Reliability Engineering and System Safety 156 (2016) 134–147

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Algorithms for Bayesian network modeling and reliability assessment of infrastructure systems

Iris Tien [a,c,*], Armen Der Kiureghian [b,c]

[a] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355, USA
[b] Civil Engineering Emeritus, Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720-1710, USA
[c] American University of Armenia, Yerevan, Armenia

ARTICLE INFO

Article history:
Received 9 February 2015
Received in revised form
19 July 2016
Accepted 26 July 2016
Available online 2 August 2016

Keywords:
Bayesian networks
Systems modeling
Algorithms
Reliability assessment
Risk analysis
Infrastructure systems

ABSTRACT

Novel algorithms are developed to enable the modeling of large, complex infrastructure systems as Bayesian networks (BNs). These include a compression algorithm that significantly reduces the memory storage required to construct the BN model, and an updating algorithm that performs inference on compressed matrices. These algorithms address one of the major obstacles to widespread use of BNs for system reliability assessment, namely the exponentially increasing amount of information that needs to be stored as the number of components in the system increases. The proposed compression and inference algorithms are described and applied to example systems to investigate their performance compared to that of existing algorithms. Orders of magnitude savings in memory storage requirement are demonstrated using the new algorithms, enabling BN modeling and reliability analysis of larger infrastructure systems.

1. Introduction

Infrastructure systems are essential for a functioning society, from distributing the water we drink, to delivering the electricity we consume, to enabling transport of people and goods from source to destination points. Our nation's infrastructure, however, is aging and becoming increasingly unreliable with potentially severe consequences. Given a complex infrastructure network comprised of many interconnected components, system reliability analysis is required to identify the critical components and make decisions regarding inspection, repair, and replacement to minimize the risk of system failure.

The Bayesian network (BN) is a useful probabilistic tool for system reliability assessment. It is a graphical tool that offers a transparent modeling scheme, allowing easy checking of the model even by non-experts in systems analysis and probabilistic methods. In an environment where information about a system is evolving in time and is subject to uncertainty, BNs are able to update the reliability state of the system as new information, e.g., from observations, inspections, or repair actions, becomes available. Infrastructure systems are subject to high degrees of uncertainty, including discrepancies between initial design and construction, uncertain degradation of system components over time, and exposure to stochastic hazards. BNs provide the proper probabilistic framework to handle such information for engineering decision making.

A major obstacle to widespread use of BNs for system reliability assessment, however, is the limited size of the system that can be tractably modeled as a BN. This is due to the exponentially increasing amount of information that needs to be stored as the number of components in the system increases. This paper proposes a method to address this limitation.

The main contributions of this paper are novel compression and inference algorithms that enable the modeling of larger systems as BNs than has been previously possible. The paper is organized as follows: Section 2 provides a brief background on BNs, including the advantages of using BNs for system reliability analysis and the current limitations in BN modeling of large systems. Section 3 introduces the proposed compression algorithm for constructing and storing the conditional probability tables (CPTs) required by the BN. Section 4 describes the inference algorithm for system reliability analysis, which uses the compressed CPTs without decompressing them. Section 5 demonstrates the proposed algorithms through application to a test system. Results for memory storage and computation time are presented.

* Corresponding author.
E-mail address: itien@ce.gatech.edu (I. Tien).

http://dx.doi.org/10.1016/j.ress.2016.07.022

2. Background

2.1. Methods for system reliability assessment, including Bayesian networks (BNs)

Over the years, many methods have been developed to assess system reliability. While not intended to be an exhaustive list, these include reliability block diagrams (RBDs), fault trees (FTs), event trees (ETs), binary decision diagrams (BDDs), and Bayesian networks (BNs). RBDs, FTs, and ETs are symbolic models showing the logical relationships between component states and system outcomes. These can be extended into BDDs, graphical models representing Boolean relationships between variables. General relationships between random variables can be modeled graphically using BNs.

RBDs are useful to show the components of a system and their relationships; however, they are not efficient for reliability analysis of complex infrastructure systems [17,20]. FTs have been used extensively in the nuclear industry [33]. They are constructed for a particular undesired system outcome; hence, a single FT cannot model all possible modes or causes of system failure [4]. ETs trace forward through a causal chain to assess the probability of occurrence of different system outcomes. The size of an event tree, however, can grow exponentially with the number of sequential events [22]. BDDs are useful for modeling Boolean functions, as they occur in system reliability analysis [1,7]. The number of nodes and paths in a BDD is exponential with the number of variables in the domain of the Boolean function [19]. For reliability analysis of an infrastructure system, this implies a BDD of exponentially increasing size as the number of components in the system increases.

A Bayesian network (BN) is a graphical model comprised of nodes and links. Each node represents a random variable and each link describes the probabilistic dependency between two variables. Each BN node is assigned a set of mutually exclusive and collectively exhaustive states. In our application, the nodes represent the states of the system components and the overall system performance, and the links describe the probabilistic dependencies between component and system performance. The reader is referred to texts such as [13] for further information on BNs.

As stated earlier, the capability of BNs for updating and handling of uncertain information and their graphical modeling representation makes them particularly well suited for reliability assessment of infrastructure systems under evolving states of information [30]. There are both exact and approximate methods for inference in BNs. These methods are applicable given a BN structure, to update probability assessments over the network in light of new information. For the case where system topology changes, as can occur in post-disaster scenarios, first a restructuring of the BN, then performing inference over the new BN is necessary. Approximate inference methods are generally sampling based, including importance sampling [23,35] and Markov chain Monte Carlo [11]. In theory, these methods converge to the exact solution for a sufficiently large number of samples. In practice, however, the rate of convergence is unknown and can be slow [27]. This is especially true when simulating events that are a priori unlikely. Exact inference methods are, therefore, preferred. The algorithm described in Section 4 is for exact inference.

2.2. Current limitations in BN modeling of large systems

The use of BNs for system reliability assessment has been limited by the size and complexity of the system that can be tractably modeled. Systems analyzed in previous studies have been small, typically comprised of 5–10 components. This includes studies on generating BNs from conventional system modeling methods, e.g., RBDs [14,32] and FTs [5]. Mahadevan et al. [16] demonstrate the ability of the BN to use system-level test data to update information at the component level. They note that the computational effort increases significantly with the number of system components. They introduce an approach characterized as "branch and bound," whereby events of relatively low probability are ignored, to apply the BN to larger systems. The example given, however, is for a system consisting of only 8 components, and the willful discarding of available information, leading to a subsequent loss of accuracy in the result, is not ideal.

Boudali and Dugan [6] use BNs to model the reliability of slightly larger systems, including a system of 16 components. However, the authors state that this "large number" of components makes it "practically impossible" to solve the network without resorting to simplifying assumptions or approximations. Clearly, even a system of 16 components is not enough to create a full model of many real-world infrastructures. Nielsen et al. [19] propose a method utilizing Reduced Ordered Binary Decision Diagrams (ROBDDs) to efficiently perform inference in BNs representing large systems with binary components. However, a troubleshooting model is considered, which includes a major assumption of single-fault failures, i.e., the malfunction of exactly one component causes the system to be at fault. In general, the number of paths in the ROBDD is exponentially increasing with the number of components. It is the single-fault assumption that bounds the size of the ROBDD. For general systems, including infrastructure systems, this single-fault assumption cannot be guaranteed. Therefore, the gains from using the ROBDD may not be applicable.

Finally, a topology optimization algorithm is proposed in Bensi et al. [3] to address the inefficiency of a converging BN structure as shown in Fig. 1. The authors develop a discrete optimization program to create a more efficient, chain-like BN model of the system based on survival- or failure-path sequences. The proposed optimization program, however, must consider the permutation of all component indices and, therefore, may become intractably large for large systems.

2.3. Conditional probability tables in construction of BN

The system size limitation arises due to the conditional probability tables (CPTs) that must be associated with each node in the BN. In the BN terminology, the CPT of a child node provides the probability mass function of the variable represented by that node given each of the mutually exclusive combinations of the states of the parent nodes. For an infrastructure system, the state of the system is dependent on the states of each of its constituent components, as shown in Fig. 1. The BN can include parent nodes of the components, as indicated by the dashed arrows, representing common hazards, characteristics, or demands among components. The focus of this study is on the system description

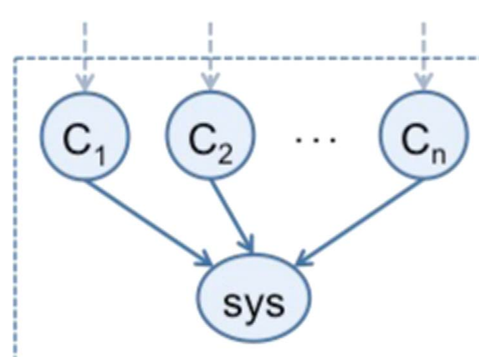

Fig. 1. BN of a system comprised of n components.