Available online at www.sciencedirect.com
SCIENCE DIRECT®

Reliability Engineering and System Safety 91 (2006) 849–856

RELIABILITY ENGINEERING & SYSTEM SAFETY

www.elsevier.com/locate/ress

Designing a Bayesian network for preventive maintenance from expert opinions in a rapid and reliable way

G. Celeux [a], F. Corset [b,*], A. Lannoy [c], B. Ricard [c]

[a] *INRIA Futurs 91405 Orsay cedex, France*
[b] *LabSAD, EA 3698, UPMF, 1251 Avenue centrale BP47 38040 Grenoble cedex 9, France*
[c] *EDF R&D Département MRI 6, quai Watier, 78401 Chatou cedex, France*

Received 12 November 2003; accepted 18 August 2005
Available online 7 October 2005

Abstract

In this study, a Bayesian Network (BN) is considered to represent a nuclear plant mechanical system degradation. It describes a causal representation of the phenomena involved in the degradation process. Inference from such a BN needs to specify a great number of marginal and conditional probabilities. As, in the present context, information is based essentially on expert knowledge, this task becomes very complex and rapidly impossible. We present a solution, which consists of considering the BN as a log-linear model on which simplification constraints are assumed. This approach results in a considerable decrease in the number of probabilities to be given by experts. In addition, we give some simple rules to choose the most reliable probabilities. We show that making use of those rules allows to check the consistency of the derived probabilities. Moreover, we propose a feedback procedure to eliminate inconsistent probabilities. Finally, the derived probabilities that we propose to solve the equations involved in a realistic Bayesian network are expected to be reliable. The resulting methodology to design a significant and powerful BN is applied to a reactor coolant sub-component in EDF Nuclear plants in an illustrative purpose.

Keywords: Bayesian Network; Degradation process; Log-linear model; Expert opinion; Complexity reduction; Maintenance

1. Introduction

Preventive maintenance is considered in a lot of industries because costs due to failures and repairs could be very important. Moreover, since system safety is an important goal for industries, the knowledge of the degradation processes has become essential. Preventive maintenance of a system is taking into account expert knowledge, feedback observations and degradations in order

- to model the system lifetime and to quantify the degradation or failure probability,
- to detect important variables involved in the degradation process and to design maintenance tasks in order to differ or eliminate ageing,
- to quantify the effect of maintenance actions on the system behavior,
- to propose diagnosis and decision help,
- to propose data mining and sensibility analysis.

Bayesian Networks, abbreviated BNs in the following, could be thought of as useful to help engineers to fulfill those purposes of preventive maintenance. BNs provide an efficient way to represent the degradation process of an industrial system or machine. Bayesian Networks are some specific graphical models introduced by Pearl [11], and Lauritzen and Spiegelhalter [10]. Graphical Models introduce a relation between graph and probability theories. The random variables of a probabilistic model are described with the vertices of a graph, where edges describe their dependencies measured with conditional probabilities. A great interest of BNs is to provide an efficient tool for modelling in a simple and readable way the most probable links between events of different nature (expert opinion, feedback experience, …) using conditional independence between random variables. BNs which can also be regarded as a way to introduce randomness in influence diagrams can be expected to be useful for preventive

* Corresponding author.
E-mail addresses: gilles.celeux@math.u-psud.fr (G. Celeux), franck.corset@upmf-grenoble.fr (F. Corset), andre.lannoy@planetis.com (A. Lannoy), benoit.ricard@edf.fr (B. Ricard).

0951-8320/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2005.08.007

maintenance in the future. The articles of [12] and [13] are recent examples of such maintenance modelling with influence diagrams.

Bayesian Networks, by describing the main conditional probabilities between variables, allow to compute easily the joint probability distribution of all the variables involved in a complex process. However, in order to obtain this joint probability distribution, the number of required probabilities increases exponentially with the number of variables in the model. In the last decade, authors developed algorithms in order to facilitate calculations in graphical models. We can cite the well-known 'junction tree' algorithm (see Lauritzen and Spiegelhalter[10], Cowell [6] for instance). Thus many efficient algorithms to deal with computations on more and more complex graphical models are available. These algorithms are included in many software programs like, for instance, the Denmark software Hugin expert [16] and Netica [15] of Norsys Software Corp. If it is essential to develop efficient algorithms to take full advantage of the knowledge contained in graphical models, we must keep in mind that if the knowledge base is poor or badly managed the best algorithm becomes useless (see the introduction section in Cowell et al. [7]). The aim of the present paper is to deal with the practical difficulties encountered when designing a BN for a real maintenance modelling problem. All the ideas we present are motivated and illustrated with the modelling of a nuclear mechanical system degradation with a Bayesian Network.

First, we address the problem of deriving reliable information from experts in a graphical model setting and propose a method to obtain honest probability values in a simple and possibly interactive way from experts knowledge. Then we deal with a second difficulty. In order to compute the joint probability of the variable, the classical BN approach consists of asking for the marginal probabilities of the entry variables and the conditional probabilities of the other variables knowing all the possible combinations between their parents. Owing to the great number of parameters a BN can involve in practice, it can become a formidable task, and we propose to approximate the graphical model with an unsaturated log-linear model (see for instance [14], chapter 7). This procedure leads to consider that some variables are conditionally independent. In the maintenance modelling context, the experts are asked to provide the marginal probabilities of all the variables and not only of the parent variables as in the classical approach. We choose this method because generally experts can more easily provide such simpler marginal probabilities than conditional probabilities. This strategy leads to a system with more equations than unknown quantities. Thus, it allows us to assess the consistency of the required probabilities. According to Bayes theorem, the properly weighted summation of conditional probabilities of a vertex N, knowing a parent vertex, is equal to its marginal probability. After checking the consistency of the given probability values, a feedback procedure is proposed when some inconsistency is encountered, and the experts can choose the reliable given probabilities from their own viewpoint. Thus, the probabilities finally selected to solve the system of n equations with n unknown quantities are expected to be the most reliable ones.

This paper is devoted to present our heuristic strategy to build practically a graphical model in a realistic and relevant way. It is organized as follows. In Section 2, Bayesian Networks models are introduced and a parallel between graphical models and log-linear models which is helpful to reduce the BNs complexity is presented. Section 3 is devoted to the presentation of our methodology to get information from experts in a simple and reliable way. We indicate possible drawbacks of an over restrictive strategy and present ways to remedy these drawbacks. An illustrative application concerning nuclear mechanical system degradation is presented in Section 4 and a short conclusion section ends this paper.

2. Bayesian Networks

Bayesian Networks are powerful graphical models to describe conditional independence and analyze probable causal influence between random variables. In our study, variables are all discrete and most of them are binary. These random variables are represented by the vertices of the graph, and the probable influence between two random variables is represented by an edge between the corresponding vertices. We now give some definitions.

Definition 1. *A directed graph is a couple* $\mathcal{G}=(V, E)$, *where* $V=(X_1,\ldots,X_n)$ *denotes the vertices of the graph and* $E=(e_1,\ldots,e_m)$ *denotes a part of Cartesian product* $V\times V$, *where* e_i *is called the edges of the graph.*

If (X_i, X_j) lies in E, then this element is called an edge. It is denoted $X_i\rightarrow X_j$, X_i is called the source and X_j the target of the edge. For directed graph, the parents and the children of the vertices are defined as follows:

Definition 2. *If a directed edge has source* X_i *and target* X_j, *then* X_i *is called the parent of* X_j *and* X_j *is called the son or child of* X_i. *The set of the parents of* X_j *is denoted* $pa(X_j)$ *and the set of children of* X_i *is denoted* $ch(X_i)$.

In a directed graph, the oriented paths are defined as follows:

Definition 3. *An oriented path is a set of distinct vertices* $X_i,\ldots, X_j$ *such that* (X_{k-1}, X_k) *is an edge for all* $k=i+1,\ldots,j$. *This path is denoted* $X_i \mapsto X_j$. *A cycle is a path such that* $X_i=X_j$.

Directed graph without cycle are called Directed Acyclic Graphs (DAG). We can now define a Bayesian Network.

Definition 4. *A Bayesian Network is*

- *a set of variables V, defining the vertices, and a set of edges between variables E,*
- *each variable has a finite number of exclusive states,*
- *variables and edges define a directed acyclic graph, denoted* $G=(V, E)$,
- *for each variable Y with its parents* $X_1,\ldots,X_n$, *is associated a conditional probability* $P(Y|X_1,\ldots X_n)$. *When a variable has*