Probability Propagation Method for Reliability Assessment of Acyclic Directed Networks

Yanjie Tong[1] and Iris Tien, Ph.D., A.M.ASCE[2]

Abstract: Many civil infrastructure systems that deliver resources from source points to sinks, e.g., power distribution and gas pipeline networks, can be described as acyclic directed networks comprising nodes and links. Reliability assessment of these systems can be challenging, particularly for systems of increasing size and complexity and if the probabilities of rare events are of interest. This paper proposes a new analytical probability propagation method for reliability assessment of acyclic directed networks called the directed probability propagation method (dPrPm). Through a link-adding sequence to propagate a message consisting of the marginal and pairwise node reliabilities from source nodes to sink nodes, the method results in the upper and lower bounds of all sink node reliabilities. Reliability of a sink node is measured by the probability of reaching that node from a source node. Compared with previous methods, dPrPm addresses the case of multiple-sink networks, results in guaranteed reliability bounds, and analyzes acyclic directed networks as relevant for infrastructure systems. Proofs are provided guaranteeing the accuracy of dPrPm, and computation time is significantly reduced from typical exponential increases with system size to a polynomial increase. To assess performance, the proposed method was applied to three test applications: a directed grid network, a power distribution network, and a more complex gas pipeline network under seismic hazard. Results were compared with the exact solution and Monte Carlo simulations to evaluate accuracy and computational cost. Results showed that dPrPm performs equally well in terms of accuracy across network reliabilities and achieved order-of-magnitude increases in computational efficiency to obtain exact bounds on reliability assessments at all system sink nodes. **DOI: 10.1061/AJRUA6.0001017.**

Author keywords: Reliability analysis; Acyclic directed networks; Infrastructure system reliability; Probability propagation; Reliability bounds.

Introduction

Many civil infrastructure systems, e.g., power distribution and gas pipeline networks, are characterized by flow across the system. Considering system components as nodes and the connections between them as links, these infrastructures can be modeled as acyclic directed networks, with resources directed through the network from supplies to distribution points. Policies regarding inspection, maintenance, and replacement of components in the system rely heavily on reliability assessment of node accessibility, measuring the performance and resilience of the network.

Methods for reliability analysis include analytical and simulation-based methods. Whereas analytical approaches usually produce an exact result, enumeration of minimum link sets (MLS), minimum cut sets (MCS), or possible component states are often necessary to conduct an analysis. Increasing the number of nodes in the network typically yields an exponential increase in computational cost. For simulation-based methods, one of the challenges is to capture rare events. For networks that are highly reliable or highly unreliable, failure to capture sample points of rare events will result in high errors. Both the efficiency and accuracy of simulation-based methods vary on the reliability of links.

To address these challenges in system reliability assessment, a new analytical method, called the probability propagation method (PrPm), was proposed (Tong and Tien 2019). The method provides an estimate of network reliability for general complex networks. However, there are no guarantees of the accuracy of the estimate. Rather than providing an approximation, the present paper proposes an advancement of PrPm, the directed probability propagation method (dPrPm), to provide reliability assessment of directed acyclic networks with guaranteed upper and lower bounds. This requires new definitions of the propagation sequence and probability updating rules for message propagation. The outcome of dPrPm is the upper and lower bounds of all sink node reliabilities, given with 100% confidence level. The computational cost is independent of link reliabilities. dPrPm reduces the cost from a typical exponential increase with the number of nodes in a network to a polynomial increase.

The rest of the paper is organized as follows. The following section provides background on analytical and sampling-based system reliability assessment methods, and the origin of the idea of dPrPm. The proposed dPrPm is detailed in the next section, including descriptions of propagation sequences and probability updating rules. Proofs are given to guarantee the upper and lower bounds for acyclic directed networks. The paper then applies dPrPm to three examples: a directed grid network, power distribution network, and gas pipeline network. Results are shown to compare performance between the proposed method and existing methods in terms of accuracy and computation time for network reliability analysis.

[1]Ph.D. Student, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355 (corresponding author). Email: yjtong@gatech.edu

[2]Assistant Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355. ORCID: https://orcid.org/0000-0002-1410-632X. Email: itien@ce.gatech.edu

Note. This manuscript was submitted on September 21, 2018; approved on January 18, 2019; published online on July 13, 2019. Discussion period open until December 13, 2019; separate discussions must be submitted for individual papers. This paper is part of the ***ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering***, © ASCE, ISSN 2376-7642.

Background and Related Work

This section briefly introduces some existing methods for system reliability analysis. More in-depth study is available (e.g., Birolini 2004). Systems that can be reduced to fundamental parallel and series configurations can be readily analyzed. Analysis becomes complicated under complex configurations, e.g., general infrastructure networks represented as acyclic directed networks. Among analytical methods, one brute force approach is to enumerate all possible combinations of component states and determine their outcomes accordingly. Because the computational cost increases exponentially with the size of the system, total enumeration becomes computationally intractable for large networks.

To address this issue, efforts have been made to improve the computational efficiency and storage requirements of enumeration-based methods. Tien and Der Kiureghian (2016) and Tong and Tien (2017) proposed compression and inference algorithms to facilitate inference in binary and multistate Bayesian network representations of infrastructure systems. However, computational limits on system size still exist, particularly for networks with a large number of parent nodes (Tien and Der Kiureghian 2017). Ebeling (2010) determined the reliability bounds by taking all MLS in parallel (survival of any link set yields survival of the system) and all MCS in series (failure of any cut set yields failure of the system). The gap between the upper and lower bounds can be wide, however, because of the independency assumption among all MLS and MCS. In addition, the use of both Bayesian networks and the bound-finding method require identification of the MLS or MCS of a system. Although methods have been proposed to do this efficiently, e.g., the edge cut algorithm, also known as EG-CUT algorithm, to generate MCS (Shin and Koh 1998) and a recursive depth-first search MLS identification algorithm (Applegate and Tien 2019), this is a NP-hard problem (Suh and Chang 2000) for general networks.

Recursive decomposition algorithm (RDA) (e.g., Dotson and Gobien 1979) provides an alternative analytical approach. Lim and Song (2012) proposed selective RDA to improve the efficiency of original RDA by identifying the most reliable paths. However, the computational cost may still exponentially increase, particularly when the most reliable paths are not significantly more dominant than others. Kim and Kang (2013) extended application of RDA to multiple-source-multiple-sink situations. Multiple sinks are connected to one aggregated sink node, however, which loses the reliabilities of individual sink nodes. Similar limitations are found in Liu and Li (2009). In essence, the focus is still on a one-sink node network. In addition, the number of subgraphs created by eliminating the failed components still increases exponentially with the system size.

For simulation-based methods, it is often a challenge to efficiently capture the occurrence of rare events. To address this, several sampling methods, e.g., refined stratified sampling strategy (Shields et al. 2015), rejection sampling from predetermined distribution (Cheng et al. 2015), and random walks on graphs (Cheng et al. 2017) have been proposed. Although the sampling size and efficiency have been improved, challenges still exist in determining the system outcome. In comparison, the efficiency and accuracy of PrPm (Tong and Tien 2019) and dPrPm as described in this paper are not limited by the ability to assess low-probability events. In Bulteau and El Khadiri (1998) and Zuev et al. (2015), the result of a sampling point depends on an indicator function, e.g., by comparison with MLS or MCS, which can be expensive to evaluate. One advantage of the proposed directed probability propagation method is the absence of an indicator function or need to identify the MLS or MCS of a system.

The idea of PrPm and dPrPm originated from belief propagation. Details of belief propagation for network graphs were provided by Coughlan (2009) and Barber (2012). In the present paper, the message that is passed from node to node refers to the marginal and pairwise node reliabilities. To increase computational efficiency, we do not consider the full nodal joint distributions. Accuracy is guaranteed through definition of the propagation sequence to create intermediate structures for final inference and the corresponding probability updating rules.

PrPm was first introduced by Tong and Tien (2019). The method described in the present paper advances upon that work in three main ways. First, although PrPm could be used for single-source-single-sink networks and multiple-sources-single-sink networks, the case of multiple-sources-multiple-sinks networks was not addressed. In an infrastructure system, the multiple sinks scenario is particularly relevant to calculate probabilities of providing the infrastructure resource at multiple end-point distribution nodes. The directed PrPm, i.e., dPrPm, proposed in this paper addresses the multiple-sinks case and provides reliabilities at all sink nodes in the network. Second, although PrPm provided approximations of the network reliability that were shown to be accurate compared to the exact solution, there were no guarantees of the accuracy of the estimates, nor were there bounds that could be placed on the results. Results were empirically shown to be accurate rather than being able to provide proofs of their accuracy. dPrPm provides 100% confidence bounds on the reliabilities at sink nodes and proofs are provided guaranteeing their accuracy. Third, this paper focuses on reliability assessment of acyclic directed networks as applicable to infrastructure systems. In these systems, resources are distributed through a network from supply or generation nodes to distribution nodes without cycles in the network. The newly described propagation sequences and probability updating rules reflect this case.

Proposed Directed Probability Propagation Method

Networks of Interest

Due to the characteristics of flows in networks, we consider the links connecting the nodes in the system as directed. The reliability of a sink node in the system is defined as its accessibility from a source node. The links in the system are assumed to be independent or conditionally independent of each other. In our proposed method, dPrPm, there is no limitation on the number of source nodes and sink nodes. An example 16-node network is shown in Fig. 1, in which solid black circles are source nodes and all hollow circles are sink nodes. Arrows on links indicate directionality of

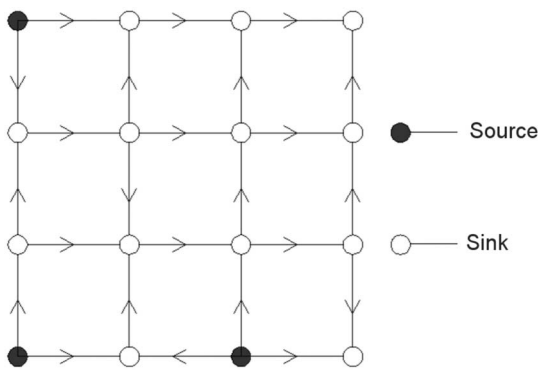

Fig. 1. Example network of interest.

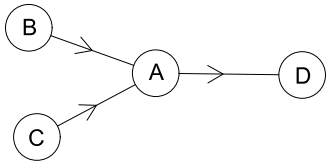

Fig. 2. Illustration of belief propagation.

flow. In the following sections, we use the network in Fig. 1 as an example to illustrate the procedure of dPrPm.

Overall Method

In belief propagation, messages are passed from node to node in the network to perform inference. An example is shown in Fig. 2, in which Node A receives the message from Node B and Node C. It updates the message and then passes the message to Node D. In the case in which the structure is a tree, belief propagation yields an exact marginal distribution.

In dPrPm, we start the message propagation from source nodes in the network. The message being passed here refers to the marginal node reliability $\Pr(\alpha)$ and pairwise node reliability $\Pr(\alpha, \beta)$. To begin, we remove all links in the network and keep only the source nodes [Fig. 3(a)]. Then we add link one at a time [Figs. 3(b and c)] to restore the original connectivity of the network. The configurations in Figs. 3(b and c) are referred to as intermediate structures. Every time a link is added, we update the message $\Pr(\alpha)$ and $\Pr(\alpha, \beta)$ for all nodes. In the end, once all links are added [Fig. 3(d)], the reliability of each sink node is obtained.

Heuristic Propagation Sequences for Acyclic Directed Networks

The message propagates through the network through a link-adding sequence. The defined sequence affects the accuracy of the result. When a link is added to create a new intermediate structure, e.g., adding the dashed link $l_{\alpha\to\beta_1}$ to connect Nodes α and β_1 [Figs. 4(a and b)], the reliabilities of multiple sink nodes are influenced. For example, in the case in Fig. 4, after link $l_{\alpha\to\beta_1}$ is added, additional paths to Nodes β_1, β_2, and β_3 are created. Thus, the message relating to these three nodes (β_1, β_2, and β_3) needs updating. However, we cannot obtain exact values for these updating terms because the message inherited from the previous step only includes the marginal and pairwise node reliabilities. Approximations are necessary for estimating a three-node joint distribution.

To yield a more accurate solution, the objective is to make fewer approximations during message propagation. A heuristic for making fewer approximations is to limit the number of nodes influenced when links are added. As an example, in Fig. 5, Node β is the only node influenced after adding link $l_{\alpha\to\beta}$. Thus, we need only to update the message relating to Node β.

For an acyclic directed network, we can find certain link-adding sequences that ensure that only one node is influenced every time a link is added. To illustrate this, we classify all nodes into one of three types in each intermediate step:

Type A node: in the intermediate structure, all links connected to the node have been added.

Type B node: in the intermediate structure, all incoming links to the node have been added.

Type C node: all other nodes.

Fig. 6(a) shows the original connections in the network; Figs. 6(b and c) are intermediate structures. Node types are marked next to each node. At the start, no links are directed toward the source nodes, i.e., they have no incoming links that need to be added [Fig. 6(b)]. The source nodes have outgoing links that have not yet been added. Therefore, all source nodes belong to Type B. In Fig. 6(c), all links connected to Type A nodes have been added. All incoming links to Type B nodes have been added. The remaining nodes are Type C.

We now provide two proofs showing the ability to define propagation sequences for acyclic directed networks such that only one

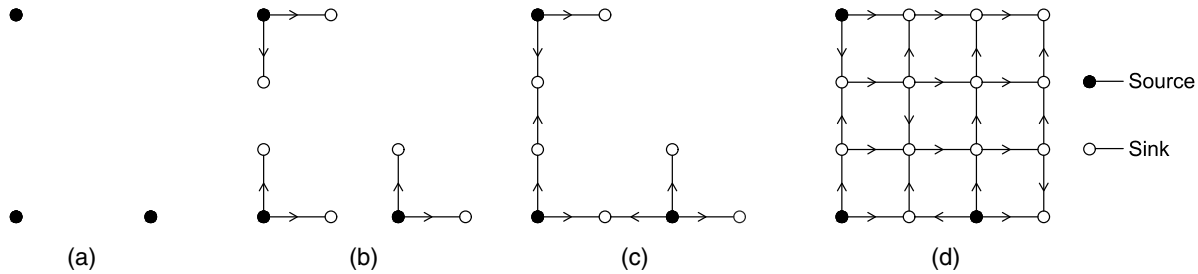

Fig. 3. Constructing intermediate structures for directed probability propagation method (dPrPm) to obtain reliabilities at all sink nodes.

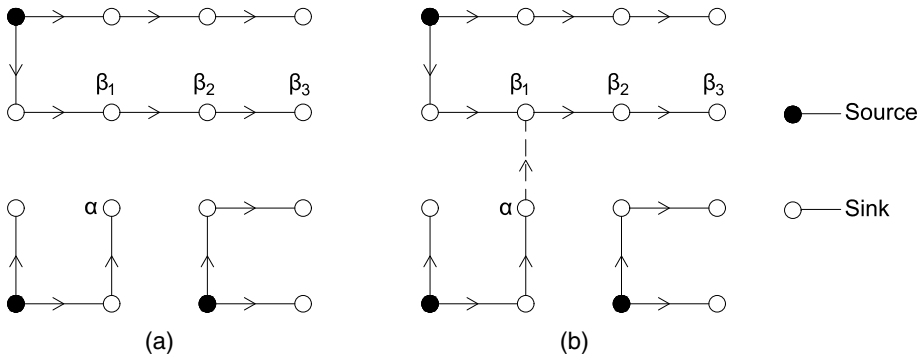

Fig. 4. Adding link $l_{\alpha\to\beta_1}$ to form a new intermediate structure.

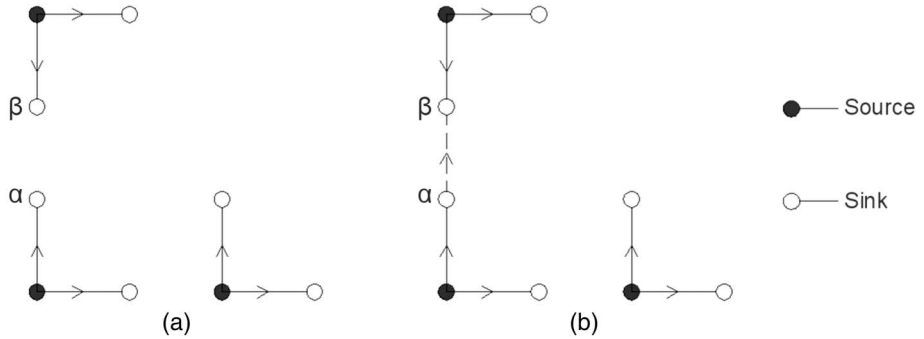

Fig. 5. Adding link heuristically to reduce number of nodes influenced.

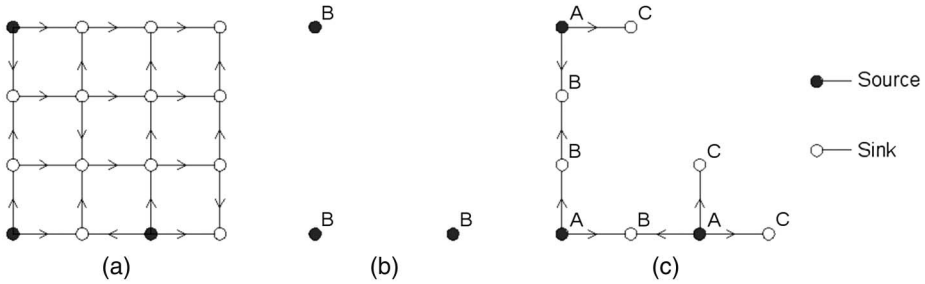

Fig. 6. Example node classifications in the network: (a) original connections; and (b and c) intermediate structures.

node is influenced at each step of the link-adding sequence. This is done by always adding links originating from Type B nodes.

Theorem 1. All intermediate structures have at least one Type B node.

Proof. Suppose that for a given intermediate structure, all nodes belong to either Type A or Type C. Let N be the number of Type C nodes [Fig. 7(a)], with the Type C nodes in the structure indicated by the set $\{\alpha_1, \ldots, \alpha_N\}$. Therefore, $N \geq 1$; otherwise, all nodes are Type A and all links have been added. Based on the definition of a Type C node, there is a path $pa_{\beta_1 \to \alpha_1}$ directed to α_1. Because the network is acyclic, β_1 cannot be α_1 and we set it as α_2. Similarly, for Node α_2, there is a path $pa_{\beta_2 \to \alpha_2}$ directed to α_2. For the acyclic network, β_2 cannot be α_1 or α_2 and we set it as α_3. Continuing the deduction to α_N, there is a path $pa_{\beta_N \to \alpha_N}$ directed to α_N. Because $\beta_N \in \{\alpha_1, \ldots, \alpha_N\}$, it creates a loop, contradicting the acyclic assumption. Thus, at least one of the nodes in set $\{\alpha_1, \ldots, \alpha_N\}$ belongs to Type B.

Theorem 2. If we prioritize the links that originate from Type B nodes in adding sequences, only one node is influenced every time a link is added to the intermediate structure.

Proof. Suppose a link $l_{\alpha \to \beta}$ is added [Fig. 7(b)]. Node α is Type B, and Node β must be Type C. This is because if we are adding a link into β, then clearly not all links into β have been added, and β is by definition a Type C node. Because we do not add links originating from Type C nodes (in the example, Node β), then β has no outgoing links and no node can be reached from Node β in the current intermediate structure. Thus, Node β is the only node influenced.

In sum, for acyclic directed networks, there is always a link-adding sequence for which only one node is influenced for each added link, and hence which makes the fewest approximations. Multiple link-adding sequences can satisfy the preceding requirements.

Message Propagation: Probability Updating Rules

As we add links and create intermediate structures, updates of the message should be made simultaneously. In dPrPm, the updating rules are discussed under two different scenarios (Fig. 8). Link $l_{\alpha \to \beta}$ is added into the intermediate structure and Node γ is a random node

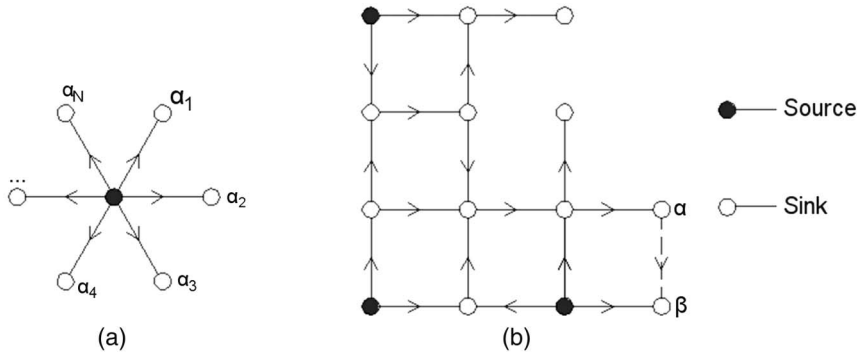

Fig. 7. Intermediate structures for acyclic directed network proofs.

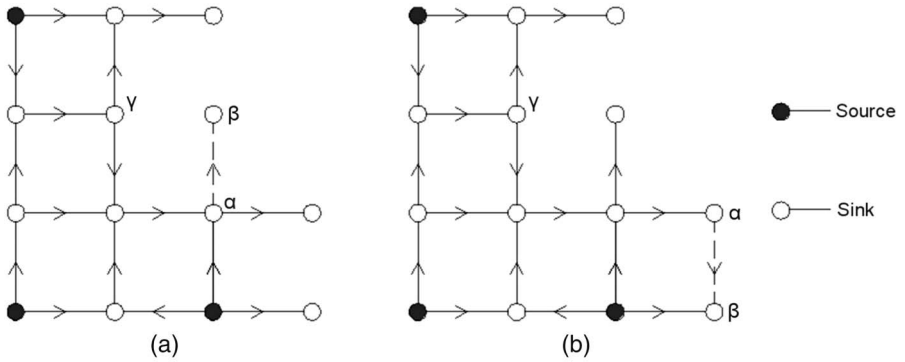

Fig. 8. Two updating scenarios for added link $l_{\alpha\to\beta}$.

other than Nodes α or β. In Fig. 8(a), β is not in the previous intermediate structure; in Fig. 8(b), it is. In both scenarios, it is easy to see that we need only to update messages relating to Node β. For notation, $\overline{\Pr}$ refers to terms after updating and Pr refers to terms before updating. $R_{\alpha\to\beta}$ is the link reliability between Node α and Node β.

In the first scenario, because Node β is not in the previous intermediate structure, no approximation is needed [Fig. 8(a)]. The updating rules are

$$\overline{\Pr(\beta)} = \Pr(\alpha)R_{\alpha\to\beta} \tag{1}$$

$$\overline{\Pr(\alpha,\beta)} = \Pr(\alpha)R_{\alpha\to\beta} \tag{2}$$

$$\overline{\Pr(\gamma,\beta)} = \Pr(\alpha,\gamma)R_{\alpha\to\beta} \tag{3}$$

In the second scenario, Node β is already part of the previous intermediate structure [Fig. 8(b)]. To calculate the updating rules, first we construct the three-node joint distribution as shown in Table 1, which gives the probability shares for each combination of nodal states; X represents the probability share that all three nodes, α, β, and γ, are reachable, i.e., $\Pr(\alpha,\beta,\gamma)$. The remaining terms are inferred from the marginal and pairwise node distributions in the message. In Table 1, for the states of α, β, and γ, 1 denotes that the node is reachable, and 0 otherwise.

$$X_1 = 1 - \Pr(\alpha) - \Pr(\beta) - \Pr(\gamma) + \Pr(\alpha,\gamma) + \Pr(\alpha,\beta) + \Pr(\beta,\gamma) - X$$

$$X_2 = \Pr(\gamma) - \Pr(\alpha,\gamma) - \Pr(\beta,\gamma) + X$$

$$X_3 = \Pr(\beta) - \Pr(\beta,\gamma) - \Pr(\alpha,\beta) + X$$

$$X_4 = \Pr(\beta,\gamma) - X$$

$$X_5 = \Pr(\alpha) - \Pr(\alpha,\gamma) - \Pr(\alpha,\beta) + X$$

Table 1. Three-node joint distribution

α	β	γ	Probability shares
0	0	0	X_1
0	0	1	X_2
0	1	0	X_3
0	1	1	X_4
1	0	0	X_5
1	0	1	X_6
1	1	0	X_7
1	1	1	X

$$X_6 = \Pr(\alpha,\gamma) - X$$

$$X_7 = \Pr(\alpha,\beta) - X$$

$$X = \Pr(\alpha,\beta,\gamma)$$

The updating rules for the scenario in Fig. 8(b), including the unknown parameter X, are

$$\overline{\Pr(\beta)} = \Pr(\beta) + (\Pr(\alpha) - \Pr(\alpha,\beta))R_{\alpha\to\beta} \tag{4}$$

$$\overline{\Pr(\alpha,\beta)} = \Pr(\alpha,\beta) + (\Pr(\alpha) - \Pr(\alpha,\beta))R_{\alpha\to\beta} \tag{5}$$

$$\overline{\Pr(\beta,\gamma)} = \Pr(\beta,\gamma) + (\Pr(\alpha,\gamma) - X)R_{\alpha\to\beta} \tag{6}$$

To calculate the exact value of X, we would need the full three-node joint distribution. For computational and memory storage efficiency, we only focus on the marginal and pairwise node reliabilities and do not carry three-node joint distributions in the message. However, we are able to bound the value of X based on the following derivations from probability properties.

First, all the terms in the Probability Shares column in Table 1 should be greater than or equal to zero. Thus, let

$$X_1 = \max\{\Pr(\alpha,\gamma) + \Pr(\beta,\gamma) - \Pr(\gamma), \Pr(\alpha,\beta) + \Pr(\beta,\gamma) - \Pr(\beta), \Pr(\alpha,\gamma) + \Pr(\alpha,\beta) - \Pr(\alpha)\} \tag{7}$$

$$X_2 = \min\{\Pr(\alpha,\beta), \Pr(\beta,\gamma), \Pr(\alpha,\beta), 1 - \Pr(\alpha) - \Pr(\beta) - \Pr(\gamma) + \Pr(\alpha,\beta) + \Pr(\alpha,\gamma) + \Pr(\beta,\gamma)\} \tag{8}$$

Then the bounds of X are

$$X_1 \le X \le X_2 \tag{9}$$

Second, by the definition of conditional probability and the coherency property of networks

$$X = \Pr(\alpha,\beta,\gamma) = \Pr(\alpha,\beta|\gamma)\ \Pr(\gamma) \ge \Pr(\alpha,\beta)\Pr(\gamma) \tag{10}$$

Let

$$X_3 = \max\{\Pr(\alpha,\beta)\ \Pr(\gamma), \Pr(\alpha,\gamma)\ \Pr(\beta), \Pr(\beta,\gamma)\ \Pr(\alpha)\} \le X \tag{11}$$

Similarly, $\Pr(\alpha=0,\beta=0,\gamma) = \Pr(\alpha=0,\beta=0|\gamma)\Pr(\gamma) \le \Pr(\alpha=0,\beta=0)\ \Pr(\gamma)$.

Thus, $\Pr(\alpha=0,\beta=0,\gamma)=\Pr(\gamma)-\Pr(\alpha,\gamma)-\Pr(\beta,\gamma)+X\le \Pr(\alpha=0,\beta=0)\ \Pr(\gamma)$.

Let $t_1=\Pr(\alpha=0,\beta=0)\Pr(\gamma)-\Pr(\gamma)+\Pr(\alpha,\gamma)+\Pr(\beta,\gamma)$, $t_2=\Pr(\alpha=0,\gamma=0)\Pr(\beta)-\Pr(\beta)+\Pr(\alpha,\beta)+\Pr(\beta,\gamma)$, and $t_3=\Pr(\beta=0,\gamma=0)\Pr(\alpha)-\Pr(\alpha)+\Pr(\alpha,\beta)+\Pr(\alpha,\gamma)$.

Then

$$X\le X_4=\min\{t_1,t_2,t_3\} \tag{12}$$

Combining the inequality constraints in Eqs. (9), (11), and (12)

$$\max\{X_1,X_3\}\le X\le \min\{X_2,X_4\} \tag{13}$$

Thus, we obtain the upper bound and lower bound of X. These derivations enable us to guarantee bounds on the node reliability values calculated through dPrPm as shown in the following:

Theorem 3. If we assign the lower bound value to X every time we update the message, we will obtain the lower bound values of marginal node reliabilities. If we assign the upper bound value to X every time we update the message, we will obtain the upper bound values of marginal node reliabilities.

Proof. We prove this theorem separately for the two updating scenarios in Fig. 8. For the scenario in Fig. 8(a), no approximation is made. As a result, the updated terms from Eqs. (1)–(3) are not influenced by the estimation of X.

For the scenario in Fig. 8(b), if we take the lower bound value of X, then based on Eq. (6), $\overline{\Pr(\beta,\gamma)}$ is overestimated. In the future updating steps, when link $l_{\gamma\to\beta}$ is added, Eq. (4) becomes $\overline{\Pr(\beta)}=\Pr(\beta)+(\Pr(\gamma)-\Pr(\beta,\gamma))R_{\gamma\to\beta}$; $\overline{\Pr(\beta)}$ is then underestimated at the lower bound because of the overestimation on $\Pr(\beta,\gamma)$. The same logic applies to taking the upper bound value of X, which yields an overestimation and upper bound on marginal node reliabilities.

The bounds can be further refined considering the possible link-adding sequences satisfying the requirement that only one node be influenced at each link-adding step. Suppose there are N link-adding sequences available for a general acyclic directed network. For the ith sequence, we bound the marginal node reliability of Node α as: $\Pr(\underline{\alpha})_i\le\Pr(\alpha)\le\Pr(\bar{\alpha})_i$. Combining all N sequences gives a refined boundary as

$$\max_{i=1\ldots N}\{\Pr(\underline{\alpha})_i\}\le\Pr(\alpha)\le\min_{i=1\ldots N}\{\Pr(\bar{\alpha})_i\} \tag{14}$$

Memory Storage and Computational Complexity Analysis

For memory storage cost, because the message in dPrPm only includes marginal and pairwise node reliabilities, the storage cost is $O(n^2)$.

For computational cost, when adding link $l_{\alpha\to\beta}$ into the intermediate structure, we need only to update the message relating to Node β. Thus, to add one link and update the message, at most n nodes are updated. Thus, the computational cost is $O(n)$. In total, there are m links, which requires updating m times. Thus, the total computational cost is polynomial at $O(mn)$. This is in comparison to typical exponential computational costs for binary networks with n components at $O(2^n)$.

In sum, the workflow of dPrPm is shown in Fig. 9.

Test Applications

We applied the proposed dPrPm to an example network and two real-world applications. We assessed the performance of the method in terms of accuracy and computational cost. All results were based on computations run in MATLAB_R2017b on a computer with 8 GB RAM. For all three applications, results were compared with Monte Carlo simulation. The first two examples also included exact solutions for comparison. A more complex gas pipeline network, for which the exact solution is not available, was analyzed in the third example under two seismic scenarios. Both accuracy and computational efficiency were compared to assess the performance of dPrPm.

Directed Grid Network

First, we applied dPrPm to the example network shown in Fig. 10, also used for illustration of the method earlier in this paper. The network was a variation of the undirected grid network that was studied by Tong and Tien (2019) and Dueñas-Osorio (2017). Instead of looking only at the corner-to-corner reliability of the system, in this example we changed the single-source-single-sink network into a multiple-sources-multiple-sinks directed acyclic network. Nodes 1, 3, and 13 were taken as source nodes, and the remaining nodes were sink nodes. The objective was to obtain reliability estimates for the probabilities of being able to receive the resource provided at the source nodes at each of the sink nodes.

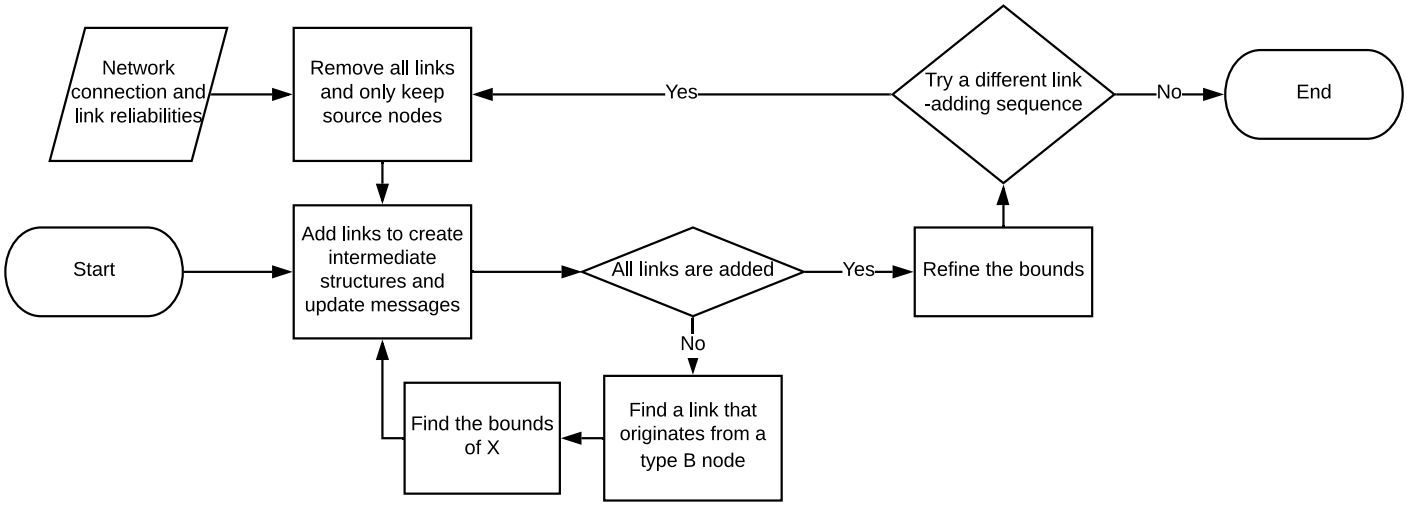

Fig. 9. dPrPm workflow.

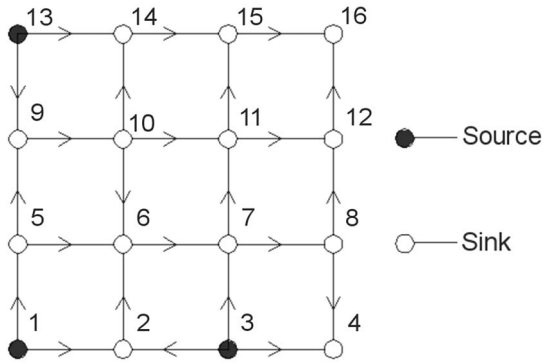

Fig. 10. Directed grid network.

For simplicity, all links were assumed to be equally reliable, with reliability R_l. Varying link reliabilities can be easily incorporated in the probability updating calculations during message passing.

We tested the performance of dPrPm under two cases: a reliable case where $R_l = 0.9$, and an unreliable case in which $R_l = 0.1$. Exact solutions were obtained through total enumeration. For dPrPm, we tried 100 different link-adding sequences to refine the reliability bounds for each sink node. Percentage error for dPrPm was calculated by taking the median value between the upper and lower bounds. For Monte Carlo simulation, we generated 100,000 runs for each node. Results are shown in Tables 2 and 3. The computational cost is compared in Table 4.

Compared with the approximated value given by Monte Carlo simulation, dPrPm provides exact upper bounds and lower bounds to the solution. For both the reliable and unreliable cases, dPrPm outperformed Monte Carlo, as evidenced by the percentage error values compared with the exact solutions (Tables 2 and 3). In addition, in terms of the ability to capture rare events, the reliability at Node 16 for $R_l = 0.1$ was the lowest-probability event. In Table 3, the percentage error given by Monte Carlo simulation increased to 14.7% for Node 16, whereas the upper and lower bounds given by dPrPm were still narrow. For dPrPm, the percentage error arises from deviations in the calculated reliability beyond the fifth decimal place.

Although total enumeration provides the exact solution, it is also computationally intensive. This example took more than 1.5 h to generate a solution (Table 4). In comparison, dPrPm took less than 0.4 s to compute the dPrPm solution 100 times. This is also two orders of magnitude faster than Monte Carlo simulation. Although dPrPm yields an approximated solution with a 100% confidence level rather than the exact solution, the gap between the bounds can be negligible depending on the accuracy requirement, and as shown in the results for this example.

Power Distribution Network

Next, we applied the proposed dPrPm to a four-substation power distribution network from Pacific Gas and Electric (Ostrom 2004) (Fig. 11), which also was investigated by Der Kiureghian and Song (2008) and Tien (2017). The system consisted of circuit breakers,

Table 2. Exact solutions, Monte Carlo simulation, and dPrPm for directed grid network when $R_l = 0.9$

		Monte Carlo		dPrPm			
Node	Exact	Result	Error (%)	Lower bound	Upper bound	Gap	Error (%)
2	0.99000	0.99016	0.01616	0.99000	0.99000	0.00000	0.00000
4	0.98015	0.98022	0.00670	0.98015	0.98015	0.00000	0.00000
5	0.90000	0.89885	−0.12778	0.90000	0.90000	0.00000	0.00000
6	0.99510	0.99475	−0.03545	0.99510	0.99510	0.00000	0.00000
7	0.98956	0.98974	0.01827	0.98956	0.98956	0.00000	0.00000
8	0.89060	0.89086	0.02882	0.89060	0.89060	0.00000	0.00000
9	0.98100	0.98138	0.03874	0.98100	0.98100	0.00000	0.00000
10	0.88290	0.88355	0.07362	0.88290	0.88290	0.00000	0.00000
11	0.97734	0.97758	0.02416	0.97734	0.97734	0.00000	0.00000
12	0.97457	0.97456	−0.00054	0.97457	0.97457	0.00000	0.00000
14	0.97946	0.97889	−0.05830	0.97946	0.97946	0.00000	0.00000
15	0.98498	0.98544	0.04671	0.98498	0.98498	0.00000	0.00000
16	0.98539	0.98592	0.05393	0.98506	0.98558	0.00048	−0.00532

Table 3. Exact solutions, Monte Carlo simulation, and dPrPm for directed grid network when $R_l = 0.1$

		Monte Carlo		dPrPm			
Node	Exact	Result	Error (%)	Lower bound	Upper bound	Gap	Error (%)
2	0.19000	0.18899	−0.53158	0.19000	0.19000	0.00000	0.00000
4	0.10092	0.10160	0.66962	0.10092	0.10092	0.00000	0.00000
5	0.10000	0.10191	1.91000	0.10000	0.10000	0.00000	0.00000
6	0.02986	0.03034	1.60529	0.02986	0.02986	0.00000	0.00000
7	0.10269	0.10242	−0.26046	0.10269	0.10269	0.00000	0.00000
8	0.01027	0.01057	2.93370	0.01027	0.01027	0.00000	0.00000
9	0.10900	0.10970	0.64220	0.10900	0.10900	0.00000	0.00000
10	0.01090	0.01169	7.24771	0.01090	0.01090	0.00000	0.00000
11	0.01135	0.01145	0.91177	0.01135	0.01135	0.00000	0.00000
12	0.00215	0.00198	−7.95624	0.00215	0.00215	0.00000	0.00000
14	0.10098	0.09956	−1.40720	0.10098	0.10098	0.00000	0.00000
15	0.01122	0.01177	4.89889	0.01122	0.01122	0.00000	0.00000
16	0.00134	0.00114	−14.65189	0.00134	0.00134	0.00000	−0.00041

Table 4. Computational cost of exact solution, Monte Carlo simulation, and dPrPm for directed grid network

Method	Time (s)
Exact solution	5,766.37
Monte Carlo	29.88
dPrPm	0.32

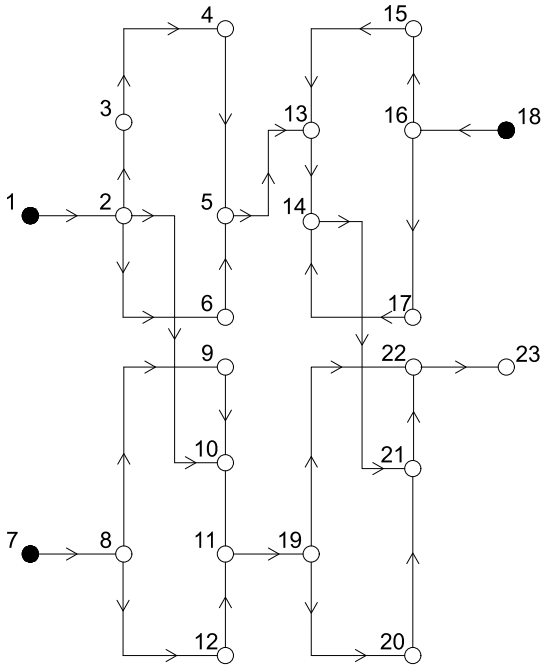

Fig. 11. Power distribution network.

switches, and transformers. The solid black circles represent three source nodes, and the sink node is Node 23. The other 19 components each represent a triplet configuration in the system of switch-breaker-switch.

Although dPrPm provides the reliability at all nodes in the network, for simplicity, we only focused on the reliability of Node 23 in this example. We assumed no nodal failure, and each link had the same reliability R_l. Varying link reliabilities can easily be incorporated. For comparison, the method of total enumeration was used to obtain the exact solution. The network was tested under three scenarios for systems of varying reliabilities: $R_l = 0.9$, 0.99, and 0.3. Monte Carlo simulation was also conducted, with 100,000 realizations for each value of R_l (Table 5). Computational cost is compared in Table 6.

dPrPm provided a highly accurate solution, with a maximum percentage error less than 0.05% for all three scenarios (Table 5). For the lowest-probability event, $R_l = 0.3$, Monte Carlo simulation had the highest error, 8.63%, whereas dPrPm produced a solution with 0.00295% error (Table 5). In terms of computation time, dPrPm ran almost 100 times faster than Monte Carlo simulation (Table 6).

In this example, we also investigated the effect of using results from multiple link-adding sequences to refine the bounds for dPrPm [Eq. (14)]. Theoretically, the total number of link-adding sequences satisfying the requirements previously described is finite. As a result, the best solution by dPrPm can be obtained by considering all of them. Fig. 12 shows the evolution of the bounds as the number of sequences considered increases. The gap between the upper bound and lower bound dropped quickly in the first

Table 5. Exact solutions, Monte Carlo simulation, and dPrPm for power distribution network at different R_l

		Monte Carlo		dPrPm			
R_l	Exact	Result	Error (%)	Lower bound	Upper bound	Gap	Error (%)
0.9	0.85741	0.85805	0.07490	0.85591	0.85807	0.00216	−0.04886
0.99	0.98969	0.98939	−0.02981	0.98968	0.98969	0.00001	−0.00018
0.3	0.00221	0.00240	8.63048	0.00221	0.00221	0.00000	0.00295

Table 6. Computational cost of exact solution, Monte Carlo simulation, and dPrPm for power distribution network

Method	Time (s)
Exact solution	36,864.37
Monte Carlo	35.97
dPrPm	0.36

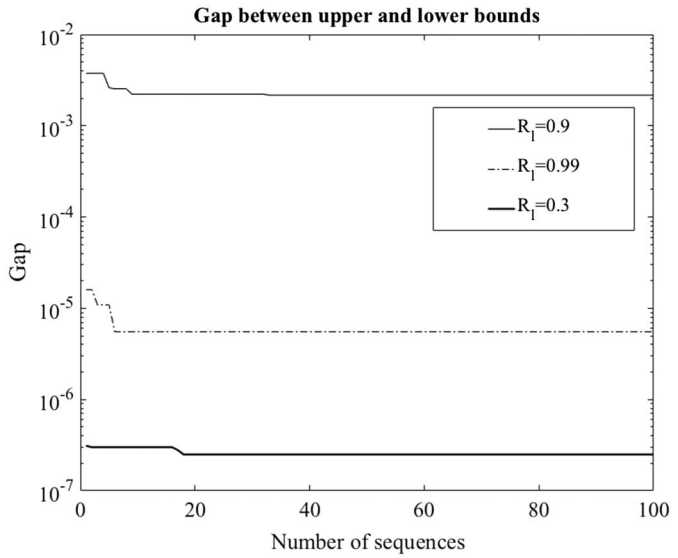

Fig. 12. Gap between upper and lower bounds with respect to number of link-adding sequences considered.

20 runs for all three cases (Fig. 12). The size of the gap then leveled off. Enumeration of all possible link-adding sequences is computationally intensive. Therefore, rather than finding all possible sequences, we set the number of sequences to 100 to balance accuracy and efficiency.

Gas Pipeline Network

Finally, dPrPm was tested for a larger and more complex gas pipeline network. The pipeline information for the Los Angeles area is available from the California Energy Commission's GIS open data website (California Energy Commission 2018). This network previously was studied to assess reliability of buried pipelines under earthquakes (e.g., Lanzano et al. 2014; Ambraseys and Menu 1998). The satellite map of the investigated area is shown in Fig. 13. The extracted layout of the network is shown in Fig. 14, with pipelines between nodes represented by straight lines, and arrows indicating the directionality of the links. We analyzed the pipeline network reliability under two scenarios: one moderate earthquake with peak

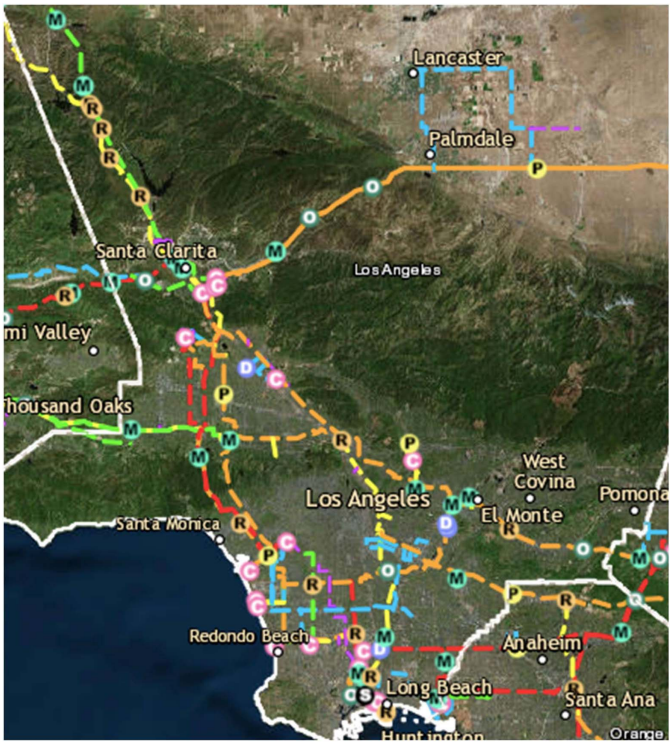

Fig. 13. Gas pipeline network (satellite) adapted from California Energy Commission's GIS open data website.

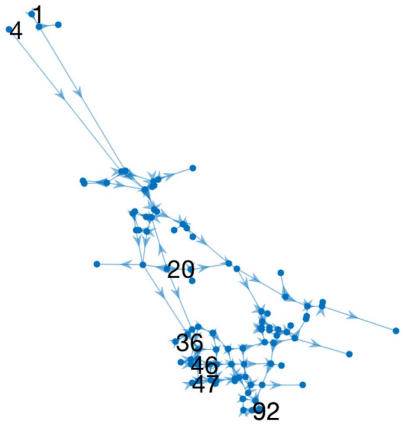

Fig. 14. Gas pipeline network (extracted).

ground acceleration (PGA) $0.1g$, and a more severe earthquake with PGA $0.35g$. The network had 96 nodes and 123 directed links. We assumed that Nodes 1, 4, 20, 92, 36, 46, 47, and 92 were source nodes (Fig. 14). The remaining nodes were considered as sink nodes. In total, there were 8 source nodes and 88 sink nodes.

In this example, we first found the link reliabilities under the two seismic scenarios as described by Lanzano et al. (2014) and Ambraseys and Menu (1998). Reliabilities at each sink node were then obtained by running dPrPm 100 times and Monte Carlo simulation with 100,000 runs under the two earthquakes. The results of the two methods are compared in Figs. 15 and 16 for PGAs of $0.1g$ and $0.35g$, respectively. For clarity, nodes are ordered by increasing reliability. The exact solution is not available for this network. In all plots, the bold solid line denotes the upper bound given by dPrPm. The thin solid line denotes the lower bound. The dashed line denotes the result given by Monte Carlo simulation. In Fig. 15, for a $0.1g$ earthquake, the upper bound and lower bound overlap, indicating highly confined bounds, and therefore only the upper bound line can be seen. Figs. 15(a) and 16(a) show the difference between the dPrPm reliability bounds and results from Monte Carlo. The median value between the upper and lower dPrPm bounds was used as the baseline, and the difference from the median value is shown. Figs. 15(b) and 16(b) show the values of the upper and lower bounds by dPrPm and the solution by Monte Carlo simulation.

In Fig. 15, the narrowness of the bounds obtained by dPrPm is observed. With the guaranteed accuracy of dPrPm, the value of the bounds can be taken as close to the exact solution. In addition, for the approximated results provided by Monte Carlo, it is unknown if the simulations underestimate or overestimate the exact solution, as can be seen by the randomness in the positive or negative differences from the median value in Fig. 15(a). In Fig. 16, the widest gap between bounds for all nodes was about 1.6%. Whereas dPrPm gave guaranteed upper and lower bounds to the solution, the 100% confidence level was unachievable for Monte Carlo simulation. Table 7 compares the computation time for calculation. Consistent with the previous applications, dPrPm took two orders of magnitude less computation time to obtain a result.

Extension to Dependent Case and Cascading Failures

Although dPrPm is designed for networks with conditionally independent links, it can also be extended to dependent cases. One approach to do this is to condition on the parent nodes governing the links. For example, in the gas pipeline network application, links were conditioned on seismic intensity. Because the computational cost is fairly low (in this case, 1.80 s for a network of 123 links), the prior system reliability can be found by conducting inference over all enumerated parental node combinations, if tractable.

The special dependency case in which the link reliability $R_{\alpha\to\beta}$ depends on the state of Node α is also considered. When link $l_{\alpha\to\beta}$ is added to the intermediate structure, the updating rules in Eqs. (1)–(6) assume that the link reliability $R_{\alpha\to\beta}$ is independent of the node state. By replacing $R_{\alpha\to\beta}$ with the link reliability conditioned on the node state ($R_{\alpha\to\beta|\alpha}$), similar updating rules for this special dependent case are built. For example, Eq. (3) is modified to $\overline{\Pr(\gamma,\beta)} = \Pr(\alpha,\gamma)R_{\alpha\to\beta|\alpha=1}$, and Eq. (6) is modified to $\overline{\Pr(\beta,\gamma)} = \Pr(\beta,\gamma) + (\Pr(\alpha,\gamma) - X)R_{\alpha\to\beta|\alpha=1}$. These modified updating rules enable dPrPm to directly address this case.

In practice, the failures of many networked infrastructure systems are the result of failures propagating or cascading through a network. To capture these effects with dPrPm, we treat a failure as an observation of the system. Based on the observation, we update the link reliabilities accordingly. With the redefined link reliabilities, we then rerun dPrPm to update the reliabilities at all sink nodes. Any new failures that are calculated are the cascaded result from previous node failures.

Conclusion

This paper proposed a new analytical method called the directed probability propagation method (dPrPm) to evaluate the reliability of acyclic directed networks. Through a defined propagation sequence and accompanying probability updating rules, the method results in guaranteed upper and lower bounds of reliabilities at all sink nodes in a network. The benefits of dPrPm are summarized in five aspects

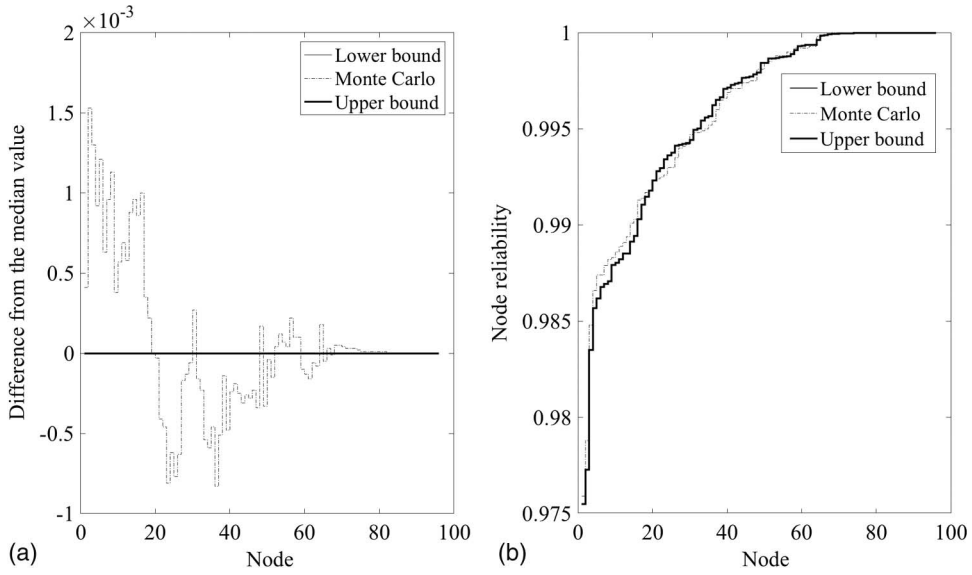

Fig. 15. dPrPm and Monte Carlo simulation of gas pipeline network at PGA $= 0.1g$ (nodes ordered by increasing reliability).

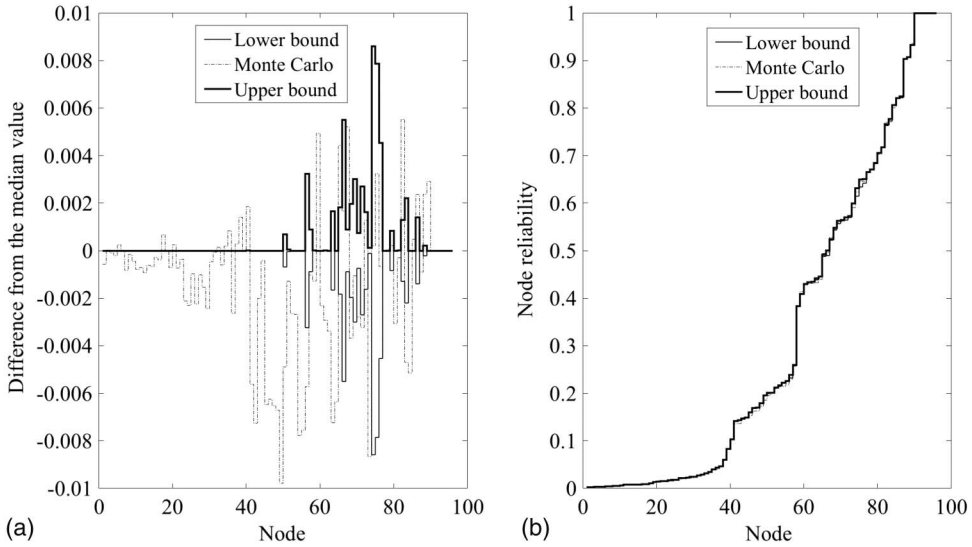

Fig. 16. dPrPm and Monte Carlo simulation of gas pipeline network at PGA $= 0.35g$ (nodes ordered by increasing reliability).

Table 7. Computational cost of Monte Carlo simulation and dPrPm for gas pipeline network

Method	Time (s)
Monte Carlo	180.22
dPrPm	1.80

1. No minimum cut sets or minimum link sets are needed to compute network reliabilities. Whereas many other analytical methods rely on MLS or MCS, e.g., Bayesian network analysis or recursive decomposition algorithms (RDA), dPrPm does not require the computationally intensive enumeration of component states, MCS, or MLS to determine the system outcome.
2. The method is applicable to the multiple-sources-multiple-sinks problem. Previous work investigated node accessibility as a measure of network reliability. However, these methods are often limited to the one-sink problem. To assess the reliability of all sink nodes in the network, researchers have to run the analysis multiple times. In dPrPm, because the message contains the marginal node reliabilities, each run gives the reliabilities of all sink nodes.
3. dPrPm is computationally efficient. Compared with existing analytical algorithms such as RDA and inference in Bayesian networks, computational complexity is reduced from an exponential increase $O(2^n)$ with system size to a polynomial increase $O(mn)$. Time consumption comparisons in the three test applications showed orders of magnitude savings in computation time.
4. Results given by dPrPm are exact bounds. Whereas many other methods, e.g., Monte Carlo simulation, give an approximated answer, a 100% confidence level is guaranteed by dPrPm. In many cases, as shown in the test applications, these bounds are narrow, resulting in solutions with very low percentage errors compared with the exact solution.
5. Performance of dPrPm is independent of link reliabilities. Many sampling-based approaches are limited in the ability to analyze rare events or by computational efficiency if the probabilities of

rare events are of interest. dPrPm is an analytical method, and its efficiency is independent of link or network reliabilities.

Acknowledgments

Support for this work by the National Science Foundation through Grant No. CNS-1541074 is acknowledged.

Notation

The following symbols are used in this paper:

$l_{\alpha\to\beta}$ = link connecting Node α to Node β;
m = number of links in network;
n = number of nodes in network;
α = lower case Greek letter refers to a node in the network;
A = upper case letter refers to node type;
$\overline{\Pr}$ = probability after updating;
$\Pr(\alpha)$ = probability that Node α is reachable;
$\Pr(\bar{\alpha})$ = upper bound of $Pr(\alpha)$;
$\Pr(\underline{\alpha})$ = lower bound of $Pr(\alpha)$;
$\Pr(\alpha,\beta)$ = probability that both Node α and Node β are reachable;
$\Pr(\alpha,\beta=0)$ = probability that Node α is reachable but Node β is not reachable;
$\Pr(\alpha|\beta)$ = probability that Node α is reachable given Node β is reachable;
$pa_{\alpha\to\beta}$ = path from Node α to Node β; and
$R_{\alpha\to\beta}$ = reliability of the link connecting Node α to Node β.

References

Ambraseys, N. N., and J. M. Menu. 1988. "Earthquake-induced ground displacements." *Earthquake Eng. Struct. Dyn.* 16 (7): 985–1006. https://doi.org/10.1002/eqe.4290160704.

Applegate, C., and I. Tien. 2019. "Framework for probabilistic vulnerability analysis of interdependent infrastructure systems." *J. Comput. Civ. Eng.* 33 (1): 04018058. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000801.

Barber, D. 2012. *Bayesian reasoning and machine learning*. Cambridge, UK: Cambridge University Press.

Birolini, A., 2004. *Reliability engineering: Theory and practice*. 4th ed. Berlin: Springer.

Bulteau, S., and M. El Khadiri. 1998. "A Monte Carlo simulation of the flow network reliability using importance and stratified sampling." Ph.D. dissertation, Dept. of RAIRO-Operations Research, INRIA.

California Energy Commission. 2018. "California energy commission: GIS open data." Accessed June 25, 2018. https://cecgis-caenergy.opendata.arcgis.com/.

Cheng, L., Z. Lu, and L. Zhang. 2015. "Application of rejection sampling based methodology to variance based parametric sensitivity analysis." *Reliab. Eng. Syst. Saf.* 142: 9–18. https://doi.org/10.1016/j.ress.2015.04.020.

Cheng, W., J. Cox, and P. Whitlock. 2017. "Random walks on graphs and Monte Carlo methods." *Math. Comput. Simul.* 135: 86–94. https://doi.org/10.1016/j.matcom.2015.12.006.

Coughlan, J. 2009. *A tutorial introduction to belief propagation*. San Francisco: Smith-Kettlewell Eye Research Institute.

Der Kiureghian, A., and J. Song. 2008. "Multi-scale reliability analysis and updating of complex systems by use of linear programming." *Reliab. Eng. Syst. Saf.* 93 (2): 288–297. https://doi.org/10.1016/j.ress.2006.10.022.

Dotson, W., and J. O. Gobien. 1979. "A new analysis technique for probabilistic graphs." *IEEE Trans. Circuits Syst.* 26 (10): 855–865. https://doi.org/10.1109/TCS.1979.1084573.

Dueñas-Osorio, L. 2017. "Reliability of grid networks and recursive decomposition algorithms." Accessed April 14, 2017. https://duenas-osorio.rice.edu/sites/g/files/bxs2181/f/docs/tutorial_grid25.html.

Ebeling, C. E. 2010. *An introduction to reliability and maintainability engineering*. 2nd ed. Long Grove, IL: Waveland.

Kim, Y., and W. Kang. 2013. "Network reliability analysis of complex systems using a non-simulation-based method." *Reliab. Eng. Syst. Saf.* 110: 80–88. https://doi.org/10.1016/j.ress.2012.09.012.

Lanzano, G., E. Salzano, F. Santucci de Magistris, and G. Fabbrocino. 2014. "Seismic vulnerability of gas and liquid buried pipelines." *J. Loss Prev. Process Ind.* 28: 72–78. https://doi.org/10.1016/j.jlp.2013.03.010.

Lim, H.-W., and J. Song. 2012. "Efficient risk assessment of lifeline networks under spatially correlated ground motions using selective recursive decomposition algorithm." *Earthquake Eng. Struct. Dyn.* 41 (13): 1861–1882. https://doi.org/10.1002/eqe.2162.

Liu, W., and J. Li. 2009. "An improved recursive decomposition algorithm for reliability evaluation of lifeline networks." *Earthquake Eng. Vib.* 8 (3): 409–419. https://doi.org/10.1007/s11803-009-8152-2.

Ostrom, D. 2004. "Database of seismic parameters of equipment in substations." Accessed January 10, 2017. https://apps.peer.berkeley.edu/lifelines/lifelines_pre_2006/final_reports/413-FR.pdf.

Shields, M. D., K. Teferra, A. Hapij, and R. P. Daddazio. 2015. "Refined stratified sampling for efficient Monte Carlo based uncertainty quantification." *Reliab. Eng. Syst. Saf.* 142: 310–325. https://doi.org/10.1016/j.ress.2015.05.023.

Shin, Y. Y., and J. S. Koh. 1998. "An algorithm for generating minimal cutsets of undirected graphs." *Korean J. Comput. Appl. Math.* 5 (3): 681–693. https://doi.org/10.1007/BF03008891.

Suh, H., and C. K. Chang. 2000. "Algorithms for the minimal cutsets enumeration of networks by graph search and branch addition." In *Proc. 25th Annual IEEE Conf. on Local Computer Networks*, 100–10. Piscataway, NJ: IEEE.

Tien, I. 2017. "Bayesian network methods for modeling and reliability assessment of infrastructure systems." In *Risk and reliability analysis: Theory and applications: Springer Series in Reliability Engineering*. Berlin: Springer.

Tien, I., and A. Der Kiureghian. 2016. "Algorithms for Bayesian network modeling and reliability assessment of infrastructure systems." *Reliab. Eng. Syst. Saf.* 156: 134–147. https://doi.org/10.1016/j.ress.2016.07.022.

Tien, I., and A. Der Kiureghian. 2017. "Reliability assessment of critical infrastructure using Bayesian networks." *J. Infrastruct. Syst.* 23 (4): 04017025. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000384.

Tong, Y., and I. Tien. 2017. "Algorithms for Bayesian network modeling, inference, and reliability assessment for multi-state flow networks." *J. Comput. Civ. Eng.* 31 (5): 04017051. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000699.

Tong, Y., and I. Tien. 2019. "Analytical probability propagation method for reliability analysis of general complex networks." *Reliab. Eng. Syst. Saf.* 189: 21–30. https://doi.org/10.1016/j.ress.2019.04.013.

Zuev, K., S. Wu, and J. Beck. 2015. "General network reliability problem and its efficient solution by Subset Simulation." *Probab. Eng. Mech.* 40: 25–35. https://doi.org/10.1016/j.probengmech.2015.02.002.