A probabilistic relational model approach for fault trees modeling

Thierno Kante, Philippe Leray

A Probabilistic Relational Model approach for Fault Tree modeling

Thierno Kante[1,2] and Philippe Leray[1]

[1] LS2N (UMR CNRS 6004), DUKe Research Group, University of Nantes, France.
[2] EDICIA, Carquefou, France.
thierno-sadou.kante@univ-nantes.fr, philippe.leray@univ-nantes.fr

Abstract. Fault Trees or Bow Tie Diagrams are widely used for system dependability assessment. Some probabilistic extensions have been proposed by using Bayesian network formalism. This article proposes a general modeling approach under the form of a probabilistic relational model (PRM), relational extension of Bayesian networks, that can represent any fault tree, defined as an event tree with possible safety barriers, simply described in a relational database. We first describe an underlying relational schema describing a generic fault tree, and the probabilistic dependencies needed to model the existence of an event given the possible existence of its related causes and eventual safety barriers.

Keywords: Fault trees, Bow Tie diagram, Bayesian Network, Probabilistic Relational Model

1 Introduction

Risk prevention has always been a major concern in many areas such as industrial system, offshore and public security... Nowadays, different approaches for risk analysis in the areas of Dependability (operating reliability) have been proposed in the literature such as preliminary risk analysis (PRA), Petri networks, Bow-Tie method, Fault Tree method.

Fault Trees or Bow Tie Diagrams are widely used for system dependability assessment. Some probabilistic extensions have been proposed by using Bayesian network formalism. This article proposes a general modeling approach under the form of a probabilistic relational model (PRM), relational extension of Bayesian networks, that can represent any fault tree, defined as an event tree with possible safety barriers, simply described in a relational database. We first describe the underlying relational schema describing a generic fault tree, and the probabilistic dependencies needed to model the existence of an event given the possible existence of its related causes and eventual safety barriers.

The rest of the paper is organized as follows: Section 2 presents the background about Fault Trees, Probabilistic Graphical Models (PGMs) such as Probabilistic Relational Models, and related works about use of PGMs for reliability. Section 3 presents our contribution, starting from a description of a Fault Tree,

and building a generic PRM from its relational schema to the probabilistic dependencies modeling the initial fault tree. We summarize our work and discuss open questions and perspectives offered by our contribution in Section 5.

2 Background

2.1 Fault Trees (FTs)

Several formalisms have been proposed in Systems Dependability Assessment. They are classified into two categories: combinatorial models (as fault trees or reliability block diagrams) and state-space models (as Markov chain or Petri nets). Among the combinatorial models, FT is one of the most popular and diffused formalisms for analysis of large, safety critical systems [9]. A FT is synthetically defined by all combinations of events that can lead to failure. This search of combinations of events that can cause a failure continues with a search of minimum cut-sets (sets of basic events, or conditions, necessary and sufficient to produce the failure) and then an evaluation of the likelihood of the occurrence of the failure from the combination of the likelihood that elementary events occur. The FT modeling is based on a descending approach (top-down approach). It is based on the following assumptions: (i) events are binary events (working/not-working); (ii) events are statistically independent; and (iii) relationships between events and causes are represented by means of logical gates [1]. In risk analysis, to assess the impacts of an undesired event, an event tree (ET) is added to the FT, resulting in bow tie model. The goal is to place and assess barriers to prevent or protect from the undesired event. An example of Fault tree associated with safety barriers is described in Figure 1.

However, the FTs and Bow tie diagrams are static models. Dynamic fault trees have been proposed to extend standard FTs to dynamic systems [5].

2.2 Probabilistic Relational models

For representing uncertain knowledge, Probabilistic Graphical Models, in particular, Bayesian networks (BNs) are increasingly used in the field of artificial intelligence [11]. They are a powerful modeling and analysis tool that has been

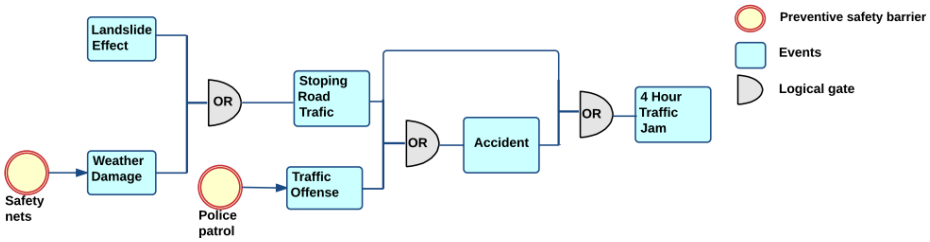

Fig. 1. Example of a Fault Tree with safety barriers.

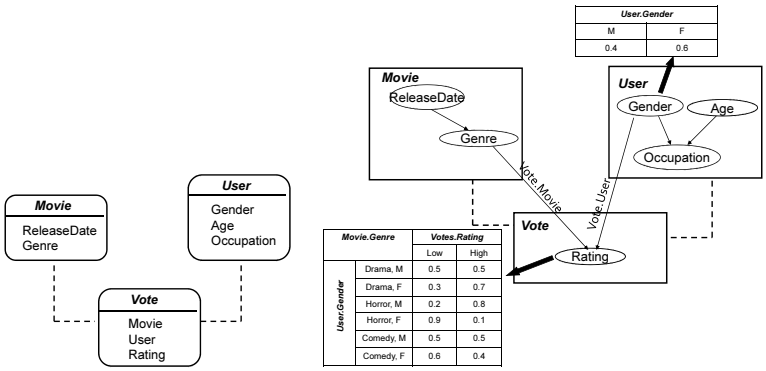

Fig. 2. (left) An example of relational schema. (right) An example of Probabilistic Relational Model.

applied in a variety of real-world tasks. Bayesian networks have been extended in order to model more complex problems, such as dynamic ones with Dynamic Bayesian networks, object-oriented ones or relational ones with Probabilistic Relational Models.

As defined in [7], A Probabilistic Relational Model (PRM) Π for a relational schema $\mathcal{R}$ (i.e., set of entities and relations) is defined through a qualitative dependency structure $\mathcal{S}$ and a set of parameters associated with it $\theta_{\mathcal{S}}$. The relational schema $\mathcal{R}$ describes a set of classes $\mathcal{X} = \{X_1, \ldots, X_1\}$, each of which has a set of descriptive attributes denoted by $\mathcal{A}(X)$, which take on a range of values $\mathcal{V}(X.A)$ and a set of reference slots denoted by $\mathcal{R}(X) = \{\rho_1 \ldots \rho_k\}$. Each $X.\rho$ has X as domain type and Y as a range type, where $Y \in \mathcal{X}$. A sequence of slots $\rho_1 \ldots \rho_k$, where $\forall i, Range[\rho_i] = Dom[\rho_{i+1}]$ defines a slot chain K. The notion of aggregation is also adopted from the database theory: an aggregate γ takes a multi-set of values of some ground type, and returns a summary of it, a single-valued attribute is derived from the aggregation function.

Formally, a PRM Π is defined as follows. For each class $X \in \mathcal{X}$ and each descriptive attribute $A \in \mathcal{A}(X)$, we have:

- A set of parents $Pa(X.A) = \{U_1, \ldots, U_l\}$, where each U_i has the form $X.B$ if it is a simple attribute in the same relation or $\gamma(X.K.B)$, where K is a slot chain and γ is an aggregation function.
- A legal conditional probability distribution (CPD), $P(X.A|Pa(X.A))$.

An example is described in Figure 2 (right) for the relational schema depicted in Figure 2 (left). Probabilistic inference is performed on a *Ground Bayesian Network* (GBN) obtained from a PRM for the given database instance $\mathcal{I}$. A GBN is generated by a process (also called unrolling) of copying the associated PRM for every object in $\mathcal{I}$. Thus a GBN will have a node for every attribute of every object in $\mathcal{I}$ and probabilistic dependencies and CPDs as in the PRM.

2.3 Related work

The formalism of BNs is well suited to represent complex multi-state systems. Recent works have shown that reliability formalisms such as event trees, fault trees (FTs) or Bow Tie diagram are easy to model by an equivalent BN. For example, [13] has shown that a reliability structure represented as a reliability block diagram can be transformed into a Bayesian network model. This approach makes it possible to compute the reliability of the system using probabilistic inference in the equivalent BN. Similar works have proposed a language allowing to transform fault trees or Bow Tie diagrams into Bayesian networks [1, 6, 8]. However, these approaches do not allow to model the dynamic aspect of the system. In [2, 12, 14], a description of a Dynamic fault tree (DFT) with a Dynamic Bayesian network (DBN) has been proposed. These works consider time as a discrete variable and describe temporal probabilistic dependencies with Markov chain. A generalization to continuous time has been proposed by [10, 3].

The majority of the previously cited methods deal with Boolean variables (existence of an event). A few of them consider the notion of barrier, and when this barrier is defined, its existence is also Boolean. In addition, BNs are not adapted to model large and complex domains because the structure of the network is fixed in advance. Thus, no part is reusable and therefore explicitly requires rewriting structure or parameter regularity. The data and the model are not decoupled, so taking into account a new component requires updating BN model by an expert or a complete learning of the model. As in the oriented-object framework used in [14], Probabilistic relational models (PRMs) improve the possibilities of generalization in this direction.

3 Contribution

We propose here a general modeling approach under the form of a probabilistic relational model (PRM) that can represent any fault tree, defined as an event tree with possible safety barriers, simply described in a relational database.

3.1 Fault Tree modeling

We suppose that our FT is defined by a triplet $(\mathcal{E}, \mathcal{G}, \mathcal{B})$.

$\mathcal{E} = \{E_i\}$ is a set of events, with a prior probability $PriorStrength(E_i)$ defined in a set of ordered discrete values $\{absent, low, ...strong\}$.

$\mathcal{G} = \{G_j\}$ is a set of gates, with $Inputs(G_i) \subset \mathcal{E}$, $Output(G_i) \in \mathcal{E}$, $Type(G_i) \in \{OR, AND, ...\}$ and $DependencyStrength(E_i, output(G_j)) \in \{$*absent, low, ... strong*$\}$ for each $E_i \in Inputs(G_j)$.

$\mathcal{B} = \{B_k, (E_i, G_j)\}$ is a set of barriers. A barrier B_k is associated to one specific event E_i appearing as an input of a given gate G_j, with a $BarrierStrength(B_k, E_i, G_j) \in \{$*absent, low, ... strong*$\}$. We can notice here that our definition of barrier is related to one association input-output for a given gate, more general than the usual one where a barrier is only describing an effect on a gate output.

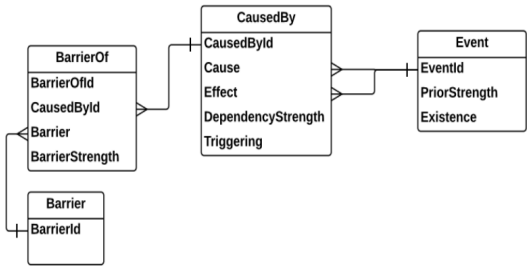

Fig. 3. Relational schema for Fault Tree modeling.

3.2 A PRM for Fault Tree modeling

Relational schema From the previously defined Fault Tree, we propose a relational schema described in Figure 3 with two entity classes *Event*, class of events, and *Barrier*, class of barriers. and two association classes *CausedBy*, association between events, and *BarrierOf* association between a barrier and one association of events.

Instances of the classes are defined by the following rules: (i) one instance of *Event* for each $E_i \in \mathcal{E}$, (ii) one instance of *Barrier* for each $B_k \in \mathcal{B}$, (iii) one *CausedBy* instance for each $G_j \in \mathcal{G}$ and $E_i \in Inputs(G_j)$ with $Cause = E_i$ and $Effect = Output(G_j)$, and (iv) one *BarrierOf* instance for each $\{B_k, (E_i, G_j)\} \in \mathcal{B}$ with $Barr = B_k$ and $CauseById$ is the instance related to gate G_j and input E_i.

Probabilistic dependencies The probabilistic dependencies are defined over the corresponding attributes of the previous classes. $Event.existence \in \{absent,$ *low, ... strong* $\}$ represents the potential existence of an event. This attribute can be observed, or will be estimated depending on its prior strength ($Event.PriorStrength$) and the existence of the events than can raise it in the fault tree.

We propose to model the logical gate between an event and its possible causes by an ICI (independence of causal influence) model [4] by adding $CausedBy.Triggering$ attribute as an inhibitor node between each cause and the effect node. This attribute $Triggering$ has the same domain than $Event.Existence$. This model will correspond to probabilistic dependencies between $CausedBy.Triggering$ and $CausedBy.Cause.Existence$, and deterministic function γ (determined by gate type) between $Event.existence$ and the set of possible triggering associations $Event.Effect^{-1}.Triggering$. For instance the deterministic function corresponding to an OR gate is the max function.

The triggering will be weighted by $CausedBy.DependencyStrength$ or inhibited by the strength of the associated barriers $CausedBy.CauseById^{-1}.BarrierStrength$. As this association can possibly be inhibited by several barriers, we decide here to use the max aggregation function to merge the effects of these possible barriers.

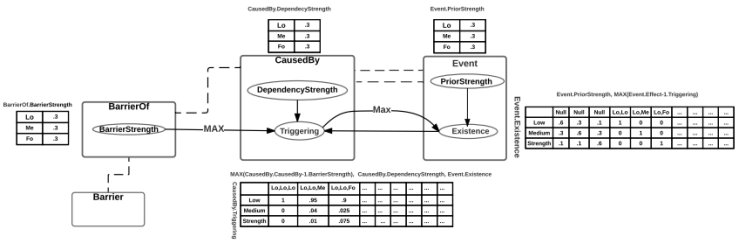

Fig. 4. Probabilistic relational model for Fault Tree modeling.

Conditional Probability distributions The conditional probability distribution (CPD) $P(Event.Existence \mid PriorStrength, \gamma(Event.Effect^{-1}.Triggering))$ is defined by a simple dependency. When the corresponding event is a root event, $\gamma(Event.Effect^{-1}.Triggering)) = NULL$ and this CPD is an increasing function depending only on $PriorStrength$. In the opposite, when this event is not a root event, this CPD is independent from the $PriorStrength$ and corresponds only to the deterministic function γ.

The conditional probability distribution concerning $CausedBy.Triggering$, $P(CausedBy.Triggering \mid CausedBy.Cause.Existence, ...$ $CausedBy.DependencyStrength, max(CausedBy.CauseById^{-1}.BarrierStrength))$, is defined by two components. First the dependency between $Caused.Triggering$ and $CausedBy.Cause.existence$ is parametrized like in any ICI model, where the strength of each cause is here weighted by the $CausedBy.DependencyStrength$ or inhibited by the several possible $BarrierStrength$.

The distributions $P(Event.PriorStrength)$, $P(CausedBy.DependencyStrength)$, $P(Barrier.BarrierStrength)$ correspond to probability distribution of observed root attributes, so their exact definition has no impact in our model. We choose here uniform distributions.

Ground Bayesian network Figure 4 shows the corresponding probabilistic relational model (PRM) defined with its relational schema, its associated probabilistic dependencies and conditional probability distributions. As defined in section 2.2, probabilistic inference is performed on the Ground Bayesian Network obtained from a PRM by unrolling the PRM template model for each instance of each class in the database.

We present here a simple example with the description of a FT (Figure 1) in the database of Figure 5 with 6 events, 3 OR gates and 2 barriers. Figure 6 describes the ground BN obtained from our PRM for this FT. Given a set of $Event.Existence$ and possible $BarrierOf.BarrierStrength$, the GBN can finally be queried to estimate the probability of other $Event.Existence$.

Figure 6 presents two scenarios in the same context where *Landslide effect* is low, *Weather damage* is medium and *Traffic offense* is strong. In the first scenario, we consider that the two barriers are low, and we observe that the

Event

EventId	PriorStrength	Existence
Landslide effect	low	?
Weather damage	low	?
Stopping road traffic	medium	?
4 hour traffic jam	medium	?
Traffic offense	medium	?
Accident	low	?

BarrierOf

BarrierOfId	BarrierId	CausedById	BarrierStrength
BO1	Safety nets	C2	?
BO2	Police patrol	C5	?

Barrier

BarrierId
Safety nets
Police patrol

CausedBy

CausedById	Cause	Effect	DependencyStrength	Triggering
C1	Landslide effect	Stopping road traffic	strong	?
C2	Weather damage	Stopping road traffic	strong	?
C3	Stopping road traffic	Accident	medium	?
C4	Stopping road traffic	4 hour traffic jam	strong	?
C5	Traffic offense	Accident	medium	?
C6	Accident	4 hour traffic jam	strong	?

Fig. 5. Instantiations of the relational schema describing the FT model of Figure 1.

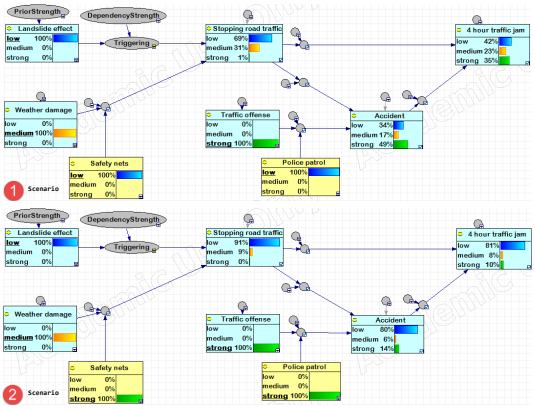

Fig. 6. Ground Bayesian network obtained by unrolling the PRM of Figure 4 on the instance given in Figure 5, with two scenarios of probabilistic inference.

probability of having a low *Stopping road traffic* is 69%, the probability of a strong *Accident* is 49%, and the probability of a low *4 hour traffic jam* is 42%.

In the second scenario, we consider strong barriers, and we observe that the probability of having a low *Stopping road traffic* is 91% (increasing because of the *Safety net* barrier), the probability of a strong *Accident* is 14% (decreasing because of the *Police patrol* barrier), and the probability of a low *4 hour traffic jam* is 81% (because of the cumulative effects of both barriers).

4 Conclusion and Perspectives

This preliminary work proposes a general modeling approach under the form of a probabilistic relational model, that can represent any fault tree, defined as an

event tree with possible safety barriers, simply described in a relational database. We first describe the underlying relational schema used to model a generic fault tree, and the probabilistic dependencies needed to model the existence of an event given the possible existence of its related causes and eventual barriers.

The way we model barriers in this work is more general than a simple inhibition of a gate output. Our barriers can (totally or partially) inhibit any input of a logical gate. As already proposed in the literature, we also use ICI models (such as NoisyMax) in order to deal with probabilistic extensions of the logical gates used in Fault Trees. With our proposal, adding new events, gates or barriers simply consists in adding new instances in the database, and generating a new ground Bayesian network where probabilistic inference can then be performed.

References

1. A. Bobbio, L. Portinale, M. Minichino, and E. Ciancamerla. Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *RESS*, 71(3):249–260, 2001.
2. H. Boudali and J.B. Dugan. A discrete-time Bayesian network reliability modeling and analysis framework. *RESS*, 87(3):337–349, 2005.
3. H. Boudali and J.B. Dugan. A continuous-time Bayesian network reliability modeling, and analysis framework. *IEEE Transaction on reliability*, 55(1):86–97, 2006.
4. F.J. Díez and M.J. Druzdzel. Canonical probabilistic models for knowledge engineering. Technical report, Research Centre on Intelligent Decision-Support Systems, 2000.
5. J.B. Dugan, S.J. Bavuso, and M.A Boyd. Dynamic fault-tree models for fault-tolerant computer systems. *IEEE Transactions on reliability*, 41(3):363–377, 1992.
6. C. Duval, G. Fallet-Fidry, B. Iung, P. Weber, and E. Levrat. A Bayesian network-based integrated risk analysis approach for industrial systems: application to heat sink system and prospects development. *Proc. of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, page 1748006X12451091, 2012.
7. N. Friedman, L. Getoor, D. Koller, and A. Pfeffer. Learning probabilistic relational models. In *IJCAI*, volume 99, pages 1300–1309, 1999.
8. N. Khakzad, F. Khan, and P. Amyotte. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection*, 91(1):46–53, 2013.
9. N.G. Leveson. *System safety and computers*. Addison-Wesley, 1995.
10. S. Montani, L. Portinale, A. Bobbio, and D Codetta-Raiteri. Radyban: A tool for reliability analysis of dynamic fault trees through conversion into dynamic Bayesian networks. *RESS*, 93(7):922–932, 2008.
11. J. Pearl. *Probabilistic reasoning in intelligent systems: Networks of plausible reasoning*. Morgan Kaufmann Publishers, Los Altos, 1988.
12. L. Portinale, D. Codetta-Raiteri, and S. Montani. Supporting reliability engineers in exploiting the power of dynamic Bayesian networks. *International journal of approximate reasoning*, 51(2):179–195, 2010.
13. J.G. Torres-Toledano and L.E. Sucar. Bayesian networks for reliability analysis of complex systems. In *Ibero-American Conference on Artificial Intelligence*, pages 195–206. Springer, 1998.
14. P. Weber and L. Jouffe. Complex system reliability modelling with dynamic object oriented Bayesian networks (doobn). *RESS*, 91(2):149–162, 2006.