Reliability Engineering and System Safety 94 (2009) 542–550

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

A generic method for estimating system reliability using Bayesian networks

Ozge Doguc, Jose Emmanuel Ramirez-Marquez *

Stevens Institute of Technology, Hoboken, NJ 07030, USA

ARTICLE INFO

Article history:
Received 6 August 2007
Received in revised form
26 March 2008
Accepted 14 June 2008
Available online 19 June 2008

Keywords:
System reliability
Reliability methods
Reliability analysis
Reliability modeling

ABSTRACT

This study presents a holistic method for constructing a Bayesian network (BN) model for estimating system reliability. BN is a probabilistic approach that is used to model and predict the behavior of a system based on observed stochastic events. The BN model is a directed acyclic graph (DAG) where the nodes represent system components and arcs represent relationships among them. Although recent studies on using BN for estimating system reliability have been proposed, they are based on the assumption that a pre-built BN has been designed to represent the system. In these studies, the task of building the BN is typically left to a group of specialists who are BN and domain experts. The BN experts should learn about the domain before building the BN, which is generally very time consuming and may lead to incorrect deductions. As there are no existing studies to eliminate the need for a human expert in the process of system reliability estimation, this paper introduces a method that uses historical data about the system to be modeled as a BN and provides efficient techniques for automated construction of the BN model, and hence estimation of the system reliability. In this respect *K*2, a data mining algorithm, is used for finding associations between system components, and thus building the BN model. This algorithm uses a heuristic to provide efficient and accurate results while searching for associations. Moreover, no human intervention is necessary during the process of BN construction and reliability estimation. The paper provides a step-by-step illustration of the method and evaluation of the approach with literature case examples.

1. Introduction

System reliability can be defined as the probability that a system will perform its intended function for a specified period of time under stated conditions [1]. Estimating system reliability is an important and challenging problem for system engineers. It is important because a company's reputation, customer satisfaction and system design costs can be directly related to the failures experienced by the system [2]. It is also challenging since current estimation techniques require a high level of background in system reliability analysis, and thus familiarity with the system.

Traditionally, engineers estimate reliability by understanding how the different components in a system interact to guarantee system success. Typically, based on this understanding, a graphical model (usually in the form of a fault tree, a reliability block diagram or a network graph) is proposed to represent how component interaction affects system functioning. Once the graphical model is obtained, different analysis methods [3–5] (minimal cut sets, minimal path sets, Boolean truth tables, etc.) can be used to quantitatively represent system reliability. Finally, the reliability characteristics of the components in the system are introduced into the mathematical representation in order to obtain a system-level reliability estimate.

This traditional perspective aims to provide accurate predictions about the system reliability using historical or test data. This approach is valid whenever the system success or failure behavior is well understood. However, for new complex systems, both in the design phase or already deployed, understanding component interaction may prove to be a challenging problem, which usually requires intervention of a domain expert.

To address these challenges, Bayesian networks (BNs) have been proposed as an alternative to traditional reliability estimation approaches [1,6,7]. In this respect, Barlow [8] and Almond [9] introduced some of the earliest studies for reliability estimation via BN. Barlow [8] provided a case study on spherical pressure vessels by comparing Bayesian and non-Bayesian approaches to estimate system reliability, while Almond [9] introduced the graphical-belief environment, which is used for risk prediction in large complex systems. Based on these studies, it became clear that for some systems, BN had significant advantages—in terms of efficiency in evaluating associations and simplicity in providing a system model—over traditional reliability evaluation frameworks,

Abbreviations: BN, Bayesian network; CPT, conditional probability table; *K*2, named after Kutató 2; HRP, Halden project

* Corresponding author.
E-mail addresses: jmarquez@stevens.edu, Jose.Ramirez-Marquez@stevens.edu (J.E. Ramirez-Marquez).

0951-8320/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2008.06.009

Nomenclature

$p(A)$	probability of an event A
$p(H\|E)$	probability of event H given the evidence E
X_i	node in BN representing component i in the system
u	upper bound on the number of parents of a node in a BN
n	number of components in a system
F	scoring function used in the $K2$ algorithm
π_i	set of candidate parents for X_i
j	a parent in π_i
k	an instantiation of parent X_j
α_{ijk}	number of observations on component X_i, where both component X_i and parent X_j are instantiated as k
q_i	total number of unique parent set instantiations for node X_i
d_i	number of instantiations that node X_i can have
α_{ij}	sum of α_{ijk}'s for component X_i and its parent X_j
$p(X_i=1)$	probability of node $X_i=1$
t	number of observations on a system
A_I	number of incorrect associations decided by the $K2$ algorithm in a BN
A_{FP}	number of incorrect associations decided by the $K2$ algorithm in a BN
A_{FN}	number of associations in a BN that cannot be discovered by the $K2$ algorithm
A_T	total number of associations in a BN
ρ	error rate of the $K2$ algorithm

due to their ease of use in interaction with domain experts in the reliability field [10].

From a reliability perspective, a BN can be represented as a directed graph with (1) nodes that represent system components and (2) edges that show relationships among the nodes. To evaluate the BN, via Bayes' rule, edges within the graph are assigned with a value that shows the degree (or strength) of the relationship it represents—a detailed description of BN will be provided in Section 3. Currently, approaches for system reliability analysis via a BN [6,10–13] use specialized networks designed for a specific system. That is, the BN structure (i.e. the nodes, their associations and the strength of these associations) used for estimating system reliability must be known *a priori*. Inherently, this assumption supposes that an expert with "adequate" knowledge about the system behavior can build the BN. However, finding such an expert may not be possible at all times for every system under consideration. The number of such experts is limited, and finding one is usually difficult and costly [14]. Moreover, human intervention is always open to unintentional mistakes, which could cause discrepancy in the results.

These issues are particularly true in complex systems, where the number of components and interactions is large, and thus the likelihood of miscalculations can be substantial. As a result, such an assumption may not be realistic enough to be applicable. Additionally, modern systems evolve with time by adding and removing new and obsolete components [15]. Thus, the original BN model may not be accurate throughout a system life cycle, forcing human intervention to be needed right after every change in the system. Therefore, in the applications where an expert builds the BN, there is always the need for keeping that same expert as the system evolves.

To address these issues, this study introduces a holistic method for estimating system reliability. The main contribution of this paper is to relieve the need of an expert by linking BN construction from raw component and system data. In essence, the method replaces the need of an expert to find associations among the components with raw data related to component and system behavior. These data are then used to develop (via association rule mining) and evaluate (via conditional probabilities and Baye's theorem) a BN that describes the relationships and interactions of components with system success behavior. Based on the extensive literature review, this is the first study that incorporates these methods for estimating system reliability by eliminating the need for human intervention in BN construction.

The proposed method automates the process of BN construction by feeding raw system behavior data to the $K2$ algorithm (a commonly used association rule mining algorithm). This algorithm is a machine-learning algorithm that uses canonically ordered sets of variables and identifies the associations among them by using a predefined scoring function and a heuristic. The $K2$ algorithm has proven to be efficient and accurate for finding associations [16] from a dataset of historical data about the system. The $K2$ algorithm reduces the algorithmic complexity of finding associations from exponential (2^n) to quadratic (n^2) [16] with respect to the number of components (n) in the system. Therefore, the method proposed in this paper is efficient for complex systems with large datasets. Moreover, unlike previous approaches, the proposed solution is not system specific, it can be applied to systems following any kind of configuration (two terminal, k-terminal, all terminal, etc.) and behavior (binary, capacitated and multi-state) without the need for an expert to assess and identify component relationships. That is, the approach can build a BN and estimate reliability for any system when observed system failure data are available.

This paper is organized as follows: Section 2 gives a brief summary about related work in the literature. Section 3 provides information about BNs and the $K2$ algorithm, while Section 4 presents a comprehensive illustration of BN construction and system reliability estimation techniques. In Section 5 experimental analysis on the accuracy and performance of the methodology is provided. Finally, in Section 6 conclusions and future research directives are provided.

2. Literature survey

Estimation of systems reliability using BN dates back as early as 1988, when it was first defined in [8]. The idea of using BN in systems reliability has mainly gained acceptance because of the simplicity it allows to represent systems and the efficiency for obtaining component associations. The concept of BN has been discussed in several earlier studies [17–19]. More recently, BN have found applications in software reliability [20,21], fault finding systems [18] and general reliability modeling [11].

Currently, predefined BN are used for reliability estimation for specific systems. Gran and Helminen [1] provide a BN for nuclear power plants and introduce a hybrid method for estimating the reliability of the plant. The nuclear plant is composed of a software system and the plant hardware; therefore, they combined two BNs that were currently being used for corresponding systems: (1) the Halden project (HRP) [22] uses a BN for risk assessment based on disparate evidences. (2) The VTT automation [23] focuses on the reliability of software-based systems using BN. Moreover, there is another challenge: each BN uses a different modeling and simulation environment. The HRP uses HUGIN [24] and SERENE [25], which uses conditional probability table (CPT)