New Method for Complex Network Reliability Analysis through Probability Propagation

Yanjie Tong
Ph.D. Candidate, Dept. of Civil Engineering, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355, USA

Iris Tien
Assistant Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355, USA

ABSTRACT: Reliability analysis of complex networks is often limited by increasing dimensionality of the problem as the number of nodes and possible paths in the network increases. This is true particularly for reliability analysis problems that exponentially increase in computational requirements with system size. In this paper, we present a new method for complex network reliability analysis. We call this the probability propagation method (PrPm). The idea originates from the concept of belief propagation for inference in network graphs. In PrPm, the message passed between nodes is a joint probability distribution. At each step, the distribution is updated and passed as the message to its direct neighbors. After the message passes to the terminal node, an estimation of the network reliability is found. The method results in an analytical solution for system reliability. We present the derived updating rules for message passing and apply the method to two test applications: a system distribution network and general grid network. In the message passing, some approximations are made. Results from the applications show high accuracy for the proposed method compared to exact solutions where possible for comparison. In addition, PrPm achieves orders of magnitude increases in computational efficiency compared to existing approaches. This includes reducing the computational cost for analyses from an exponential increase in computation time with the size of the system to a quartic increase. The method enables the accurate and computationally tractable calculation of failure probabilities of large, generally connected systems.

Reliability analysis of complex networks is often limited by increasing dimensionality of the problem as the number of nodes and possible paths in the network increases. Analytical approaches typically result in exact calculations of network reliability. These include total enumeration, which lists all possible combinations of network components and their corresponding outcomes in the system. Other analytical methods include recursive decomposition algorithms (Dotson and Gobien 1979, Lim and Song 2012, Kim and Kang 2013) and the use of Bayesian networks to model and analyze complex networks (Tien and Der Kiureghian 2017, Tong and Tien 2017). Recent studies have advanced methods to increase the efficiency of these approaches. However, in many cases, they continue to be characterized by exponentially increasing computational requirements as the size of the system increases.

Alternatives to analytical approaches are simulation-based methods. These utilize a series of samples to compute system reliabilities (Bulteau and El Khadiri 1998, Shields et al. 2015, Zuev et al. 2015). Challenges with these methods include selecting distributions from which to sample, enumerating states such as minimum link sets or minimum cut sets to compare samples against, and calculating indicator functions to convert samples to network reliabilities.

In this paper, we propose an analytical method to calculate the reliability of complex networks. The idea originates from the concept of belief propagation for inference in network graphs. The message that is passed between nodes in the network is a joint probability distribution. At each step, the distribution is updated and passed as the message to its direct neighbors. After the message passes to the terminal node, an estimation of the network reliability is found. The method results in an approximated analytical solution for system reliability. As a probability distribution is passed through the network, we call this method the probability propagation method (PrPm). The method is applicable for both general networks (Tong and Tien 2019a) and more specifically for directed acyclic networks as is characteristic of many infrastructure systems. We call the latter the directed probability propagation method (dPrPm) (Tong and Tien 2019b). PrPm for general complex networks is described here.

The rest of this paper describes the process of PrPm. This includes selection of the propagation sequence and updating rules during message passing. In the message passing, some approximations are made. We present the derived updating rules for message passing. We then apply PrPm to two test networks: a system distribution network and general grid network to investigate its performance in terms of accuracy and computation time compared to existing methods.

1. METHODOLOGY: PROPOSED PROBABILITY PROPAGATION METHOD (PrPm)

1.1. Propagation sequence

Proper selection of the sequence of nodes to propagate the probability distribution message ensures correctness of the final reliability assessment and informs the accuracy of the result. Correctness is ensured by guaranteeing that all nodes in the network are reached in the sequence. Prioritizing propagation to nodes with one direct neighbor (compared to multiple direct neighbors) improves accuracy of the reliability assessment because the approximations that exist in the general case become exact calculations in the one direct neighbor case. The rules for selecting the propagation sequence are as follows:

First, newly defined propagated nodes must be the direct neighbors of propagated nodes. Second, newly defined propagated nodes should not separate any two non-propagated nodes. This guarantees that every node in the network is considered. Third, newly defined propagated nodes should not connect with each other. This guarantees that every link in the network is considered. For a given network, more than one propagation sequence may be identified satisfying these three rules. In that case, nodes with one direct neighbor should be prioritized for propagation as this yields no approximation in the calculation.

1.2. Updating rules for message passing

The probability distribution message is passed through the network according to the selected propagation sequence. The message is updated at each step according to the rules described in this subsection. We assume a binary network, i.e., the nodes in the network can be in one of two states such as 0 or 1 indicating failure or survival, respectively. We also assume that each node receives messages from at most two direct neighbors. The situation where a node has more than two direct neighbors is addressed through a nodal expansion procedure presented in the following subsection.

We provide updating rules for two cases: one where a node receives a message from one direct neighbor as shown in Figure 1, or from two direct neighbors as shown in Figure 2. In these figures, empty circles represent non-propagated nodes that have not yet received any message. Solid diamonds represent propagated nodes that will be involved in future propagation steps. We name these as boundary nodes. Solid circles represent propagated nodes that will not be involved in any future massage passing. We name these as non-boundary nodes. In Figure 1 and 2, we denote the node that receives a message as N, the direct neighbors that pass the message as A and B, and a general boundary node that is not a direct neighbor to N as C. Let us consider the one direct neighbor propagation case first.

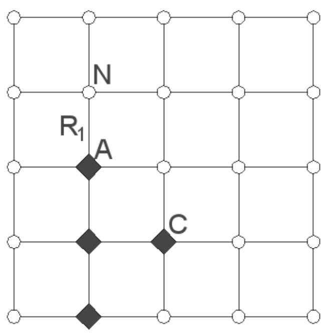


Figure 1: Message passing from one direct neighbor

In Figure 1, R_1 indicates the reliability of the link between A and N. Let the probability distribution for nodes A and C be represented by the probability values as shown in Table 1.

Table 1: Probability distribution for nodes A and C

A	C	$\Pr(\cdot)$
0	0	P_1
0	1	P_2
1	0	P_3
1	1	P_4

Let R represent the reliability of node N, i.e., $R = \Pr(N = 1)$. The updating rules to obtain the elements of the three-node joint distribution are then as shown in Table 2. Once the three-node joint distribution is calculated, we can easily define the new two-node joint distributions $p(A, N)$ and $p(C, N)$.

Table 2: Updating rules when passing message from one direct neighbor

A	C	N	Update
0	0	0	$P_1(1 - R_1 R)$
0	0	1	$P_1 R_1 R$
0	1	0	$P_2(1 - R_1 R)$
0	1	1	$P_2 R_1 R$
1	0	0	$P_3(1 - R_1 R)$
1	0	1	$P_3 R_1 R$
1	1	0	$P_4(1 - R_1 R)$
1	1	1	$P_4 R_1 R$

Figure 2 shows the case of message passing from two direct neighbors. R_1 and R_2 indicate the reliabilities of the links between A and N, and between B and N, respectively. The probability distribution values for nodes A, B, and C are shown in Table 2. The updating rules for the two direct neighbors case are shown in Table 3.

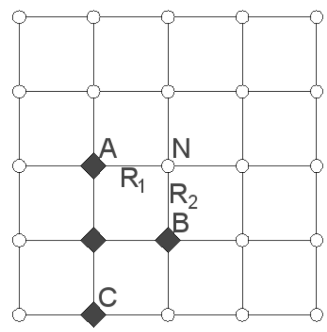


Figure 2: Message passing from two direct neighbors

Table 3: Probability distribution for nodes A, B, and C

A	B	C	$\Pr(\cdot)$
0	0	0	P_1
0	0	1	P_2
0	1	0	P_3
0	1	1	P_4
1	0	0	P_5
1	0	1	P_6
1	1	0	P_7
1	1	1	P_8

Table 4: Updating rules when passing message from two direct neighbors

A	B	C	N	Update
0	0	0	0	P_1
0	0	0	1	0
0	0	1	0	P_2
0	0	1	1	0
0	1	0	0	$P_3(1 - R_2 R)$

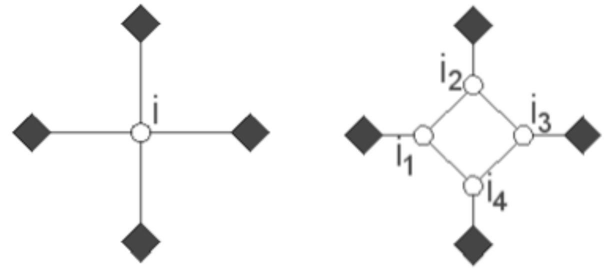
0	1	0	1	$P_3(1-R_1)R_2R$
0	1	1	0	$P_4(1-R_2R)$
0	1	1	1	$P_4(1-R_1)R_2R$
1	0	0	0	$P_5(1-R_1R)$
1	0	0	1	$P_5RR_1(1-R_2)$
1	0	1	0	$P_6(1-R_1R)$
1	0	1	1	$P_6(1-R_2)R_1R$
1	1	0	0	$P_7\{1-[1-(1-R_1)(1-R_2)]R\}$
1	1	0	1	$P_3R_1R_2R+P_5RR_1R_2+P_7[1-(1-R_1)(1-R_2)]R$
1	1	1	0	$P_8\{1-[1-(1-R_1)(1-R_2)]R\}$
1	1	1	1	$P_4R_1R_2R+P_6RR_1R_2+P_8[1-(1-R_1)(1-R_2)]R$

From the updated four-node joint distribution shown in Table 4, the new joint distributions $p(A,N)$, $p(B,N)$, and $p(C,N)$ for future propagation steps can be defined accordingly. $p(A,B)$, $p(A,C)$, and $p(B,C)$ are updated as well.

One important result from the updating rules given in Table 2 and Table 4 is that we need the joint distributions of only two nodes rather than all nodes during the message-passing process. While this yields an approximated solution, PrPm reduces the computational cost from an exponential increase with the number of nodes in the network $O(2^n)$ if a general full distribution is considered to a quartic increase $O(n^4)$. A detailed computational complexity analysis is presented in the following subsection.

1.3. Nodal expansion

In general, a node may have more than two direct neighbors. In these cases, we expand the nodes such that each node has at most two direct neighbors, and the previous updating rules apply. An example where a node i that has four direct neighbors (a) is expanded to have at most two direct neighbors (b) is shown in Figure 3.



(a) (b)

Figure 3: Nodal expansion example from four direct neighbors (a) to two direct neighbors (b)

For the nodal expansion case shown in Figure 3, instead of updating node i directly, we update the node sequentially $i_1 \rightarrow i_2 \rightarrow i_3 \rightarrow i_4$. Nodal expansion is performed before beginning the message passing. Without affecting the connectivity of the original network, the additional links created by nodal expansion are set to be 100% reliable. PrPm then proceeds with the expanded network.

The nodal expansion in combination with the updating rules governs the computational complexity of PrPm. For a network of n nodes, as we need to expand the node to ensure that every node receives information from at most two direct neighbors, the newly defined propagated nodes connect to $O(n)$ neighbors with the maximum number being n for a fully connected network. Thus, there will be $O(n^2)$ nodes in total. According to the updating rules, for each newly defined propagated node, the computational cost for that node is $O(n^2)$. This is due to the number of C nodes being $O(n^2)$ with the maximum number being n^2-2, excluding node A and node N as shown in Figure 1. Therefore, the total computational cost is the combined individual computational costs, $O(n^2)O(n^2)=O(n^4)$.

2. NETWORK TEST APPLICATIONS

To assess performance of the proposed PrPm, we apply it to two test networks. We are interested in the accuracy of the method as well as its computational cost. Both networks have exact solutions for comparison of accuracy. All results for computation times are based on computations run in MATLAB_R2016b on a 16 GB RAM computer.

The first application is a system distribution network with multiple sources for a resource to flow to a single sink. The performance of PrPm is evaluated across a range of link reliabilities for this case. The second application is a highly connected general grid network to assess the performance of the proposed method in terms of both accuracy and efficiency for systems of increasing size.

2.1. System distribution network

Figure 4 shows the system distribution network test application, previously investigated in Der Kiureghian and Song (2008) and Tien (2017). Nodes 1, 7, and 18 are sources in the network and shown as solid diamonds in the first propagation step, i.e., they are propagated nodes that will be involved in future propagation steps. Node T is the terminal sink node. The reliability at node T, or the probability of providing the system resource at node T, is of interest.

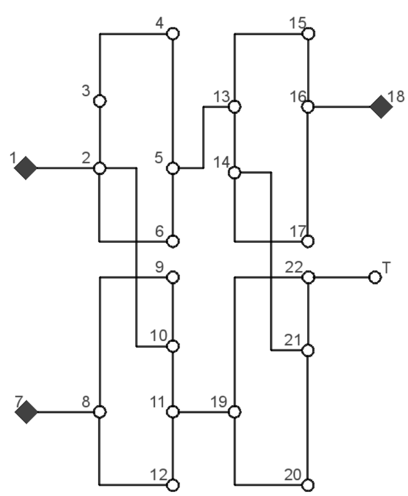


Figure 4: Configuration of system distribution network test application

For this example, all components are assumed to be independent and no nodal failure is considered. Previous studies assess network reliability based on varying nodal failure probabilities. Here, we convert to link failure probabilities. Compared with the original network, links $1-2$, $3-10$, $5-13$, $7-8$, $11-19$, $14-21$ and $16-18$ are assumed to be perfectly reliable as there are no additional elements on these links. All other links, which in the original power distribution system example have circuit breakers, switches, and transformers located on them, have a probability of failure p_f.

The results of PrPm to calculate the reliability at node T are shown in Table 5. Network reliability is indicated as "Rel" in the table. For this network, the existence of the C node as shown in Figure 2 during the message-passing process introduces errors into the propagation. Therefore, the results obtained by PrPm are an approximation in this case. In Table 5, the accuracy of PrPm is compared with the exact solution obtained by total enumeration and the sampling-based solution from Monte Carlo with 10000 realizations. The percent error, indicated "% err" in the table, is calculated relative to the network failure probability. Results are shown across a range of link reliabilities.

Table 5: Accuracy of PrPm compared with exact solution and Monte Carlo across varying link reliabilities

p_f	Exact	Monte Carlo		PrPm	
		Rel	% err	Rel	% err
0.01	0.9899	0.9912	12.8713	0.9899	0
0.05	0.9476	0.9454	4.1985	0.9474	0.3817
0.10	0.8900	0.8936	3.2727	0.8888	1.0909
0.15	0.8264	0.8211	3.0530	0.8233	1.7857
0.20	0.7551	0.7635	3.4300	0.7503	1.9600

In addition to accuracy, computational efficiency is investigated. Results for the required computation time to calculate network reliability for the exact solution, Monte Carlo, and PrPm are shown in Table 6.

Table 6: Computation time (s) for PrPm compared with exact solution and Monte Carlo

Exact	Monte Carlo		PrPm	
Time	Time	Ratio	Time	Ratio
113.63	1.1013	103.18	0.0629	1806.52

From Table 5 and Table 6, PrPm outperforms Monte Carlo simulation in both accuracy and computation time. As expected for simulation-based methods, the percentage error for Monte Carlo increases as the failure probabilities decrease. In comparison, PrPm increases in accuracy as failure probabilities decrease. In terms of computation, as PrPm provides an analytical solution, the burden of the method remains constant across system failure probabilities. Therefore, for all cases, PrPm increases the efficiency of obtaining the solution by one order of magnitude compared to Monte Carlo and more than three orders of magnitude compared to the exact solution.

2.2. General grid network

To assess performance of PrPm for general systems, a general grid network is analyzed and performance evaluated for grids of increasing size. An example 5x5 grid is shown in Figure 5. The corner-to-corner reliability from the source node (indicated as S) to the terminal node (indicated as T) is of interest. For the problem, link reliability is assumed to be 0.9 with node reliability 1.

Figure 5: Example 5x5 grid network

In PrPm, the increase in computational efficiency compared to the exact analytical solution is due to propagating only the two-node joint probability distribution during message passing compared to the full joint distribution. This results in an approximation error. To assess the effect of this approximation on the accuracy of the solution, Table 7 shows the estimated corner-to-corner network reliability obtained from PrPm compared to consideration of the full joint distribution. The obtained bounds from the full joint distribution are guaranteed to include the exact solution. Percentage error is calculated for the PrPm approximation result for network reliability compared to the median of the reliability bounds. Consideration of the full distribution is intractable for grids of sizes larger than 12x12. The computation time comparison is shown in Table 8.

Table 7: Accuracy of PrPm compared with full distribution bounds for grids of increasing size to calculate corner-to-corner reliability

Size	PrPm	Full distribution		% error
		Upper	Lower	
3×3	0.9833	0.9725	0.9724	1.1157
4×4	0.9872	0.9751	0.9750	1.2461
5×5	0.9877	0.9756	0.9755	1.2455
6×6	0.9878	0.9756	0.9756	1.2505
7×7	0.9878	0.9757	0.9757	1.2401
8×8	0.9878	0.9757	0.9757	1.2401
9×9	0.9878	0.9757	0.9757	1.2401
10×10	0.9878	0.9757	0.9757	1.2401
11×11	0.9878	0.9757	0.9757	1.2401
12×12	0.9878	0.9757	0.9757	1.2401
20×20	0.9878	/	/	/
30×30	0.9878	/	/	/
40×40	0.9878	/	/	/
50×50	0.9878	/	/	/
75×75	0.9878	/	/	/
100×100	0.9878	/	/	/

Table 8: Computation time (s) of PrPm compared with full distribution bounds for grids of increasing size

Size	PrPm	Full	Ratio
3×3	0.1003	0.16	1.60
4×4	0.1092	0.26	2.39
5×5	0.1109	0.72	6.49
6×6	0.1162	2.83	24.40
7×7	0.1162	10.88	93.79
8×8	0.1162	42.98	370.52
9×9	0.1162	174.56	1504.83
10×10	0.1162	723.55	6237.50

11×11	0.1174	2986.20	25523.08
12×12	0.1287	12688.82	98362.95
20×20	0.2188	/	/
30×30	0.6423	/	/
40×40	1.7913	/	/
50×50	4.6558	/	/
75×75	46.1766	/	/
100×100	196.0232	/	/

From Table 8, we see that when considering the full joint distribution, there is an exponential increase in computation time as the size of the grid increases. For a propagation step with n_b boundary nodes, calculating the joint distribution requires the storage and updating of 2^{n_b} elements, resulting in an exponentially increasing computational complexity with n at $O(2^{n_b})$. With the consideration of the two-node joint distribution, the time complexity of computation for the proposed PrPm is quartic at $O(n^4)$. With this, the accuracy of the result is slightly lowered by 1.24% as shown in Table 7. However, this error remains stable as grid size increases, and the computational cost is reduced by several orders of magnitude compared to the exact solution, with computational savings increasing as the size of the network increases.

3. CONCLUSIONS

The proposed PrPm provides a method to conduct efficient and accurate analysis of the reliability of complex networks. While the method results in an approximation of the network reliability, strategic selection of the propagation sequence improves its accuracy. The analysis of networks comprising nodes with at most two direct neighbors results in exact solutions of reliability. Ongoing work is investigating the applicability of the method to networks of different types.

4. REFERENCES

Bulteau S., and El Khadiri M. (1998). "A Monte Carlo simulation of the flow network reliability using importance and stratified sampling," *PhD diss., INRIA*, 32, 271-287.

Der Kiureghian A., and Song J. (2008). "Multi-scale reliability analysis and updating of complex systems by use of linear programming," *Reliability Engineering and System Safety*, 93, 288-297.

Dotson, W., and Gobien, J.O. (1979). "A New Analysis Technique for Probabilistic Graphs," *IEEE Transactions on Circuits and Systems*, 26(10), 855–865.

Kim, Y., and Kang, W. (2013). "Network reliability analysis of complex systems using a non-simulation-based method," *Reliability Engineering and System Safety*, 110, 80-88.

Lim, H.-W., and 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.

Shields, M. D., Teferra, K., Hapij, A., and Daddazio, R.P. (2015). "Refined stratified sampling for efficient Monte Carlo based uncertainty quantification." *Reliability Engineering and System Safety*, 142, 310-325.

Tien, I. (2017). "Bayesian network methods for modeling and reliability assessment of infrastructure systems," In *Risk and Reliability Analysis: Theory and Applications*, Springer International Publishing, 417-452.

Tien, I., and Der Kiureghian, A. (2017). "Reliability assessment of critical infrastructure using Bayesian networks," *Journal of Infrastructure Systems*, 23(4).

Tong, Y., and Tien, I. (2017). "Algorithms for Bayesian network modeling, inference, and reliability assessment for multi-state flow networks," *Journal of Computing in Civil Engineering*, 31(5).

Tong, Y., and Tien, I. (2019a) "Analytical Probability Propagation Method for Reliability Analysis of General Complex Networks," *Reliability Engineering and System Safety*,

Tong, Y., and Tien, I. (2019b). "Probability Propagation Method for Reliability Assessment of Acyclic Directed Networks," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems*.

Zuev, K., Wu, S., and Beck, J. (2015). "General network reliability problem and its efficient solution by Subset Simulation", *Probabilistic Engineering Mechanics,* 40, 25-35.