

BAYESIAN NETWORKS FOR RELIABILITY ANALYSIS OF COMPLEX SYSTEMS

José Gerardo Torres Toledano

Instituto de Investigaciones Eléctricas
Unidad de Resultados de Ingeniería Mecánica
Apdo.Postal 1-475 C.P. 62001
Cuernavaca, Morerlos, México
jgtorres@iie.org.mx.

Luis Enrique Sucar Succar

ITESM - Campus Morelos.
Departamento de Computación
División de Ingeniería y Ciencias
Apdo.Postal 99-C, C.P. 62050
Cuernavaca, Morelos, México
esucar@campus.mor.itesm.mx.

ABSTRACT

This paper presents an extension of Bayesian networks (BN) applied to reliability analysis. We developed a general methodology for modelling reliability of complex systems based on Bayesian networks. A reliability structure represented as a reliability block diagram is transformed to a Bayesian network representation, and with this, the reliability of the system can be obtained using probability propagation techniques. This allows for modelling complex systems, such as a bridge type, and dependencies between failures, which are difficult to obtain with conventional reliability analysis techniques. The relation between a BN and fault tree, and some advantages of BN for modeling system reliability are shown. We present some examples of the application of this methodology in solving difficult cases, which occur in reliability analysis of real systems, such as power plants.

Keywords: knowledge representation, model-based reasoning, probabilistic reasoning, Bayesian networks, reliability analysis.

BAYESIAN NETWORKS FOR RELIABILITY ANALYSIS OF COMPLEX SYSTEMS

ABSTRACT

This paper presents an extension of Bayesian networks (BN) applied to reliability analysis. We developed a general methodology for modelling reliability of complex systems based on Bayesian networks. A reliability structure represented as a reliability block diagram is transformed to a Bayesian network representation, and with this, the reliability of the system can be obtained using probability propagation techniques. This allows for modelling complex systems, such as a bridge type, and dependencies between failures, which are difficult to obtain with conventional reliability analysis techniques. The relation between a BN and fault tree, and some advantages of BN for modeling system reliability are shown. We show some examples of the application of this methodology in solving difficult cases, which occur in reliability analysis of power plants.

Keywords: knowledge representation, model-based reasoning, probabilistic reasoning, Bayesian networks, reliability analysis.

I. INTRODUCCION

Complex industrial plants and equipment for critical applications, such as power plants, require a high reliability, i.e., a very low probability of failure. For this, there are statistical techniques that can predict the reliability of a complex system based on its structure and the reliability of each component. Some traditional techniques for reliability analysis have several important limitations, including the assumption that all the failures are independent and that the rate of failure is constant (exponential model). Also, building the model used to calculate the reliability of the system is a difficult and complex task, so an expert reliability engineer is usually required.

It should be clear that failure prediction is difficult to be done. However, for a period of time given, the probability of failure can be obtained by applying probability theory. In the context of this work, reliability is the probability that the equipment performs its intended functions satisfactorily or without failure, for a mission time under specific design and environmental conditions. The reliability of complex equipment depends on the individual reliability of its elements.

The motivation for developing this work is to obtain a computational method that can incorporate explicitly dependencies between failures and include the effects of maintenance in the reliability analysis of complex systems in operation. A Bayesian network is used to represent the system reliability structure, and obtain its reliability via probability propagation. With this representation the limitations of other techniques are avoided, so it is possible to manage dependencies and non-exponential distributions.

The document is divided in six parts. The second part summarizes the theory of Bayesian networks and general aspects of reliability analysis. The third part focuses on dependency between failures in reliability analysis and the fourth part presents a procedure for systems reliability modelling supported by Bayesian networks. The following part presents an application to reliability analysis of power plants. Finally, the conclusions and future work are presented.

II. GENERAL ASPECTS OF BAYESIAN NETWORKS AND RELIABILITY ANALYSIS

A. Bayesian Network

Bayesian networks are directed acyclic graphs (DAG), see figure 1, in which the nodes represent propositions (or variables), the arcs signify direct dependencies between the linked propositions, and the strength of these dependencies are quantified by conditional probabilities [Pearl 86]. Such graphical structures, known also as influence diagrams or belief networks, are used for representing expert knowledge. The graph represents a set of random variables and its dependency and independency relations. It is used to estimate the posterior probability of unknown variables given other variables (evidence), through a process known as *probabilistic reasoning*. This generates recommendations or conclusions about a particular problem, and can be used for explanation, the process of communicating the relevant information to user.

When the graph is used for diagnosis it is called a *probabilistic expert system*. In this context, the reasoning is based on dependency relations: fault-symptoms, cause-effects, hypothesis-evidence. Every fault and symptom is modeled by random variables with a finite range of possible values. A graph is constructed with a node for each variable and it has an edge (arc) from one node, which represents a fault to another node which represents a symptom, e.g. the symptom C is dependent on A in figure 1.

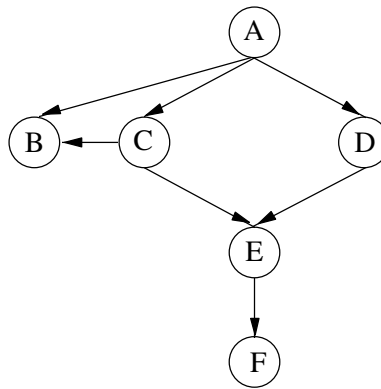


Figure 1. Example of a directed acyclic graph.

Probability propagation

The topology of a network gives the dependency relations between the variables implicated. It represents which variables are conditionally independent given another variable. Following figure 1, C is conditionally independent of D , given E if

$$P(C|D,E)=P(C|E)$$

The main advantage of Bayesian networks is the representation of the joint probability distribution of the variables, this probability can be expressed as a product of the conditional

distributions of each node given its direct influences (parents) in the graph. Hence, letting $pa(v)$ denote the parents of node v , the graph implies that the joint distribution $p(v)$ has the form:

$$p(V) = \prod_{v \in V} p(v|pa(v)) \quad (1)$$

This is also known as a recursive model with respect to some DAG. Thus, for example, the model for figure 1 is equivalent to:

$$P(A,B,C,D,E,F) = P(F|E) \cdot P(E|C,D) \cdot P(D|A) \cdot P(C|A) \cdot P(B|C,D) \cdot P(A) \quad (2)$$

The reasoning mechanism used for diagnosis is deductive. It consists in instantiating the input variables (symptoms or evidences) and propagating their effect through the network to update the probability of the hypothesis (fault) variables. The propagation procedure is based on Bayes theorem and the structure of dependencies of the network. This can be implemented through communications between neighboring nodes, by local operations, and by sending messages between connected nodes in the network [Pearl 88]. The posterior probability of any variable can be calculated by the product of the evidence (V) from its parents and its sons, using Bayes theorem.

Propagation in trees

A tree structured network has only one node, called root node, without parents and the rest of the nodes have only one parent.

In a tree, any node (C) can be a point of division in two independent sub-trees. A subtree contains as root the node of division and is denoted by ($-$), the data contained in this subtree represents the evidence V , the remainder of the tree is denoted by ($+$) with evidence V^+ [Neapolitan 90]. Therefore, the posterior probability of any variable (C) can be obtained by Bayes theorem as:

$$P(C_i|V) = P(C_i)P(V^+, V|C_i)/P(V) \quad (3)$$

But since both sub-trees are independent, and with Bayes theorem further applied, we have:

$$P(C_i|V) = \alpha P(C_i|V^+)P(V^-|C_i) \quad (4)$$

Where α is a normalization constant.

If we define:

$$\pi(C_i) = P(C_i|V^+) \quad (5)$$

$$\lambda(C_i) = P(V^-|C_i) \quad (6)$$

Replacing the equations (5) and (6) in (4), we obtain:

$$P(C_i | V) = \propto \pi(C_i) \lambda(C_i) \quad (7)$$

The above equation offers a way to update the probabilities of any node C as a product of the evidence V of its ascendant nodes (π) and of its descendant nodes (λ).

Propagation in polytrees or simply networks connected

A polytree is a network in which a node can have more than one parent, without multiple paths between nodes (figure 2).

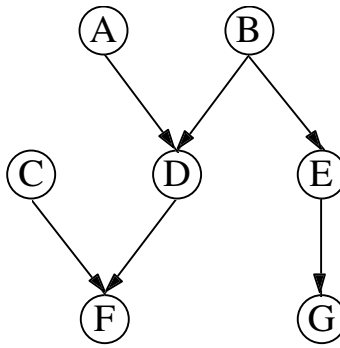


Figure 2. Example of polytree.

The propagation of probabilities in polytree structures is very similar to the case of tree networks [Neapolitan 90]. The principal difference is that polytrees require the conditional probability of each node given all its parents nodes. In similar form that the treatment of networks with tree structure, for the case of polytrees an expression to obtain the probability of any node given some evidence can be deduced [Pearl 88].

Consider a typical fragment of a singly-connected network, consisting of a node B_i , the set of all parents of B_i , $V^+ = \{V_1^+, \dots, V_l^+\}$, and the set of all children of B_i , $V^- = \{V_1^-, \dots, V_m^-\}$. As before, let V be the total evidence obtained, so that:

$$P(B_i | V) = \alpha P(B_i | V_1^+, \dots, V_l^+) P(V_n^- | B_i) \cdot P(V_m^- | B_i) \quad (8)$$

Dividing the polytree in two parts, V^+ and V^- , it is possible to obtain a mechanism for local probability propagation similar to the one for trees.

The algorithm for probability propagation in polytrees is very efficient, so the computation time is nearly proportional to the diameter (largest path) of the network. For multiconnected networks probability propagation is more complex and there are several algorithms based on clustering, conditioning and stochastic simulation [Pearl 88, Neapolitan 90].

B. System Reliability

In reliability analysis, we can distinguish three characteristic types of failures which may be inherent in the behavior of the equipment [Bazovsky 61]. First, there are the failures which occur early in the life of a component, they are called *early failures* and in the majority of the cases are the result of a poor manufacturing and quality-control techniques during the production process. Secondly, there are failures which are caused by wearout of parts. These occur in an equipment only if it is not properly maintained or not maintained at all. Thirdly, there are the so called "chance" failures. These failures are caused by sudden cumulative stress summation beyond the design strength of the component. Chance failures occur at random intervals, irregularly and unexpectedly.

Reliability analysis differentiates between early, wearout, and chance failures for two main reasons. First, each one of these types of failures follows a specific statistical distribution and therefore requires a different mathematical treatment. Second, different methods must be used for their elimination or correction.

In reliability analysis of a complex system, is nearly impossible to model the complete system. The logical process to accomplish this is to divide the system in smaller elements, units, subsystems, or components. The main assumption is that every entity has two states, success and failure (although some times three or more are needed). The subdivision generates a "block diagram" that is similar to the description of systems in operation [Fullwood, Hall 88]. The models are then fixed to this structure, and they utilize probabilistic techniques to calculate the reliability of the system in terms of the reliability of the subdivisions [Shooman 68].

To evaluate the adequate performance, an observation of inadequate performance in operation is required, therefore, the frequency at which malfunctions and failures occurs it is used as a parameter for a mathematical formulation of reliability. This parameter is called *failure rate*; it is usually measured in number of failures per unit operating hour. Its reciprocal value is called the *mean time between failures* and this is measured in hours [Bazovsky 61].

III. DEPENDENCY BETWEEN FAILURES IN RELIABILITY ANALYSIS

Our objective is to build a versatile computational tool capable of evaluating the reliability during useful life or wear out of complex systems. Traditionally, fault trees [Fullwood, Hall 88] are used for reliability analysis. However, this technique has its limitations. It usually assumes independent events and it is difficult to model dependency between events or faults.

Dependent events can be found in reliability analysis in the following cases:

- 1) **Common causes.** Some condition or event which provokes multiple elemental failures is called a common cause. For instance, fire or flood may cause simultaneous failures of sets of components. Thus, under these conditions, component failures are no longer independent. Other sources of common cause are aging, human error and system environment in general.

- 2) **Mutually exclusive primary events.** Consider the basic events: “switch fails to close” and “switch fails to open”. These two basic events are *mutually exclusive*, i.e., the occurrence of one basic event precludes another. Thus we encounter dependent basic events when a fault tree involves mutually exclusive basic events.
- 3) **Standby redundancies.** When an operating component fails, a standby component is put into operation, and the redundant configuration continues to function. Thus, components failures are not statistically independent, since the failure of an operating component causes a standby component to be more susceptible to failure.
- 4) **Components supporting loads.** Assume that a set of components supports loads such as stresses, currents, etc. A failure of one component increases the load supported by the other components. Consequently, the remaining components are more likely to fail, and we can not assume statistical independence of components.

Bayesian networks allow to represent explicitly dependencies between failures as above mentioned. We suggest to employ this approach to solve reliability analysis of complex systems, in particular when there are dependent failures.

IV. PROCEDURE FOR SYSTEM RELIABILITY MODELING

The procedure suggested for reliability analysis based on Bayesian networks consists in defining the conditional probability matrix equivalent to series-parallel configurations of simple systems (as the AND-OR gates utilized in fault trees). Being the reliability block diagram a methodology commonly used for reliability analysis, we will refer it to introduce the representation with BN. Reliability analysis begins with the construction of a reliability block diagram of the system which is a graphic representation where every component is represented as a block or rectangle connected to other components, in series or in parallel form.

Considering a series or parallel system with only two components, figure 3, its representation as Bayesian network looks like figure 4 with one additional node, X . We use circles for representing series systems and squares for parallel system.

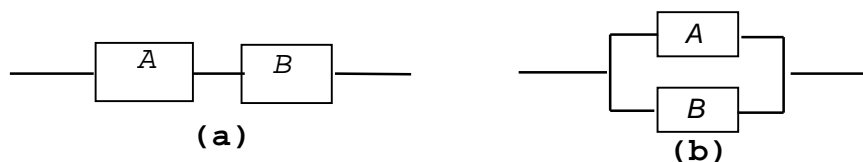


Figure 3. System Reliability Block Diagram: (a) series, (b) parallel.

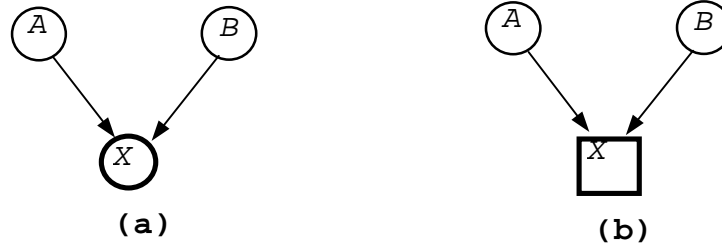


Figure 4. Bayesian Network for two components (a) series system, (b) parallel system

According to equation (1), the joint probability of the series system is:

$$P(X,A,B) = P(X/A,B) P(A,B) \quad (9)$$

where the elements of the columns of the conditional probability matrix, $P(X/A,B)$, are the combination of the parent nodes states: A,B are operating, A is operating and B is failed, A is failed and B is operating, and A, B are failed. The first row represents the success probability of the system given the information of A and B . This matrix is equivalent to an AND gate:

$$P(X/A,B) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (10)$$

The elements of the $P(A,B)$ matrix are taken from the marginal probability $P(A)$ and $P(B)$, for example $P(a' b)$ is the probability that component A is in a failure state and component B in an operating state.

$$P(A,B) = \begin{bmatrix} a & b \\ a'b \\ a b' \\ a'b' \end{bmatrix} \quad (11)$$

In the parallel case only the conditional probability matrix is modified, such matrix is equivalent to an OR gate [Schwarzblat *et al* 80]:

$$P(X / A,B) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (12)$$

Following the above scheme, the generalization for multiple components is not difficult. For instance, for a three component system, which require two components functioning, the probability matrix will be:

$$P(X/A,B,C) = \begin{bmatrix} 11110000 \\ 00001111 \end{bmatrix}$$

V. APPLICATIONS TO RELIABILITY ANALYSIS

A. Complex combination of serie-parallel system

In order to exemplify the advantages of using BN, consider the schematic reliability block diagram in figure 5. This system is known as bridge type. The system is operable if at least one of the paths AC, BD, AED o BEC are good.

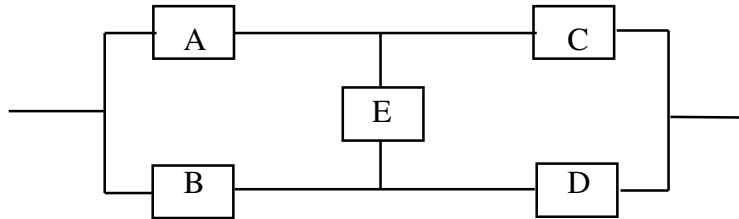


Figure 5. Reliability block diagram.

The usual method to compute the reliability is selecting a component, and consider two alternatives: the component is working (good) or the component has failed (bad) [Bazovsky 1961]. In this case the E element is chosen, which is the best choice to simplify the solution. The system is divided in two subsystems, one when E is considered as good, and other where E has failed.

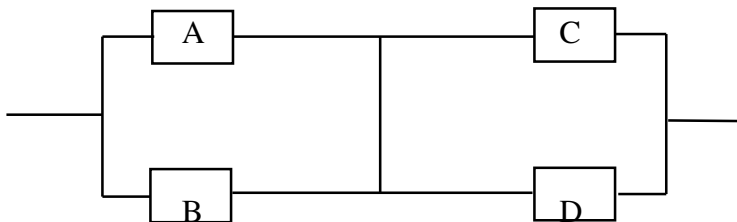


Figure 6 (a). Subsystem considering E "good".

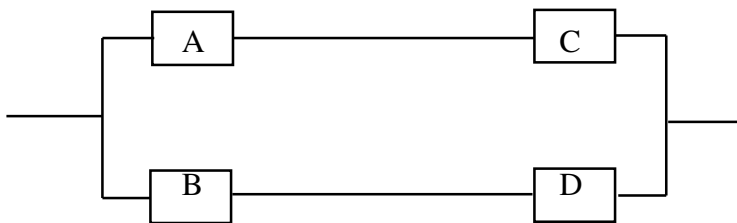


Figure 6 (b). Subsystem considering E "bad".

When a set of subsystems are defined utilizing series-parallel connected configurations, the total reliability could be evaluate applying Bayes' theorem. The probability distribution in terms of conditional probabilities is $P(X) = P(X/E=good) P(E=good) + P(X/E=bad) P(E=bad)$. See figure 6(a) and 6(b).

The preceding method is laborious. However, using a BN approach the solution is simplified so the system reliability can be obtained from single network. A graphic representation for the bridge system in the scheme of BN is shown in figure 7. We have developed an algorithm for buiding automatically a BN representation from the reliability block diagram.

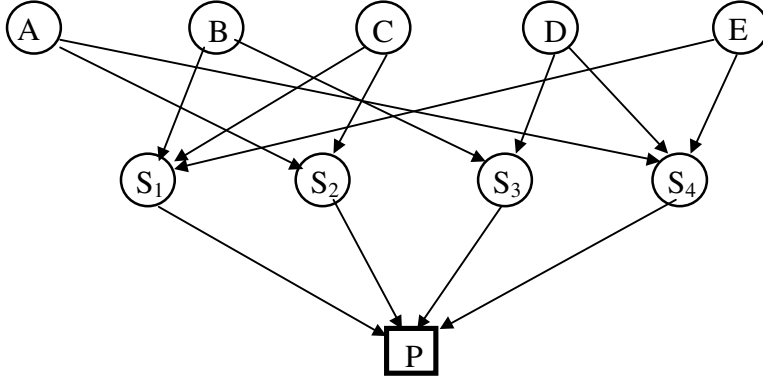


Figure 7. Bayesian Network of a complex system.

For example, if the A,B,C,D, and E components have 0.1 as success probability value, the system reliability is **0.9785** and its failure probability **0.0215**. To obtain these values, we apply probability propagation techniques (in this case, it is a multiconnected network) to the network in figure 7, obtaining the probabilities for the intermediate nodes (S_i) and for the complete system (P). For this particular case, the results for the subsystems are:

$$\begin{aligned} P(S_1) &= P(S_4) = (0.729, 0.271) \\ P(S_2) &= P(S_3) = (0.810, 0.190) \end{aligned}$$

B. Reliability of dependent components

Suppose three independent sources of shock are present in the environment [Henley and Kumamoto, 1992]. A shock from source 1 destroys component 1; it occurs at a random time U_1 , where $P[U_1 > t] = e^{-\lambda_1 t}$. A shock from source 2 destroys component 2; it occurs at random time U_2 , $P[U_2 > t] = e^{-\lambda_2 t}$. Finally a shock from source 3 destroys *both* components, it occurs at random time U_{12} , where $P[U_{12} > t] = e^{-\lambda_{12} t}$. Thus the random life length T_1 of component 1 satisfies:

$$T_1 = \min(U_1, U_{12}),$$

while the random life length T_2 of component 2 satisfies:

$$T_2 = \min(U_2, U_{12})$$

A BN model for this example of dependent failures is shown in figure 8, where S_i represents the i -th source and C_i the i -th component. The system states are assigned to X . In this case, all the conditional probability matrix are defined equivalent to AND gates.

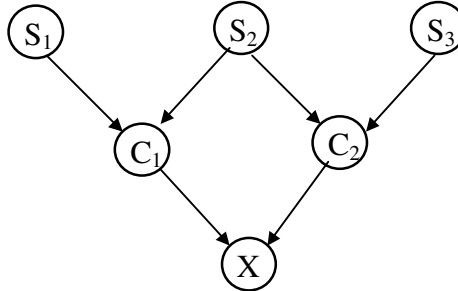


Figure 8. Bayesian network of a system with common cause failures.

Reliability and failure probability of the system is obtain by applying the conventional procedure for probability propagation for multiconnected networks. Reliability results for particular values are shown in table 1.

Node	Reliability
S_1	0.9417
S_2	0.9048
S_3	0.9980
C_1	0.9398
C_2	0.9030
X	0.8504

Table 1. Reliability (probability of success) for a system with common cause failures.

VI. CONCLUSIONS

Bayesian networks are an alternative technique for the systems reliability analysis with an ample potential of application. They are based on the management of conditional probability and on probability propagation. BN have a strong similirity to fault trees. In fact, fault trees could be viewed as a specific case of BN. One of the advantages of using Bayesian networks is the explicit representation of dependencies.

In this paper we have presented a general methodology for modelling reliability of complex systems based on Bayesian networks. A reliability structure represented as a reliability block diagram can be transformed to a Bayesian network representation, and with this, the reliability of the system can be obtained using probability propagation techniques. This allows for modelling complex systems, such as a bridge type, and dependencies between failures, which are difficult to obtain with conventional reliability analysis techniques.

This approach also allows a combination of information sources (objective and subjective) and the selection of the best probabilistic model according to the distribution and the structure of the system. The combination of information sources could be applied to avoid the lack of information in data bases of certain areas for reliability analysis. For example, in the case of the majority of the power plants, the information is augmented with the estimates obtained by operators or maintenance personal. The combination of these sources permits to increase the precision of the system reliability estimation.

Another future direction for research is to use this type of models for design. In this case, we can set the desired reliability of the system and obtain the required reliability of each component, using the same probability propagation techniques.

REFERENCES

[Bazovsky 61] Igor Bazovsky, (1961). Reliability Theory and Practice. Prentice Hall.

[Fullwood, Hall 88] Ralph R. Fullwood and Robert E. Hall, (1988). Probabilistic risk Assessment in the nuclear power industry, Fundamentals & Applications. Pergamon Press.

[Henley, Kumamoto 92] E. J. Henley and H. Kumamoto, (1992). Probabilistic risk Assessment. IEEE Press.

[IEEE 87] IEEE, (197). IEEE Guide for General Principles of Reliability Analysis of Nuclear Power Generating Station Safety Systems. ANSI/IEEE, Std 352-1987.

[Neapolitan 90] Richard E. Neapolitan, (1990). Probabilistic Reasoning in Expert Systems, theory and algorithms. John Wiley & Sons.

[Pearl 86] Judea Pearl, (1986). Fusion, Propagation, and Structuring in Belief Networks. Artificial Intelligence 29(3), 241-288.

[Pearl 88] Judea Pearl, (1988). Probabilistic Reasoning in Intelligent Systems: Network of Plausible Inference. Morgan Kaufmann.

[Shooman 68] Martin L. Shooman, (1968). Probabilistic Reliability: An Engineering Approach. McGraw-Hill.

[Schwarzblat *et al* 80] M. Schwarzblat, J. Arellano, G. Martínez, (1980). Análisis de confiabilidad en el IIE, Seminario sobre confiabilidad en plantas termoeléctricas, Instituto de Investigaciones Eléctricas, Cuernavaca, Mor., México (in Spanish).