Reliability Engineering and System Safety 95 (2010) 412–425

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Improved reliability modeling using Bayesian networks and dynamic discretization

David Marquez [a], Martin Neil [a,b,*], Norman Fenton [a,b]

[a] Department of Computer Science, Queen Mary, University of London, UK
[b] Agena Ltd., London, UK

ARTICLE INFO

Article history:
Received 28 November 2008
Received in revised form
19 November 2009
Accepted 22 November 2009
Available online 27 November 2009

Keywords:
Bayesian networks
Systems reliability
Dynamic fault trees
Dynamic discretization

ABSTRACT

This paper shows how recent Bayesian network (BN) algorithms can be used to model time to failure distributions and perform reliability analysis of complex systems in a simple unified way. The algorithms work for so-called hybrid BNs, which are BNs that can contain a mixture of both discrete and continuous variables. Our BN approach extends fault trees by defining the time-to-failure of the fault tree constructs as deterministic functions of the corresponding input components' time-to-failure. This helps solve any configuration of static and dynamic gates with general time-to-failure distributions. Unlike other approaches (which tend to be restricted to using exponential failure distributions) our approach can use any parametric or empirical distribution for the time-to-failure of the system components. We demonstrate that the approach produces results equivalent to the state of the practice and art for small examples; more importantly our approach produces solutions hitherto unobtainable for more complex examples, involving non-standard assumptions.. The approach offers a powerful framework for analysts and decision makers to successfully perform robust reliability assessment. Sensitivity, uncertainty, diagnosis analysis, common cause failures and warranty analysis can also be easily performed within this framework.

1. Introduction

Most published reliability analysis methods are based on parametric and non-parametric statistical models of time-to-failure data and their associated metrics [1]. The underlying assumption of these methods is that a coherent, statistical model of system failure time can be developed that will prove stable enough to accurately predict a system's behaviour over its lifetime. However, given the increasing complexity of the component dependencies and failure behaviours of today's real-time safety and mission critical systems, the statistical models may not be feasible to build or computationally tractable. This has led to an increasing interest in more flexible modeling frameworks for reliability analysis. The most notable such frameworks are combinatorial models such as fault trees (FTs), Markov chain based approaches such as dynamic fault trees (DFTs), which are described in Section 2.1 and Bayesian networks (BNs), which are described in Sections 2.2 and 2.3.

While the DFT approach [2,3] is very flexible, in practice it has severe limitations, such as the problem of state space explosion and the inability to handle non-standard statistical distributions. The Bayesian network (BN) framework has provided a compact representation of the event-dependent failure behaviours, characteristic of fault-tolerant systems, avoiding the state space explosion problem of the Markov chain based approaches, [4,5]. However, for real world applications, the BN models necessary for reliability assessment are invariantly 'hybrid' models [6–8], meaning that they contain both discrete variables (e.g., the possible system states) and continuous variables (e.g., operating and environmental covariates influencing the reliability of a system). Due to limitations in inference algorithms, previous attempts to apply reliability assessment to such models have been unreasonably constrained, such as for example having to assume Gaussian or exponential distributions for all continuous variables that are not adequately handled .

This paper presents (in Section 3) a simple event-based hybrid BN modeling method for reliability assessment that scales up to large, complex dynamic systems, and overcomes the limitations of both dynamic fault trees and previous BN approaches. The basic idea of our BN approach is to define the time-to-failure of the fault tree constructs as deterministic functions of the corresponding input components' time-to-failure. The new approach incorporates a recent powerful approximate inference algorithm for hybrid BNs, based on a process of dynamic discretization of the domain of all continuous variables in the BN [6]. By efficiently

* Corresponding author at: Department of Computer Science, Queen Mary, University of London, UK.
E-mail address: martin@dcs.qmul.ac.uk (M. Neil).

0951-8320/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2009.11.012

integrating the iterative approximation scheme within existing robust propagation algorithms on BN architectures, such as junction tree (JT) [9], robust inference on complex hybrid models (without any constraints on the distributions of continuous variables) can be performed.

The power and flexibility of the approach is demonstrated (in Sections 3 and 4) by comparing the results with traditional state space approaches like DFTs, used in a number of popular reliability tools. The accuracy of the algorithm is tested on a range of classical dynamic fault trees constructs, allowing the system components to adopt any time to failure distribution occurring in practical applications. In each case the results are compared with the analytical solution of the Markov chain representation or the approximated solutions generated by numerical integration schemes, as appropriate. The results are very close to the analytical solutions and are achieved with much less effort. In several cases the approach provides predictions of situations that simply cannot be modelled by alternative approaches. This shows that the approach can not only replicate the results provided by existing algorithms but also extends the state of art by providing solutions to problems that the past generation of algorithms cannot solve.

All the example models shown in this paper are built and executed using the commercial general-purpose Bayesian network software tool AgenaRisk [10], in which the dynamic discretization algorithm is now implemented.

2. Background

This section reviews the most relevant previous work. Section 2.1 reviews the work on reliability modeling that has evolved from FT and DFT type analysis. This includes Monte Carlo methods. The rest of the section covers BNs. A brief overview of BNs is presented in Section 2.2 and a review of previous BN work in reliability modeling is presented in Section 2.3.

2.1. Reliability modeling using fault tree and Markov chain based approaches

The most popular methods for addressing the kind of complex reliability analysis problem described in the introduction are based on two main frameworks: combinatorial models, like static FTs, and state-space models, like DFTs. In the static FT framework Boolean constructs are used to model how combinations of components' failures can cause failure of subsystems or of the whole system [11,12].

The problem with static FTs is that they cannot capture complex event-dependent behaviours (sequence-dependent failures, functional dependencies and standby spares) of fault-tolerant systems. This is the problem that the state space models, like DFTs [2,13], were developed primarily to address. DFTs have increased the modeling power of FTs by taking into account not only the combinations but also the sequential ordering of occurrence of component failures that led to system failure. Analytical solutions of DFTs are obtained by automatic conversion to the equivalent continuous time Markov process, with state-space given by the combination of occurrence of all possible events, and transition probabilities characterised by the components' hazard rates [2].

While the DFT approach can model nearly any sequence-dependent system, representing dynamic tree model failures as states of a Markov process is a daunting, error-prone task that also has two major limitations. Firstly the state-space generated grows exponentially with size of the system and, secondly, there are limitations on the modeling of spares (such as warm or cold spares with non-exponential time-to-failure distributions).

Several methods have been developed to deal with these limitations. For the first limitation, modularisation algorithms have been introduced to help to break down the size of large systems into smaller independent subtrees that do not share basic events. These subtrees are then solved separately using a suitable technique according to its classification as static (Boolean) or dynamic [13,14]. However, if the top-level gate of the fault tree is dynamic, the modularisation technique cannot be applied since it does not provide an exact solution [13]. Solving several Markov processes corresponding to the independent modules is computationally more efficient than solving a single Markov model for the entire system fault tree. However, the state space explosion and resulting computational complexity remains a major limitation in using the Markov representation of DFTs, as even a relatively simple DFT can give rise to a very large Markov model. An alternative approach has been proposed using numerical integration methods to obtain approximate solutions of DFTs without converting them to a Markov model [15].

In order to overcome the second limitation, Monte Carlo simulation techniques have been adopted to solve DFTs [13,3]. These approaches have extended the modeling capabilities of DFTs by allowing the inclusion of spares components with general time-to-failure distributions (lognormal, Weibull), previously not feasible using Markov-based techniques.

2.2. Bayesian networks overview

A BN [16,17] consists of two main elements.

- *Qualitative*: This is given by a directed acyclic graph (DAG), with nodes representing random variables, which can be discrete or continuous, and may or may not be observable, and directed arcs (from parent to child) representing causal or influential relationships between variables.
- *Quantitative*: Conditional probability distributions (CPDs), $P(X|pa(X))$, which define the probabilistic relationship of each node X given its respective parents $pa(X)$. Nodes without parents, called root nodes,[1] are described according to their marginal probability distributions $P(X)$.

Together, the qualitative and quantitative parts of the BN encode all relevant information contained in a full joint probability model. The conditional independence assertions about the variables, represented by the absence of arcs, allow decomposition of the underlying joint probability distribution as a product of CPDs. This significantly reduces the complexity of inference tasks on the BN [18,19]. If the variables are discrete, the CPDs can be represented as node probability tables (NPTs), $P(X|pa(X))$, which list the probability that the child node X takes on each of its different values for each combination of values of its parents $pa(X)$. For continuous variables, the CPDs represent conditional probability density functions (PDFs).

BNs have been applied to a range of real-world dependability-type problems [20–26,36] and these studies have demonstrated both the feasibility and usefulness of the technology. In the area of software system reliability, we have also shown the advantages of BNs over traditional methods for predictive and diagnostic reasoning [27–29].

[1] Note that, in the BN terminology, the root nodes correspond to the fault tree basic events. These should not be confused with the root of the fault tree, which refers in fault tree Analysis to the undesired Top Event, which is represented in the BN model as a non-root node (i.e., a node with parents).