ICOSSAR 2021-2022
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021-2022), Sept. 13-17, 2022, Shanghai, P.R. China
ISBN 978-7-5478-6230-8

Development of Centrality-based Selective Recursive Decomposition Algorithm Using Network Simplification

Dongkyu Lee[a] and Junho Song[b]

[a]*Department of Civil and Environmental Engineering, Seoul National University, E-mail: idongkr@snu.ac.kr.*
[b]*Department of Civil and Environmental Engineering, Seoul National University, E-mail: junhosong@snu.ac.kr.*

ABSTRACT: It is essential to analyze the reliability of lifeline networks such as transportation networks and water supply systems, which play a central role in maintaining urban functions. To evaluate the reliability of lifeline networks efficiently, a non-simulation-based algorithm, termed recursive decomposition algorithm (RDA) identifies disjoint cut sets and link sets for calculating network reliability. However, there are limitations in applying RDA due to complex characteristics of real lifeline networks. To overcome this issue, the authors recently proposed a new approach utilizing a network centrality index, termed centrality-based selective recursive decomposition algorithm (CS-RDA; Lee and Song, 2021). By decomposing subgraphs containing critical components with high centrality with a priority, CS-RDA reduces the number of subgraphs required to achieve the target bound width. A clustering method utilizing edge-betweenness centrality was also introduced to handle complex networks with minimal information loss. This paper presents CS-RDA and clustering method and demonstrates efficiency by a numerical example of a water distribution network in Sejong, South Korea.

KEYWORDS: Centrality, Clustering, Modularity, Network Reliability, Non-simulation-based approach.

1 INTRODUCTION

As cities become more complex, modern societies are highly dependent on lifeline networks such as transportation networks, electricity networks, and gas networks. When a severe earthquake occurs, the reliability and performance of lifeline networks become far more critical; damaged lifeline networks can hamper prompt responses, and lead to social repercussions caused by delayed recovery. Therefore, it is essential to rapidly assess the two-terminal reliability problem in system reliability analysis (SRA) to make appropriate post-disaster decisions (Stern et al. 2017).

However, the reliability analysis of large-scale networks faces various obstacles like an excessive number of components, complex network topology, and intricate interdependence between network components. To overcome these difficulties and analyze the network reliability promptly but accurately, many simulation-based approaches that have straightforward applicability and high flexibility have been proposed. These simulation-based methods, however, have a critical limitation: a tremendous amount of computational time is required to achieve a statistically significant level of convergence for rare events. In addition, these approaches often make it intractable to perform probabilistic inferences or to quantify the contribution of each component to network failure events using the result of simulations.

Various non-simulation-based approaches have been developed to address the limitations of simulation-based approaches. For example, Li & He (2002) proposed an algorithm to decompose a network recursively, called the recursive decomposition algorithm (RDA), which decomposes a network into cut sets and link sets that are mutually exclusive to each other. These identified disjoint cut sets and link sets result in a great advantage in calculating the network reliability.

Lim & Song (2012) went further by proposing a method, termed the selective recursive decomposition algorithm (S-RDA). The algorithm decomposes critical cut sets and link sets with the most likelihood with a priority. Since the probabilities of the critical cut sets and link sets identified by S-RDA are generally much higher than those identified by the original RDA, the convergence speed of the bounds on network reliability is significantly enhanced. The computational cost saved by S-RDA facilitates analyzing large-scale lifeline networks that cannot be handled by existing algorithms. Furthermore, inter- and intra-event uncertainties in spatially correlated ground motions were incorporated into network reliability analysis by S-RDA. It was found that the assumption of statistical independence that disregards spatial correlations can over- or underestimate network reliability significantly.

Although the performance of S-RDA is more

ISBN 978-7-5478-6230-8

efficient than the original RDA, some computational limitations still exist in its direct applications to real urban lifeline networks. Besides, the calculations to consider spatial correlations may increase the computational complexity and costs. Consequently, the reduction in total calculation time may become negligible. For reasonable calculation time and memory savings and fast convergence of bounds of network reliability, network simplification was also considered in RDA-based analyses . Lim et al. (2015) proposed an SRA method through a hierarchical representation. A simplified network consists of the clusters identified by spectral clustering algorithms was utilized in the proposed multi-scale analysis.

Recently, the authors proposed an approach utilizing network centrality, termed centrality-based selective recursive decomposition algorithm (CS-RDA; Lee & Song 2021), to improve the efficiency of existing RDAs and simplify the network at the same time. By preferentially decomposing the node which is most likely to belong to min-cut within each subgraph based on centrality, the convergence time of the bounds on network reliability can be reduced. The authors also introduced a clustering algorithm based on centrality in terms of edges, which is more suitable for large-scale operations than the spectral clustering algorithm. The reliability of components that make up the simplified network is evaluated by sampling methods. As a result, the computational complexity of the SRA of the simplified network decreases exponentially. The efficiency of the proposed approaches is demonstrated by a numerical example of a large-scale network and compared with that of existing RDAs.

2 FRAMEWORK OF RECURSIVE DECOMPOSITION ALGORITHMS

Using graph theory, let us consider a lifeline network as a graph $G = (N, E)$, where N and E respectively denote the sets of the nodes and edges in the graph. The node-set N consists of nodes representing both node-type (e.g., pipelines and electricity wires) and line-type components (e.g., stations and power plants). Regardless of the component type, the failure of components in the lifeline network results in the failure of node $n \in N$, whereas edge $e \in E$ only represents the connection between two nodes, that is, network topology.

2.1 Recursive decomposition algorithm

After a natural or man-made disaster occurs, suppose each component has a binary state: operative or failed. Likewise, an origin/destination (O/D) pair can be disconnected or remain connected. To express the states of individual components and the O/D connectivity, "node functions" a_i, $i = 1, \dots, N$, and a "structure function" $\Psi(G)$, i.e., Bernoulli random variables representing the state of the i^{th} component and the O/D connectivity, respectively, are introduced as follows:

$$a_i = \begin{cases} 1, & \text{if } i^{\text{th}} \text{ component operates} \\ 0, & \text{if } i^{\text{th}} \text{ component fails,} \end{cases} \tag{1}$$

$$\Psi(G) = \begin{cases} 1, & \text{if } O/D \text{ pair is connected} \\ 0, & \text{if } O/D \text{ pair is disconnected.} \end{cases} \tag{2}$$

Consider the shortest path of an O/D pair, which is represented as $A_1 = a_1 a_2 \dots a_n$. According to the Boolean operation laws, the structure function $\Psi(G)$ is expressed as a linear function of the node and structure functions of subgraphs, in the ascending order of component numbering, i.e.,

$$\begin{aligned} \Psi(G) &= a_1 a_2 \dots a_n \Psi(G) + \overline{(a_1 a_2 \dots a_n)} \Psi(G) \\ &= a_1 a_2 \dots a_n + \bar{a}_1 \Psi(G_1) + a_1 \bar{a}_2 \Psi(G_2) \\ &\quad + \cdots + a_1 a_2 \cdots a_{n-1} \bar{a}_n \Psi(G_n), \end{aligned} \tag{3}$$

where G_i represents the subgraph of G, obtained by removing the i^{th} node from the original graph G. If G_i still has the shortest path of the O/D pair, $\Psi(G_i)$ is recursively expanded in the same way. The selection of subgraphs to explore the O/D connectivity follows a breadth-first search (BFS) ordering.

Using the structure function, the network reliability can be expressed as a probability about the structure function. If all disjoint link sets L_i, $i = 1, \dots, N_L$, are identified by an algorithm, the network reliability R, can be calculated by summing up their probabilities, i.e.,

$$R = P[\Psi(G) = 1] = P\left(\bigcup_{i=1}^{N_L} L_i\right) = \sum_{i=1}^{N_L} P(L_i). \tag{4}$$

The network failure probability P_f, can be obtained in the same way using all disjoint cut sets between the origin and the destination node, C_i, $i = 1, \dots, N_C$, i.e.,

$$P_f = P[\Psi(G) = 0] = P\left(\bigcup_{i=1}^{N_C} C_i\right) = \sum_{i=1}^{N_C} P(C_i). \tag{5}$$

However, it is highly time-consuming to identify all disjoint cut sets and link sets for a large-scale network. In this case, one can calculate the upper and lower bounds on the network failure probability. Both identified cut sets and link sets are utilized to obtain the

ICOSSAR 2021-2022
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021-2022), Sept. 13-17, 2022, Shanghai, P.R. China
ISBN 978-7-5478-6230-8

information about the bounds on the network failure probability, i.e.,

$$\sum_{i=1}^{n_C} P(C_i) \leq P_f \leq 1 - \sum_{i=1}^{n_L} P(L_i), \tag{6}$$

where n_c and n_L are the numbers of identified disjoint cut sets and link sets, respectively. The decomposition of structure functions $\Psi(G_i)$ are repeated until the bound width, i.e., the gap between the upper and lower bounds on the network failure probability, becomes smaller than a target value ε, as shown in Figure 1.

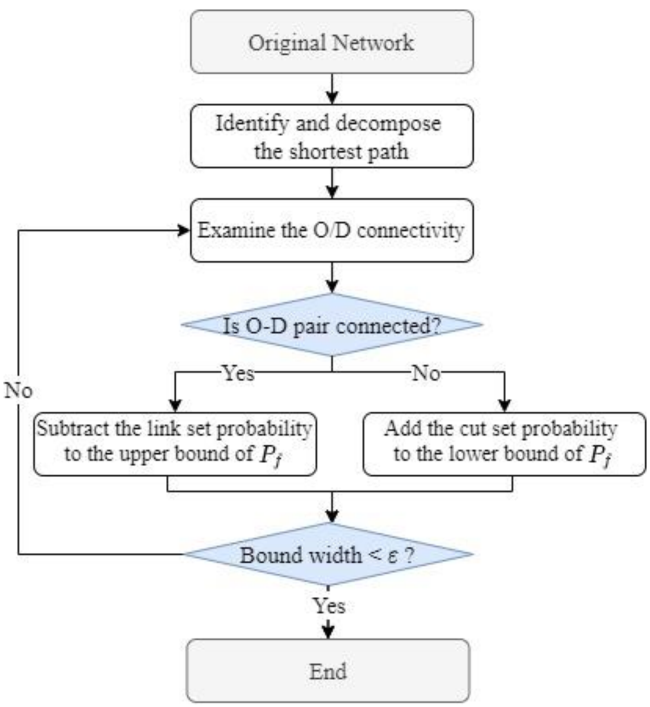

Figure 1. Flowchart of recursive decomposition algorithm.

2.2 Selective recursive decomposition algorithm

The two most notable differences between S-RDA (Lim & Song 2012) and the original RDA are: (1) path selection and (2) subgraph selection order. In S-RDA, critical disjoint link sets with primary contributions to the likelihood of the network connection are identified by finding the most reliable paths instead of the shortest paths. Furthermore, when choosing a subgraph to decompose, the one with the highest probability is identified first instead of node numbering order.

3 NETWORK RELIABILITY ASSESSMENT AND CLUSTERING IN TERMS OF NETWORK TOPOLOGY

Network centrality is used to analyze the quantitative importance of each component in the network. Various network centralities have been developed, e.g., influential people in social networks, key nodes on the Internet, and super-spreaders of diseases (Özgür et al. 2008). This paper focuses on betweenness centrality among the centrality indicators available in the literature.

3.1 Betweenness centrality

Betweenness centrality measures the probability that the shortest path between any two nodes in the graph passes through a given node. In detail, betweenness of the i^{th} component, $C_B(i)$, is given as

$$C_B(i) = \sum_{s \neq i} \sum_{t \neq s,i} \frac{v_i(s,t)}{v(s,t)}, \tag{7}$$

where $v(s,t)$ is the total number of shortest paths from component s to t, and $v_i(s,t)$ is the number of those passing through component i. In this paper, $v(s,t)$ and $v_i(s,t)$ will be substituted with the total number of the most reliable paths between component s to t, and the number of those via component i, respectively.

Betweenness can be classified into two categories depending on what components are handled: (1) node betweenness; and (2) edge betweenness. While both betweenness centralities are defined in the same principle, they focus on nodes and edges, respectively.

3.2 Centrality-based selective recursive decomposition algorithm

To improve the efficiency of SRA methods, it is important to prioritize key components of the network. To this end, we can consider two main ideas: (1) calculating critical disjoint cut sets and link sets with the highest probability, such as S-RDA, and (2) fundamentally reducing the number of expected subgraphs.

The recently proposed CS-RDA (Lee & Song 2021) utilizes node betweenness to find *critical nodes*, whose removals are expected minimize the number of branches in the following network decomposition process, e.g., O/D nodes. The algorithm is based on the observation that the removal of nodes with high betweenness accelerates network disconnection compared to the removal of randomly selected nodes (Albert et al. 2000, Iyer et al. 2013). The removal of the node with high betweenness shortens the network decomposition process exponentially owing to the reduced number of potential subgraphs. The shortened computational process not only accelerates the network reliability analysis, but also enables the

ISBN 978-7-5478-6230-8

analysis of even larger networks.

However, there is a fatal drawback in CS-RDA. Because node betweenness depends on network topology, the convergence speed of CS-RDA varies vastly over O/D pairs even within a network. When betweenness of nodes on the most reliable path between the O/D pair is exceptionally low, key nodes that are highly likely to belong to the min-cut could be pushed to lower priorities. In this case, there is little progress in convergence speed due to the low probability of the identified sets, or CS-RDA can be rather inferior to the existing algorithms. To prevent such cases, one can assign virtual nodes to each O/D node. The introduced virtual node increases not only the centrality of the O/D nodes, but also the centrality of the nodes located in the path of the two nodes, resulting in critical nodes with high centrality.

3.3 Network simplification using edge-betweenness algorithm

Even after CS-RDA improves the efficiency of network reliability analysis significantly, some large lifeline networks still exceed analyzable size. Moreover, as the number of disjoint sets within the network increases exponentially, the accuracy of calculating the probability of each disjoint set gets worse due to the accumulated computational errors caused by the high-dimensional calculation.

To address these problems of complex networks, various clustering-based network simplification schemes have been considered (Gómez et al. 2013, Lim et al. 2015). Clustering algorithms aim to minimize the computational complexity of network analysis and to preserve information about components and the topology of the original network. However, most existing clustering methods show infeasible computational complexity in large-scale networks, and there is no objective basis for judging the quality of the clustering results.

To measure the goodness of a given clustering choice, modularity Q is often used (Newman & Girvan 2004). Modularity is the normalized difference between the actual and the expected numbers of the edges connecting a pair of nodes in the same cluster, i.e.,

$$Q = \frac{1}{2m}\sum_i \sum_j \left\{\left[A_{ij} - \frac{C_D(i)C_D(j)}{2m}\right]\delta_{c_i c_j}\right\}, \qquad (8)$$

where A_{ij} is a binary variable that becomes 1 when the i^{th} and j^{th} nodes are adjacent, and 0 otherwise; $m = \sum_i \sum_j A_{ij}/2$ is the total number of the edges in the network; and $\delta_{c_i c_j}$ is the Kronecker delta, which gives 1 if both the i^{th} and j^{th} nodes belong to the same cluster, and 0 otherwise.

It is NP-hard to find the network clustering choice maximizing the modularity (Brandes et al. 2007). To solve the problem quickly and efficiently, a heuristic edge-betweenness algorithm, also kwon as the Girvan-Newman algorithm, can be employed with relatively low computational cost (Newman & Girvan 2004). The edge-betweenness algorithm removes edges with the highest edge betweenness progressively from the original network until modularity stops increasing.

The clusters identified by the edge-betweenness algorithm are represented by the simplified representation of the network introduced in Lim et al. (2015). The simplified network consists of inter-cluster nodes, inter-cluster edges, and super-edges. An inter-cluster node is a type of node located at the boundary of a cluster and is adjacent to a node within another cluster. An inter-cluster edge is an edge that connects two inter-cluster nodes located in different clusters, while a super-edge is a virtual edge representing connectivity between inter-cluster nodes within the same cluster. Since inter-cluster nodes are a part of existing network components, there is no need to evaluate their reliabilities. This is also valid for inter-cluster edges, which are a part of the edges that do not inherently have their own reliabilities. On the contrary, the reliability of super-edges requires additional calculations. There are three methods for evaluating the reliability of super-edges depending on the number of O/D nodes included in the cluster, as detailed by Lim et al. (2015).

4 EVALUATION OF SEISMIC RELIABILITY AND STATISTICAL DEPENDENCE

The assessment of seismic reliability of each component should be preceded to evaluate network failure probability, which consists of two major uncertainties: (1) the seismic demands throughout the network, and (2) the seismic capacity of each component.

4.1 Uncertainties in ground motion intensity

When peak ground acceleration (PGA) is used as ground motion intensity measures, a typical ground motion prediction equation (GMPE) is described (Abrahamson & Youngs 1992) as

$$\ln PGA_i = f(M, R_i, \boldsymbol{\lambda}_i) + \eta + \varepsilon_i, \qquad (9)$$

where PGA_i is the actual PGA demand at the i^{th} site; $f(M_w, R_i, \boldsymbol{\lambda}_i)$ is the attenuation relation in terms of the

ISBN 978-7-5478-6230-8

moment magnitude of the earthquake M_w, the distance between the earthquake source and the i^{th} site, R_i, and a set of other explanatory variables $\boldsymbol{\lambda}_i$; and η and ε_i are the inter- and intra-event residuals with zero means and standard deviations σ_η and σ_ε, respectively. η and ε_i are assumed to be statistically independent of each other and follow Gaussian distributions.

In the numerical example in Section 5, the attenuation relation model in Lim & Song (2012) is adopted to predict $f(M, R_i, \boldsymbol{\lambda}_i)$ as

$$\begin{aligned} f(M_w, R_i, \boldsymbol{\lambda}_i) &= -0.5265 - 0.0115\sqrt{R_i^2 + 1.35^2} \\ &+[0.0599(M_w - 4.5) - 0.3303]\ln(R_i^2 + 1.35^2), \end{aligned} \quad (10)$$

where R_i is the distance between the i^{th} site and the epicenter, given in km.

To consider the correlation coefficient between the seismic demands, Goda and Hong (2008) suggested models for correlation coefficient between residuals at the i^{th} and j^{th} sites. The correlation coefficient between ε_i and ε_j, $\rho_{\varepsilon_i\varepsilon_j}$, is often expressed as a function of the distance Δ_{ij} between the two sites and the spatial correlation model used in this paper is given as

$$\rho_{\varepsilon_i\varepsilon_j}(\Delta_{ij}) = \exp(-0.27\Delta_{ij}^{0.40}). \quad (11)$$

Then, one can derive the correlation between $\ln Y_i$ and $\ln Y_j$ in terms of characteristics of distance Δ_{ij}, inter-event residual, and intra-event residual, i.e.,

$$\rho_{\ln Y_i \ln Y_j}(\Delta_{ij}) = \frac{\sigma_\eta^2 + \rho_{\varepsilon_i\varepsilon_j}(\Delta_{ij})\sigma_\varepsilon^2}{\sigma_\eta^2 + \sigma_\varepsilon^2}, \quad (12)$$

where the standard deviations of the inter- and intra-event residuals are assumed to be 0.265 and 0.502, respectively, in this paper.

4.2 Seismic reliability assessment of structures

According to HAZUS-MH (FEMA 2012), the limit-state capacity of the i^{th} structure, LS_i, follows Lognormal distribution with parameters α_i^{LS} and β_i^{LS}, which are the mean and standard deviation of the natural logarithm of LS_i, respectively. In Section 5, The parameters α_i^{LS} and β_i^{LS} of limit-state capacities LS_i are assumed to be 0.85 and 0.35, respectively. Based on the capacity and demand of components, the failure probability of each structure is computed as

$$\begin{aligned} P(E_i) = \Phi(-\beta_i) &= P(\ln LS_i \le \ln Y_i) \\ &= \Phi\left(\frac{f(M_w, R_i, \boldsymbol{\lambda}_i) - \alpha_i^{LS}}{\sqrt{\sigma_\eta^2 + \sigma_\varepsilon^2 + {\beta_i^{LS}}^2}}\right), \end{aligned} \quad (13)$$

where β_i is the generalized reliability index, and $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of the standard Gaussian distribution.

Lee & Song (2021) derived a formula to calculate the correlation coefficient of two failure events E_i and E_j, ρ_{ij} as

$$\rho_{ij} = \frac{\sigma_\eta^2 + \rho_{\varepsilon_i\varepsilon_j}(\Delta_{ij})\sigma_\varepsilon^2 + \beta_i^{LS}\beta_j^{LS}\delta_{ij}}{\sqrt{\sigma_\eta^2 + \sigma_\varepsilon^2 + {\beta_i^{LS}}^2}\sqrt{\sigma_\eta^2 + \sigma_\varepsilon^2 + {\beta_j^{LS}}^2}}, \quad (14)$$

where δ_{ij} is the Kronecker delta, which is 1 if $i = j$, and 0 otherwise. The derived equation ensures the accuracy, and shortens the computational time compared to the numerical methodology.

4.3 Component importance measure

In efforts to establish a cost-effective maintenance planning for lifeline networks, it is helpful to identify components making major contributions to network reliability based on topological importance and component failure probabilities. To quantify and rank the contributions of components to a network, various component importance measures have been proposed. The conditional probability importance measure (CPIM) measures a conditional probability defined as

$$CPIM_i = P(E_i|E_{net}) = \frac{P(E_i E_{net})}{P(E_{net})}, \quad (15)$$

where E_{net} is the network failure event, that is, the event that the O/D pair is disconnected (Song & Kang 2009).

Lee & Song (2021) proposed to evaluate the component importance by the combination of node betweenness and failure probability of each node instead of a single index. The calculation of the two quantities consumes a very short time, and one can grasp the component importance in terms of vulnerability and network topology, respectively.

5 NUMERICAL EXAMPLE

A numerical example is presented to demonstrate the performance of CS-RDA and multi-scale approach: water network in Sejong, South Korea, modified from Lim et al. (2015). The computational times in this

ICOSSAR 2021-2022
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021-2022), Sept. 13-17, 2022, Shanghai, P.R. China
ISBN 978-7-5478-6230-8

paper are based on the use of MATLAB® on a personal computer with AMD Ryzen 5 3600 3.60 GHz CPU and 16GB RAM.

To compare the efficiency of the existing RDAs and proposed CS-RDA, the number of disjoint sets and computational times required to achieve the target bound width of 1% are presented. Subsequently, we discuss how the computational cost and accuracy of the analysis change in the simplified network by using the edge-betweenness algorithm.

Figure 2 shows the Sejong water distribution network that consists of 158 nodes (59 node-type and 99 line-type components) and 198 bi-directional edges; the total number of components is 554 (= 158 + 198 × 2) under a scenario earthquake with a moment magnitude of $M_w = 6.0$. The edge-betweenness algorithm identifies four clusters (blue facets) as shown in Figure 3. The simplified network has 35 nodes (red dots) and 44 bi-directional edges (thick gray line); the total number of components is reduced to 123 (= 35 + 44 × 2).

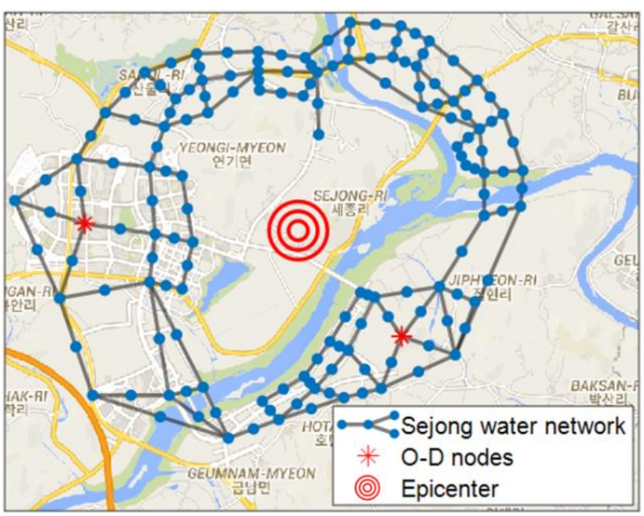

Figure 2. Water network in Sejong City (adapted from Lim et al. (2015)).

Figures 4-7 visualize and compare the bounds on the network reliability and their widths obtained by the original RDA, S-RDA and CS-RDA, in terms of the number of identified disjoint sets, and the computational times. The number of the identified disjoint sets needed for CS-RDA is 2,417, which is only about 10.08% and 22.82% of that for the original RDA and S-RDA, respectively. Comparing Figures 4-5 or Figures 6-7, the computational time ratios of CS-RDA to the original RDA and S-RDA are similar to the disjoint set ratios, respectively, which means that the computational time relies primarily on the number of disjoint sets. Even with this remarkably reduced computation time and memory

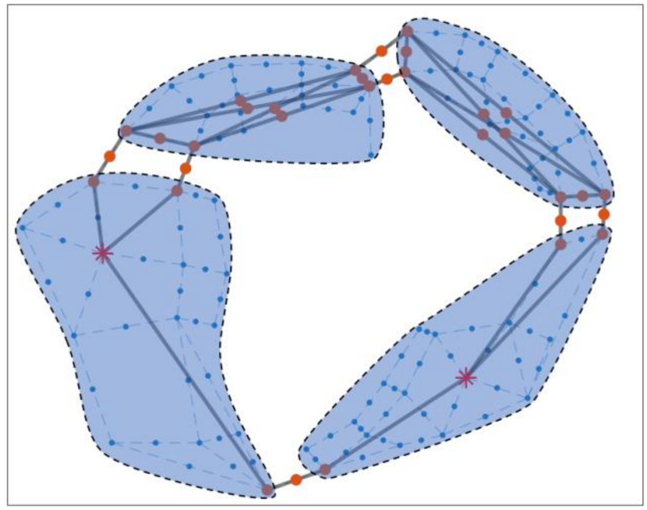

Figure 3. Simplified water network in Sejong City.

for CS-RDA, the bounds on network reliability are close to those using existing algorithms, as well as to the value $R = 0.8841$ obtained by MCS.

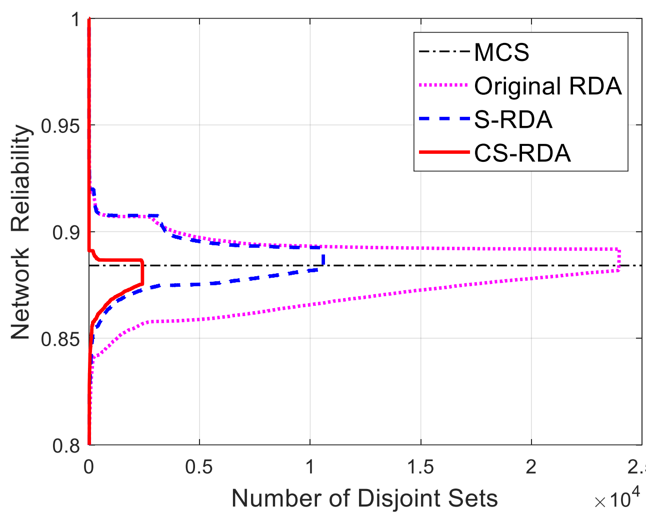

Figure 4. Bounds on network reliability in terms of the number of disjoint sets.

Table 1 summarizes the failure probability analysis results of the original and the simplified network using CS-RDA. In the network simplification, additional data preprocessing for the edge-betweenness algorithm and the evaluation of failure probability and statistical dependency of super-components is required. Despite the long preprocessing time (about 74 seconds), the total computational time decreases significantly by less than half. Considering the network failure probability obtained by MCS, there is little loss in terms of accuracy, although the analysis time excluding the preprocessing took only 8.34 seconds, which is only 2.26% of the computational time of the original network analysis.

ICOSSAR 2021-2022
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021-2022), Sept. 13-17, 2022, Shanghai, P.R. China
ISBN 978-7-5478-6230-8

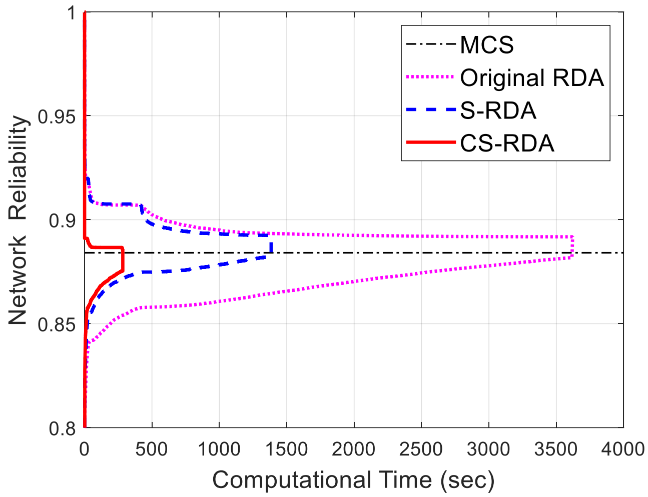

Figure 5. Bounds on network reliability in terms of computational time.

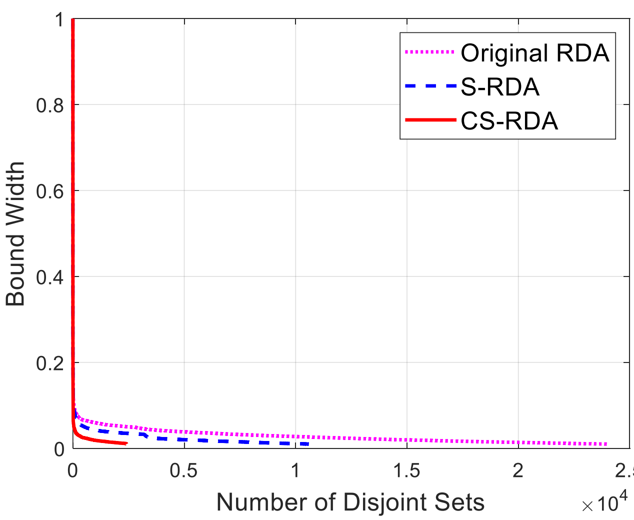

Figure 6. Bound widths in terms of the number of disjoint sets.

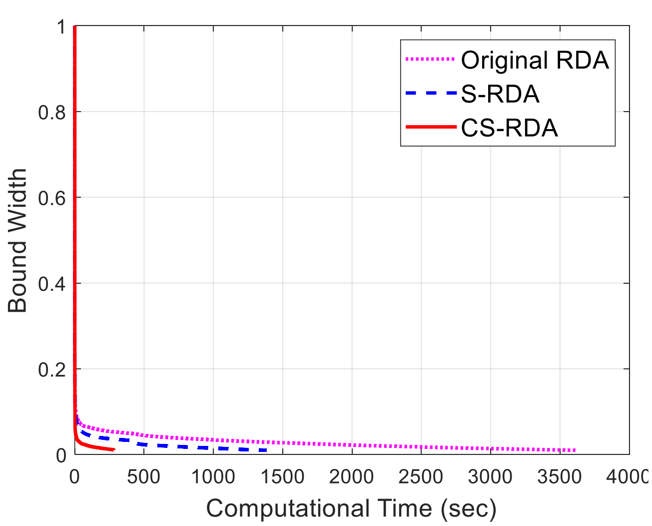

Figure 7. Bound widths in terms of computational time.

Table 1. Analysis results for water network in Sejong City.

Scale	Original Network	Simplified Network
Bound on R	[0.8777,0.8877]	[0.8732,0.8832]
No. of disjoint sets	2,414	84
Preprocessing time	-	74.1 s
Computational time	368 s	82.4 s

Table 2 shows that five nodes with the highest CPIMs and their CPIMs, node betweenness, and failure probabilities. While CPIMs comprehensively evaluate risk-related information and topological characteristics of the network to find nodes with deep impact in O/D disconnection, the computational time is considerable. In contrast, the calculations of node betweenness and node failure probabilities are very fast, and it is easy to quantify how each node is evaluated in terms of vulnerability and topological aspect, respectively.

Table 2. CPIMs, node betweenness (BC), and failure probabilities (P_f) of the nodes with the highest CPIMs in water network in Sejong City.

Component IDs	CPIM Rank	BC rank	P_f rank
16	2	1	124
21	5	52	5
46	1	3	30
54	3	130	1
139	4	107	2
Computational time	204 s	0.0209 s	0.0278 s

For example, nodes 16 and 46 with the highest CPIMs show a high level of node betweenness compared their failure probabilities, indicating that they are in a topologically key position; actually these two nodes are O/D nodes. On the other hand, nodes 21, 54 and 139 are identified as important components because of their high failure probabilities caused by its proximity to the epicenter.

6 CONCLUSIONS

In this paper, two algorithms utilizing the betweenness centrality, centrality-based selective recursive decomposition algorithm (CS-RDA) and multi-scale network simplification using the edge-betweenness algorithm, were reviewed for efficient reliability analysis of large-scale networks.

ISBN 978-7-5478-6230-8

CS-RDA arranged nodes on the most reliable path in the descending order of node betweenness to eliminate redundancy in cut sets and link sets. The algorithm further improved the efficiency by assigning more weights to O/D nodes especially when these nodes have low centralities. To deal with intractable large-scale networks, the edge-betweenness algorithm was introduced to cluster components and consequentially simplify networks. The simplified network is composed of sufficiently few nodes and edges that can be analyzed in a short time.

The numerical example of the Sejong water distribution network successfully demonstrated these proposed algorithms. CS-RDA showed a superior performance in terms of computational efficiency compared to the existing RDAs while the analysis of the simplified network took the performance of CS-RDA to the next level with little loss of accuracy. Moreover, the component importance is measured in terms of vulnerability and network topology, and compared in terms of CPIMs.

A non-simulation-based network reliability analysis and a clustering method have been unified into a network topology index called betweenness centrality. The centrality-based reliability assessment and multi-scale approach of networks are expected to help policymakers effectively deal with large and complex networks of urban communities.

ACKNOWLEDGMENTS

This research was supported by a grant (21SCIP-B146946-04) from Smart Civil Infrastructure Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

REFERENCES

[1] Abrahamson, N.A. & Youngs, R.R. 1992. A stable algorithm for regression analyses using the random effects model. *Bulletin of the Seismological Society of America* 82(1): 505-510.

[2] Albert, R., Jeong, H. & Barabási, A.L. 2000. Error and attack tolerance of complex networks. *Nature* 406(6794): 378-382.

[3] Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikoloski, Z. & Wagner, D. 2007. On modularity clustering. *IEEE transactions on knowledge and data engineering* 20(2): 172-188.

[4] Federal Emergency Management Agency (FEMA). 2012. HAZUS–MH 2.1: Multi-Hazard Loss Estimation Methodology, Earthquake Model.

[5] Goda, K. & Hong, H.P. 2008. Spatial correlation of peak ground motions and response spectra. *Bulletin of the Seismological Society of America* 98(1): 354-365.

[6] Gomez, C., Sanchez-Silva, M., Dueñas-Osorio, L. & Rosowsky, D. 2013. Hierarchical infrastructure network representation methods for risk-based decision-making. *Structure and Infrastructure Engineering* 9(3): 260-274.

[7] Iyer, S., Killingback, T., Sundaram, B. & Wang, Z. 2013. Attack robustness and centrality of complex networks. *PloS one* 8(4): e59613.

[8] Lee, D. & Song, J. 2021. Multi-scale seismic reliability assessment of networks by centrality-based selective recursive decomposition algorithm. *Earthquake Engineering and Structural Dynamics* 50(8): 2174-2194.

[9] Li, J. & He, J. 2002. A recursive decomposition algorithm for network seismic reliability evaluation. *Earthquake engineering and structural dynamics* 31(8): 1525-1539.

[10] Lim, H.W. & Song, J. 2012. Efficient risk assessment of lifeline networks under spatially correlated ground motions using selective recursive decomposition algorithm. *Earthquake Engineering and Structural Dynamics* 41(13): 1861-1882.

[11] Lim, H.W., Song, J. & Kurtz, N. 2015. Seismic reliability assessment of lifeline networks using clustering-based multi-scale approach. *Earthquake Engineering and Structural Dynamics* 44(3): 355-369.

[12] Newman, M.E. & Girvan, M. 2004. Finding and evaluating community structure in networks. *Physical review E* 69(2): 026113.

[13] Özgür, A., Vu, T., Erkan, G. & Radev, D.R. 2008. Identifying gene-disease associations using centrality on a literature mined gene-interaction network. *Bioinformatics* 24(13): i277-i285.

[14] Song, J. & Kang, W. 2009. System reliability and sensitivity under statistical dependence by matrix-based system reliability method. *Structural Safety* 31(2): 148-156.

[15] Stern, R.E., Song, J. & Work, D.B. 2017. Accelerated Monte Carlo system reliability analysis through machine-learning-based surrogate models of network connectivity. *Reliability Engineering and System Safety* 164: 1-9.