Engineering Applications of Artificial Intelligence 25 (2012) 671–682

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas

P. Weber*, G. Medina-Oliva, C. Simon, B. Iung

CRAN-Nancy-Université-CNRS, UMR7039, Boulevard des Aiguillettes, B.P. 70239, F-54506 Vandœuvre lès Nancy, France

ARTICLE INFO

Article history:
Received 21 December 2009
Received in revised form
13 May 2010
Accepted 1 June 2010
Available online 16 July 2010

Keywords:
Bayesian networks
Dependability
Risk analysis
Maintenance
Reliability
Safety

ABSTRACT

In this paper, a bibliographical review over the last decade is presented on the application of Bayesian networks to dependability, risk analysis and maintenance. It is shown an increasing trend of the literature related to these domains. This trend is due to the benefits that Bayesian networks provide in contrast with other classical methods of dependability analysis such as Markov Chains, Fault Trees and Petri Nets. Some of these benefits are the capability to model complex systems, to make predictions as well as diagnostics, to compute exactly the occurrence probability of an event, to update the calculations according to evidences, to represent multi-modal variables and to help modeling user-friendly by a graphical and compact approach. This review is based on an extraction of 200 specific references in dependability, risk analysis and maintenance applications among a database with 7000 Bayesian network references. The most representatives are presented, then discussed and some perspectives of work are provided.

1. Introduction and problem statement

The management of complex industrial systems contributes to higher competitiveness and higher performances at lower costs. In that way, the relevance of the maintenance and dependability analyses increased due to their role in improving availability, performance efficiency, products quality, on-time delivery, environment and safety requirements, and total plant cost effectiveness at high levels (Alsyouf, 2007; Kutucuoglu et al., 2001). Nowadays, one of the major problems in the dependability field is addressing the system modeling in relation to the increasing of its complexity. This modeling task underlines issues concerning the quantification of the model parameters and the representation, propagation and quantification of the uncertainty in the system behavior (Zio, 2009).

In previous years, the reliability and risk analysis of systems were studied by making assumptions simplifying the study. One of these assumptions is to focus the study only on the technical part of the system. This assumption is no longer valid, since it has been shown the importance of organizational and human factors contributions (Leveson et al., 2009). Indeed, if studies were centered on technical aspects of systems until seventies (Villemeur, 1992), several major accidents, such as the Three Miles Island nuclear accident and the Bhopal catastrophe have pointed out cause operator errors and organizational malfunctions. These accidents allowed the scientific community to present and develop, in eighties, first methods centered on the analysis of these human errors. It led to the expansion of the human reliability analysis (HRA). But other accidents (Challenger explosion, Chernobyl nuclear accident …) have emphasized, in nineties, the importance of organizational malfunctions in their occurrences and, have contributed to the emergence of different theories for the study of these organizational issues: normal accident (Perrow, 1990; Weick, 2001) and high reliability organizations (Robert, 1990; Léger et al., 2008, 2009).

As a consequence, innovative studies aim at covering the whole of these causes (technical, human and organizational). Nevertheless, such analyses are often difficult to achieve because they require a lot of resources. This matter adds complexity to the systems' modeling due to the interaction between different technical, human, organizational and nowadays environmental factors which are necessary to quantify failure scenarios and risky situations. Thus, the challenge is to formalize a model of a complex system integrating all these aspects (Trucco et al., 2008; Kim et al., 2006) (Fig. 1).

Furthermore, while modeling these factors, it is required to take into account the knowledge integration of diverse natures such as qualitative and quantitative with several abstraction levels. The organization and human analyses are more naturally modeled with a qualitative knowledge (to describe situations,

* Corresponding author.
E-mail addresses: philippe.weber@cran.uhp-nancy.fr (P. Weber), gabriela.medina-oliva@cran.uhp-nancy.fr (G. Medina-Oliva), christophe.simon@cran.uhp-nancy.fr (C. Simon), benoit.iung@cran.uhp-nancy.fr (B. Iung).

0952-1976/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2010.06.002

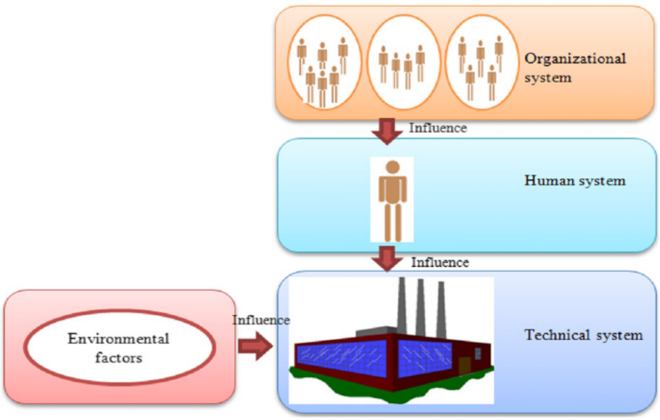

Fig. 1. Context of the complex system to be modeled.

scenarios…) such as knowledge represented in failure mode, effects, and criticality analysis (FMECA), HAZard OPerability (HAZOP), probabilistic risk assessment (PRA) analysis, etc.; and in other hand, the technical level is usually known with quantitative information (failure rates, unavailability level, Mean Time To Failure (MTTF), etc.) (Røed et al., 2008).

A complementary point of view to be modeled for the system is the temporal dimension (system dynamics) which consists in describing phenomenon such as: sequences in scenarios, degradations of components, evolution of symptoms corresponding to deterioration mechanisms, impact of preventive maintenance actions on the degradation, influence of environmental conditions and effects of the operation conditions on the evolution of the component states.

Once assessed the failure probability and risk associated to a system situation, the information is provided to support the decision making process. It implies to quantify the uncertainty and imprecision on parameters, for example, the uncertainty of the failure occurrence and its consequences (Zio, 2009).

Therefore, the main characteristics to be modeled in a system for assessing dependability and maintenance aspects are:

- the complexity and size of the system (large-scale systems) (Zio, 2009),
- the temporal aspects (Labeau et al., 2000),
- the integration of qualitative information with quantitative knowledge on different abstraction levels (Papazoglou et al., 2003; Delmotte, 2003),
- the nature of multi-state components (Griffith, 1980),
- the dependences between events such as failures (Torres-Toledano and Sucar, 1998),
- uncertainties on the parameter estimation (Zio, 2009).

For modeling these requirements, there are some classical dependability methods such as fault trees, Markov chains, dynamic fault trees, Petri nets and Bayesian networks (BN). In the recent literature, it is observed a growing interest focused on BN. This modeling method is not the solution to all problems, but it seems to be very relevant in the context of complex systems (Langseth, 2008).

Indeed some papers such as Mahadevan et al. (2001), Boudali and Dugan (2005b), Langseth and Portinale (2007) and Langseth (2008) show the increasing interest on the use of BN to estimate and to improve reliability and safety of systems over the last decade. For example, during the period 1999–2009, RESS journal (Reliability Engineering and System Safety), well known in dependability area, shows an increment of 100% of a ratio consisting on the paper number dedicated to the application of BN to reliability (or risk) divided by the total amount of papers. This type of ratio has strengthened our interest to analyze the evolution of the literature about BN and their applications on dependability, risk analysis and maintenance. For this purpose, we have built a database of references from 1990 to 2008 with different bibliographical research tools (i.e. google scholar, Sciencedirect, Web of Knowledge …). In this paper, the most relevant articles according to their citation number were referenced until 2008. Nonetheless, some citations on "hot topics" of research until 2009 are also given.

The rest of this paper is organized as follow. Section 2 is introducing the bases of BN and explaining why they are suitable to model complex systems. Section 3 shows a bibliographical review of the relevant research directions for modeling dependability, risk analysis and maintenance problems with BN. Section 4 presents a comparison of the BN modeling capabilities with other modeling methods such as Fault Tree, Markov Chains and Petri Nets. Finally, the conclusions are given by integrating also highlights future research directions.

2. BN in general

BN appear to be a solution to model complex systems because they perform the factorization of variables joint distribution based on the conditional dependencies. The main objective of BN is to compute the distribution probabilities in a set of variables according to the observation of some variables and the prior knowledge of the others. The principles of this modeling tool are explained in Jensen (1996) and Pearl (1988).

2.1. Recall of BN characteristics

A BN is a directed acyclic graph (DAG) in which the nodes represent the system variables and the arcs symbolize the dependencies or the cause–effect relationships among the variables. A BN is defined by a set of nodes and a set of directed arcs. A probability is associated to each state of the node. This probability is defined, *a priori* for a root node and computed by inference for the others.

The computation is based on the probabilities of the parent's states and the conditional probability table (CPT). For instance, let us consider two nodes A and B; with two states (S_{*1} and S_{*2}) each; structuring the BN (Fig. 2). The *a priori* probabilities of node A are defined as in Table 1.

A CPT is associated to node B. This CPT defines the conditional probabilities $P(B|A)$ attached to node B with a parent A, to define the probability distributions over the states of B given the states of A.

This CPT is defined by the probability of each state of B given the state of A (Table 2).

Thus, the BN inference computes the marginal distribution $P(B=S_{B1})$:

$$P(B=S_{B1})=P(B=S_{B1}|A=S_{A1}).P(A=S_{A1})+P(B=S_{B1}|A=S_{A2}).P(A=S_{A2}) \quad (1)$$

The added value of a BN is linked to the computation of the probabilities attached to a node state, given the state of one or

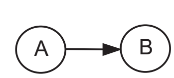

Fig. 2. Basic example of a BN.