whether edge $e \in E$ is functional. Random function π represents the physical performance of the network elements under a scenario. The concept of *π-connectivity* refers to the connectivity of elements which are functional according to π. Two vertices $u, v \in V$

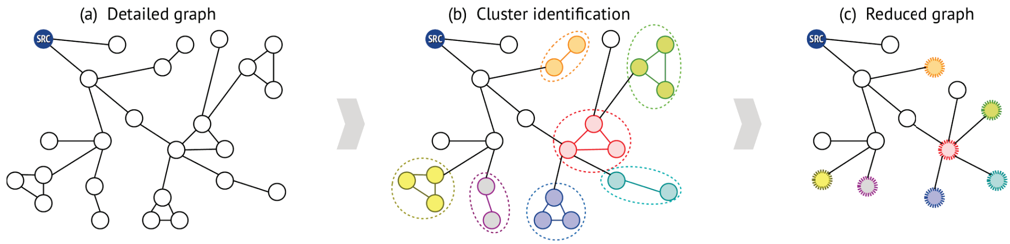

Fig. 1. Graph reduction scheme used in this work. Insets: (a) original, detailed graph; (b) classification of vertices for later merging; and (c) reduced graph obtained after applying the contractions.

are said to be π-connected if there is a path of edges $\{u, w_1\}$, $\{w_1, w_2\}$, ..., $\{w_l, v\}$ such that all its edges are functional, i.e., $\pi(\{w_i, w_{i+1}\})$, $\pi(\{u, w_1\})$, $\pi(\{w_l, v\})$, and all its vertices are also functional, i.e., $\pi(u)$, $\pi(v)$, $\pi(w_i)$.

The definition of random function π and π-connectivity has clear physical meaning in the original graph $G = (V, E)$. It essentially refers to the evaluation of the fragility curves of the elements on stochastically generated seismicity. The same does not apply to the κ-reduced graph, $G' = (V', E')$, however.

The aim of this work is to estimate π-connectivity without using the original graph $G = (V, E)$, but by instead using its κ-reduced graph $G' = (V', E')$. To do this, *equivalent* fragility curves are derived for the elements in G'. Moreover, this work only measures network performance as π-connectivity to the *source* vertices, which provide the functionality the network distributes. For example, a water reservoir would be a source in a water distribution network.

Assuming contracted elements have similar exposure and fragily, vertices in V' are associated with two fragility curves: one for the contracted vertices, p_v, being the average of the curves, and another for the contracted edges, p_e, computed the same. Edges in E' are associated with the average fragility curve of the comprised edges, q.

Having defined p_v, p_e, and q, the estimation of π-connectivity with G' is described next. Element survival is assessed when visited, using an *all-or-nothing* approach. When visited, an edge $e \in E'$ fails with probability $q(e)$, in which case, all of the contained edges fail. Else, none do so. Similarly, when vertex $v \in V'$ fails, all its elements do so, else they all survive. This is set to occur with probability

$$\begin{aligned} p(v) &= 1 - (1 - p_v(v))(1 - p_e(v)) \\ &= p_v + p_e - p_v p_e. \end{aligned} \tag{1}$$

This is not mathematically accurate, but it is motivated by the following heuristic rationale: the thick vertex survives $(1 - p(v))$ if, in average, the vertices survive $(1 - p_v(v))$ and the edges survive $(1 - p_e(v))$. The development of better formulas will be the subject of future work.

3.3. *Proposed algorithms*

The common scheme of the graph reduction algorithms is the following:

(i) Put all edges $e \in E$ in list L and sort them by the *selected score*.
(ii) Define set $S = \{\{v\} : v \in V\}$.
(iii) Pick current *best* edge $\{u, v\} \in L$ and remove it from L. Let $S_u, S_v \in S$ such that $u \in S_u$ and $v \in S_v$. If $S_u = S_v$, do nothing. If $S_u \neq S_v$, merge S_u and S_v and update S: $S \leftarrow S \cup \{S_u \cup S_v\} - \{S_u, S_v\}$.
(iv) If $|S| \geq \theta$, repeat the previous step.
(v) Enumerate the elements of S, i.e., $S = \{S_1, ..., S_\sigma\}$. Then, $\forall S_i \in S$ and $\forall v \in S_i$, define $\kappa(v) = i$ and $\kappa^{-1}(i) = S_i$.

The procedure above creates mapping κ which is then used to induce the κ-reduced graph $G' = (V', E')$ from graph $G = (V, E)$. Parameter θ indicates the number of vertices of the reduced graph, i.e., $|V'| = \theta$. However, the procedure requires a scoring method. The edge scoring methods considered herein are described next.

The first scoring method is Jaccard's similarity (Real and Vargas, 1996), which is based on the notion of neighborhood,

$$J(u, v) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}. \tag{2}$$

Jaccard's similarity permits assessing how similar two vertices are given their local context. For all $u, v \in V$, it holds that $0 \leq J(u, v) \leq 1$. If $J(u, v) = 1$, then vertices u, v are isomorphic (interchangeable).

Conversely, following the idea of preserving salient edges, a *quasi-inverse* Jaccard's may be defined to identify relevant edges in a graph,

$$Q(u, v) = \frac{1 + |N(u) \cup N(v)|}{1 + |N(u) \cap N(v)|}, \tag{3}$$

for $u, v \in V$. Contrary to Jaccard's similarity, the quasi-inverse suggests contracting edges with low $Q(u, v)$. Note that $0 < Q(u, v) \leq |V|$.

The last scoring method is based on degrees; it ranks edges by the multiplication of the degrees of the involved vertices,

$$M(u, v) = \delta(u)\delta(v), \tag{4}$$

where the less $M(u, v)$ is, the better.

The different scores were chosen so that the contracted edges do not affect the topology of the network too much. Score J identifies edges of topologically similar (redundant) vertices. Score Q identifies edges that are relevant to graph G and should be preserved. Score M identifies edges of well connected elements, which, if contracted, may introduce distortions in paths.

4. Experiments

4.1. *Networks considered*

To test the proposed algorithms, three networks are used herein: the national electric transmission network of Chile, the electric distribution network of the Greater Valparaíso, and the drinking water distribution network of the Greater Valparaíso.

For simplicity, system performance is measured by the number of vertices π-connected to the network source vertices. For the electric distribution network, the graph sources are the distribution substations. For the drinking water distribution network, the graph sources are the water reservoirs (tanks) and the elevation stations (also known as pumps or boosters). In both networks, disconnection from the network source means that a portion of the network is left without service, i.e., outage.

Electric transmission network. The model of the national electric power transmission network of Chile was built by collecting data of power generation units, substations and transmission lines, from different official Chilean data sources: the National Electrical Coordinator, the Ministry of Energy, and the Energía Abierta platform; details about the construction of this model can be found in Ferrario et al. (2019). This work only considers substations and transmission lines, and substations connected to power generation units are identified as source nodes. In total, the model comprises 994 vertices (substations), of which 274 are sources, and 1195 edges (transmission lines). The model evaluated herein only considers damage to the substations; the seismic fragility curves used were taken from FEMA (2013)

Electric distribution network. The data for the electric distribution network of The Greater Valparaíso Area was collected exclusively from online sources, following previous work (Monsalve and de la Llera, 2019). This work focuses only on the medium voltage subnetwork, which consists of 24954 edges (lines) and 24610 vertices (utility poles), with 16 being source vertices (next to substations). The retrieved data consisted of map drawings (GIS vector data) with some geometric imperfections. Thus, automated correction rules were used to correct the data. In this model, it was assumed that only the overhead lines were at risk during an earthquake; the fragility curve used was taken from Vanzi (1996).

Drinking water distribution network. The data to build the water network model was provided by the The Superintendence of Waterworks Services (SISS), a public institution that regulates and monitors private companies managing water resources in Chile. Pipelines were represented by GIS polyline shapefiles, including pipe age, material and diameter, while distribution tanks were marked as GIS point shapefiles. Pump data contains location, design pressure head and power. In this network, edges represent pipelines and vertices represent either pipe junctions, water tanks or pumping stations. A script was prepared to adjust connectivity, which is necessary for hydraulic analysis (not considered in this work). In total, 39778 edges and 36403 vertices resulted from the analysis, with 74 source vertices.

4.2. *Experimental protocol*

The networks are reduced to 25% of the original vertices using the proposed algorithms, and simulated under $10,000$ scenarios. Earthquake magnitude is sampled uniformly, to enhance infrastructure affectation, but otherwise, earthquakes were sampled according to Poulos et al. (2018).

To make the evaluation of the original and reduced networks correspond each other better, element survival is evaluated in the original graph. Then this information is reutilized in the reduced graphs. Essentially, p_v, p_e and q can be gathered from the original graph using the ratio of failed elements. If for edge $e \in E'$, $q(e) \geq 0.5$, all the comprised edges are deemed to fail. For vertex $v \in V'$, this requirement changes to $p_v(v)+p_e(v)-p_v(v)p_e(v) \geq 0.5$. Because of this translation, π-connectivity for the original graph and its approximation with the reduced graphs can be assessed on a scenario-by-scenario basis. This permits plotting and measuring the approximation error. Note that, network performance is measured as number of π-connected vertices to the network source.

4.3. *Results*

Fig. 2 depicts the experimental results on $10,000$ simulations, for each of the infrastructure networks considered and each of the reduced graphs. The plots include the root mean squared error (σ) associated with each network and graph reduction score.

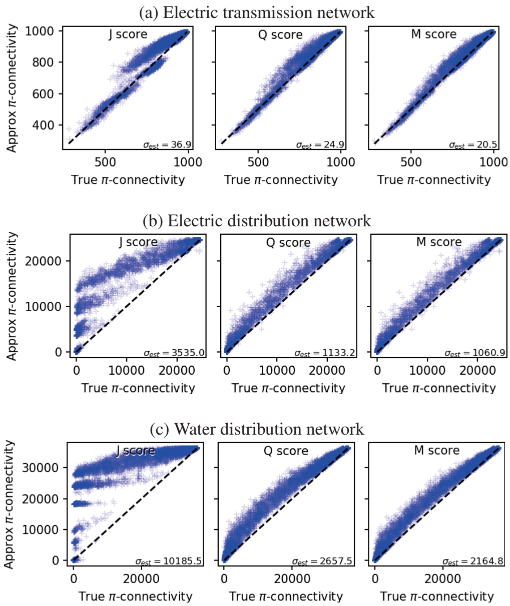

Fig. 2. Simulation results for the (a) electric transmission, (b) electric distribution, and (c) water distribution networks. The score used and the mean root squared error σ are indicated inside each scatter plot. The blue crosses represent the original performance and its estimation on a reduced graph in a scenario. The dashed black line represents the diagonal; exact estimation would leave the blue crosses in the diagonal.

Overall, the Fig. 2 shows that contracting edges according to Jaccard's similarity, J, is not particularly fruitful; it seems to be actually quite counterproductive. Estimation of the connectivity to the source seems to be quite unreliable for the electric and water distribution networks. On the other hand, using graph reductions based on Q and M scores seem to result in overall better estimations. The root mean squared error σ appears to be similar for both scores, with M yielding the lowest error.

It can be speculated that Jaccard's score, J, was not particularly adequate for graph reductions, because it might lead to the contraction of edges linking vertices that are very important to the network, i.e., that connect it. It may also be because it is likely to lead to large clusters. This stresses the ability of the approximations introduced herein (see Eq. 1) to evaluate the connectivity of thick vertices. On the contrary, Q and M scores are likely to render smaller thick vertices, which may prevent incurring in some approximation errors. Moreover, M selects edges of vertices that are generally poorly connected to the rest of the network.

5. Conclusions

This work evaluated the performance of three graph reduction algorithms for simplifying real-world critical infrastructure network models. The proposed algorithms were designed to evaluate the viability of using graph reductions to conduct probabilistic seismic risk analysis. The experiments show that the reduced graphs can actually be used to estimate network affectation of the original, detailed infrastructure system. However, more extensive experiments shall be conducted in the future; these experiments must consider more and better graph reduction algorithms, as well as testing with more infrastructure networks, size reductions, and different performance measures.

Acknowledgement

This research has been sponsored by the Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017, as well as grants ANID/FONDECYT/3170867, Modeling the impact of natural hazards on complex interdependent systems using multilayer networks and graphical models, ANID/FONDECYT/3180464, Risk Analysis for Critical Infrastructures Protection, and ANID/FONDECYT/1170836, SIBER-RISK: Simulation Based Earthquake Risk and Resilience of Interdependent Systems and Networks.

References

Ahn, K. J., S. Guha, and A. McGregor (2012). Graph sketches: sparsification, spanners, and subgraphs. In *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI symposium on Principles of Database Systems*, pp. 5–14.

Buluç, A., H. Meyerhenke, I. Safro, P. Sanders, and C. Schulz (2016). Recent advances in graph partitioning. In *Algorithm Engineering*, pp. 117–158. Springer.

Cuadra, L., S. Salcedo-Sanz, J. Del Ser, S. Jiménez-Fernández, and Z. W. Geem (2015). A critical review of robustness in power grids using complex networks concepts. *Energies 8*(9), 9211–9265.

Dai, L., B. Derudder, and X. Liu (2018). Transport network backbone extraction: A comparison of techniques. *Journal of Transport Geography 69*, 271–281.

Delling, D., R. Görke, C. Schulz, and D. Wagner (2009). Orca reduction and contraction graph clustering. In *International Conference on Algorithmic Applications in Management*, pp. 152–165. Springer.

Di Nardo, A., M. Di Natale, G. F. Santonastaso, V. G. Tzatchkov, and V. H. Alcocer-Yamanaka (2014). Water network sectorization based on graph theory and energy performance indices. *Journal of Water Resources Planning and Management 140*(5), 620–629.

Di Nardo, A., M. Di Natale, G. F. Santonastaso, and S. Venticinque (2011). Graph partitioning for automatic sectorization of a water distribution system. *Proceedings of CCWI*.

Ding, T., H. Sun, K. Sun, F. Li, and X. Zhang (2015). Graph theory based splitting strategies for power system islanding operation. In *2015 IEEE Power & Energy Society General Meeting*, pp. 1–5. IEEE.

Du, N., B. Wu, and B. Wang (2007). Backbone discovery in social networks. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI'07)*, pp. 100–103. IEEE.

Dunn, S. and S. M. Wilkinson (2013). Identifying critical components in infrastructure networks using network topology. *Journal of Infrastructure Systems 19*(2), 157–165.

FEMA (2013). Hazus–MH 2.1: Technical manual. Department of Homeland Security, Federal Emergency Management Agency, Mitigation Division Washington, D.C.

Feng, Z. (2018). Similarity-aware spectral sparsification by edge filtering. In *2018 55th ACM/ESDA/IEEE Design Automation Conference (DAC)*, pp. 1–6. IEEE.

Ferrario, E., N. Pedroni, and E. Zio (2016). Evaluation of the robustness of critical infrastructures by hierarchical graph representation, clustering and monte carlo simulation. *Reliability Engineering & System Safety 155*, 78–96.

Ferrario, E., A. Poulos, J. de la Llera, A. Lorca, A. Onetto, and C. Magnere (2019). Representation and modeling of the chilean electric power network for seismic resilience analysis. In *29th European Safety and Reliability Conference (ESREL 2019)*.

Foti, N. J., J. M. Hughes, and D. N. Rockmore (2011). Nonparametric sparsification of complex multiscale networks. *PLoS one 6*(2).

Franchin, P. and F. Cavalieri (2015). Probabilistic assessment of civil infrastructure resilience to earthquakes. *Computer-Aided Civil and Infrastructure Engineering 30*(7), 583–600.

Gdalyahu, Y., D. Weinshall, and M. Werman (1999). A randomized algorithm for pairwise clustering. In *Advances in neural information processing systems*, pp. 424–430.

Glattfelder, J. B. and S. Battiston (2009). Backbone of complex networks of corporations: The flow of control. *Physical Review E 80*(3), 036104.

Gupta, A. (1997). Fast and effective algorithms for graph partitioning and sparse-matrix ordering. *IBM Journal of Research and Development 41*(1.2), 171–183.

Hamann, M., G. Lindner, H. Meyerhenke, C. L. Staudt, and D. Wagner (2016). Structure-preserving sparsification methods for social networks. *Social Network Analysis and Mining 6*(1), 22.

Han, Y. and R. A. Davidson (2012). Probabilistic seismic hazard analysis for spatially distributed infrastructure. *Earthquake Engineering & Structural Dynamics 41*(15), 2141–2158.

Hendrickson, B. and R. W. Leland (1995). A multi-level algorithm for partitioning graphs. In *Proceedings of the 1995 ACM/IEEE conference on SuperComputing*, Volume 95, pp. 1–14.

Hernandez-Fajardo, I. and L. Dueñas-Osorio (2013). Probabilistic study of cascading failures in complex interdependent lifeline systems. *Reliability Engineering & System Safety 111*, 260 – 272.

Jayaram, N. and J. W. Baker (2010). Efficient sampling and data reduction techniques for probabilistic seismic lifeline risk assessment. *Earthquake Engineering & Structural Dynamics 39*(10), 1109–1131.

Jiang, B. (2007). A topological pattern of urban street networks: universality and peculiarity. *Physica A: Statistical Mechanics and its Applications 384*(2), 647–655.

Lallemant, D., A. Kiremidjian, and H. Burton (2015). Statistical procedures for developing earthquake damage fragility curves. *Earthquake Engineering & Structural Dynamics 44*(9), 1373–1389.

Liu, Y., N. Shah, and D. Koutra (2015). An empirical comparison of the summarization power of graph clustering methods. *arXiv preprint arXiv:1511.06820*.

Marlim, M. S., G. Jeong, and D. Kang (2019). Identification of critical pipes using a criticality index in water distribution networks. *Applied Sciences 9*(19).

Meyerhenke, H., P. Sanders, and C. Schulz (2014). Partitioning complex networks via size-constrained clustering. In *International Symposium on Experimental Algorithms*, pp. 351–363. Springer.

Miao, W., H. Jia, and Z. Tian (2013). A fast partitioning method for power system controlled splitting. *Automation of Electric Power Systems 37*, 24–30.

Monsalve, M. and J. de la Llera (2019). Identifying critical components in power distribution networks using graph theoretical measures. In *29th European Safety and Reliability Conference (ESREL 2019)*.

Montalva, G. A., N. Bastías, and A. Rodriguez-Marek (2017). Ground-Motion Prediction Equation for the Chilean Subduction Zone. *Bulletin of the Seismological Society of America 107*(2), 901–911.

Perelman, L. and A. Ostfeld (2012). Water-distribution systems simplifications through clustering. *Journal of Water Resources Planning and Management 138*(3), 218–229.

Pitilakis, K., H. Crowley, and A. M. Kaynia (2014). *SYNER-G: Typology Definition and Fragility Functions for Physical Elements at Seismic Risk: Buildings, Lifelines, Transporta-*

tion Networks and Critical Facilities. Dordrecht: Springer Netherlands.

Poulos, A., M. Monsalve, N. Zamora, and J. C. de la Llera (2018). An Updated Recurrence Model for Chilean Subduction Seismicity and Statistical Validation of Its Poisson Nature. *Bulletin of the Seismological Society of America 109*(1), 66–74.

Qiu, H. and E. R. Hancock (2007). Graph simplification and matching using commute times. *Pattern Recognition 40*(10), 2874–2889.

Real, R. and J. M. Vargas (1996). The probabilistic basis of Jaccard's index of similarity. *Systematic Biology 45*(3), 380–385.

Ron, D., I. Safro, and A. Brandt (2011). Relaxation-based coarsening and multiscale graph organization. *Multiscale Modeling & Simulation 9*(1), 407–423.

Salman, A. M. and Y. Li (2018). A probabilistic framework for multi-hazard risk mitigation for electric power transmission systems subjected to seismic and hurricane hazards. *Structure and Infrastructure Engineering 14*(11), 1499–1519.

Salzman, O., D. Shaharabani, P. K. Agarwal, and D. Halperin (2014). Sparsification of motion-planning roadmaps by edge contraction. *The International Journal of Robotics Research 33*(14), 1711–1725.

Satuluri, V., S. Parthasarathy, and Y. Ruan (2011). Local graph sparsification for scalable clustering. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pp. 721–732.

Scellato, S., A. Cardillo, V. Latora, and S. Porta (2006). The backbone of a city. *The European Physical Journal B-Condensed Matter and Complex Systems 50*(1-2), 221–225.

Serrano, M. Á., M. Boguná, and A. Vespignani (2009). Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences 106*(16), 6483–6488.

Spielman, D. A. and N. Srivastava (2011). Graph sparsification by effective resistances. *SIAM Journal on Computing 40*(6), 1913–1926.

Stewart, J. P., J. Douglas, M. Javanbarg, Y. Bozorgnia, N. A. Abrahamson, D. M. Boore, K. W. Campbell, E. Delavaud, M. Erdik, and P. J. Stafford (2015). Selection of ground motion prediction equations for the global earthquake model. *Earthquake Spectra 31*(1), 19–45.

Van, J., P. Shi, and D. Zhang (2004). Mesh simplification with hierarchical shape analysis and iterative edge contraction. *IEEE Transactions on Visualization and Computer Graphics 10*(2), 142–151.

Vanzi, I. (1996). Seismic reliability of electric power networks: methodology and application. *Structural Safety 18*(4), 311–327.

Wang, Z., M. Broccardo, and J. Song (2019). Hamiltonian monte carlo methods for subset simulation in reliability analysis. *Structural Safety 76*, 51 – 67.

Wu, A. Y., M. Garland, and J. Han (2004). Mining scale-free networks using geodesic clustering. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 719–724.

Xiao, N.-C., M. J. Zuo, and C. Zhou (2018). A new adaptive sequential sampling method to construct surrogate models for efficient reliability analysis. *Reliability Engineering & System Safety 169*, 330 – 338.

Yehia, A. and O. Swei (2020). Probabilistic infrastructure performance models: An iterative-methods approach. *Transportation Research Part C: Emerging Technologies 111*, 245 – 254.

Zhang, Q., Z. Y. Wu, M. Zhao, J. Qi, Y. Huang, and H. Zhao (2017). Automatic partitioning of water distribution networks using multiscale community detection and multiobjective optimization. *Journal of Water Resources Planning and Management 143*(9), 04017057.

Zuev, K. M., S. Wu, and J. L. Beck (2015). General network reliability problem and its efficient solution by subset simulation. *Probabilistic Engineering Mechanics 40*, 25 – 35.