

Risk Assessment Approach for EM Resilience in Complex Systems Using Bayesian Networks

Lokesh Devaraj
HORIBA MIRA Limited
Nuneaton, UK
lokesh.devaraj@horiba-mira.com

Alistair P. Duffy
De Montfort University
Leicester, UK
apd@dmu.ac.uk

Alastair R. Ruddle
HORIBA MIRA Limited
Nuneaton, UK
alastair.ruddle@horiba-mira.com

Anthony J.M. Martin
HORIBA MIRA Limited
Nuneaton, UK
anthony.martin@horiba-mira.com

Abstract—Current trends in the automotive industry are reshaping the architectures and electromagnetic characteristics of road vehicles. Increasing electrification and connectivity are enabling considerable packaging flexibility and leading to radically different electromagnetic environments. At the same time, increasing automation of driving functions will require unprecedented levels of system dependability. However, existing EMC engineering processes were developed in a very different world of low system complexity and incremental technological development. In order to adapt to rising system complexity and the increasingly rapid pace of technological change, it is considered that a more agile risk-based approach is better suited to ensure the electromagnetic resilience of future vehicles and other complex systems. This paper outlines a Bayesian network approach that allows the combination of both technical and non-technical aspects in assessing the likelihood of issues that could lead to system-level risks. This approach could be used to help achieve EM resilience from the earliest stages of product development, where the detailed information required to undertake detailed risk assessment is generally unavailable.

Keywords—Automotive, Bayesian Network, EM resilience, likelihood estimation, risk assessment.

I. INTRODUCTION

Most industries (e.g. automotive, medical, maritime etc.) currently follow a predefined rule-based approach to EMC engineering. In this scheme the system and its constituent components, which are usually procured from a large number of suppliers, are required to comply with prescriptive product test standards for electromagnetic (EM) interoperability characteristics (i.e. immunity and emission, both radiated and conducted, electrostatic discharge etc.). These test standards specify exactly what is to be measured, how it is to be tested, and the performance criteria that must be satisfied. Design guidelines, which have been developed over many years, are widely used to help engineers to achieve compliance with the test requirements. However, rapid technological change, often resulting from electrification to alleviate environmental concerns as well as increasing reliance on programmable electronic control systems and wireless communications, are leading to rapidly rising system complexity and making the traditional prescriptive assurance approach increasingly untenable.

Moreover, due to technical and economic limitations, it is not practicable to test the performance of all possible states of a complex system and its functions over the entire EM spectrum and modulations. These concerns have motivated interest in adopting a risk-based approach as an alternative to the existing prescriptive rule-based approach [1]–[2].

In the risk-based approach, system-level risks associated with dependability attributes, such as safety, security, reliability, availability etc., need to be analyzed in order to identify the potential for negative impacts on the system stakeholders. The standard for risk management, IEC 31010 [3], lists a number of tools and techniques that can be used for the estimation of the likelihood of risks. A more specific list compiled for EMC and functional safety can also be found in IEEE 1848 [4].

However, the methods proposed in this paper are of wider applicability, aiming to support:

- 1) the consideration of a wider range of potential risks to EM resilience, encompassing not just EMC for functional safety, but also possible EM impacts in terms of mission critical functionality, including communications system performance, as well as health and safety aspects such as human exposure to electromagnetic fields;
- 2) preliminary risk assessments initiated in the earliest stages, where the information required to undertake a detailed risk assessment is generally unavailable, but which can subsequently be refined over time.

The selection of suitable risk analysis technique(s) requires consideration of the system complexity as well as the potential lack of system EM knowledge and insufficient and uncertain data. A number of *technical aspects* (such as the number and range of frequencies tested for EM immunity, robustness by design, EMC verification results, intentional EMI etc.) that may be useful for estimating the risks are also influenced by *non-technical aspects* (such as time and resource constraints, provenance of the component supplier etc.). Hence, it is essential to consider both aspects in the risk analysis model for better understanding of the risks posed, and consequently greater confidence during the decision-making process.

From a safety perspective, the application of probabilistic graphical models (PGMs) for system-level risk analysis was proposed in [5]. This paper focuses on demonstrating the PGM approach to estimate the likelihood of system-level risks by combining technical and non-technical influences.

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 812790 (MSCA-ETN PETER). This publication reflects only the authors' view, exempting the European Union from any liability. Project website: <http://etn-peter.eu/>.

A simple example of such a graphical model is introduced in Section II, representing a few aspects relating to a typical item in a complex system. In Section III, the basic concepts of a Bayesian network (BN) are discussed with reference to the example model and the procedure for choosing an appropriate graphical model structure is outlined. Features of a BN model that can be used for implementing a risk-based approach are discussed in Section IV. The final section summarizes conclusions and directions for future work.

II. GRAPHICAL MODELS FOR RISK ESTIMATION

A risk-based approach to achieving EM resilience for complex systems involves identifying and estimating the likelihood of possible EMI risk factors and assessing the severity of their undesirable consequences for the stakeholders and must deal with many uncertainties. Non-technical attributes that may be indicative of such risks include the provenance of the subsystem suppliers and past customer satisfaction, which influence confidence in the quality of the EMC design, or the target operational environment. These non-technical attributes can be assessed rapidly, even before detailed design and development are undertaken.

On the other hand, collecting the relevant data for most of the technical aspects (e.g. EM immunity test results, details of the internal and external EM environment of the system, spatial location of critical components within the system etc.) can be time consuming and expensive, as well as requiring domain expertise. Nevertheless, both technical and non-technical aspects should be considered together to obtain a more comprehensive understanding of the associated risks, which can be progressively refined and improved through the product development lifecycle.

In Fig. 1, a simple graph with five nodes provides an example to illustrate the possible combination of technical and non-technical aspects for estimating the likelihood of risks relating to EMI corresponding to a fictitious item. Each node $n \in \{1 \dots N\}$, where N is the total number of nodes in the graph, is associated with a node variable X_n , representing an attribute of the model. The possibility of malfunction of the item due to EMI is represented by node variable M , variable D represents the quality of the EMC design, S the provenance of the supplier, T the novelty of the technological solution, and V the outcome of EMC verification.

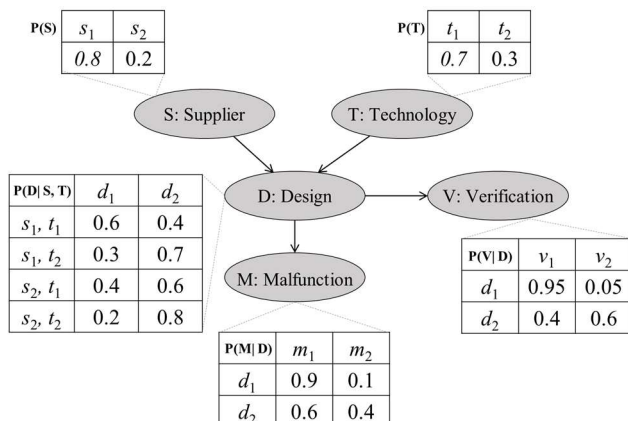


Fig. 1. A simple example of a graphical model with conditional probability distribution tables for the child nodes (D , M , V) and marginal distribution tables for the nodes with no parents (S , T).

Each of the node variables in Fig. 1 has an associated discrete state space x_n , of size $O(n)$ such that $x_n = \{x_1, x_2 \dots x_{O(n)}\}$. For example, the state space for node variable $M \in \{m_1, m_2\}$, allows for the item malfunction to be *unlikely* (m_1) or *likely* (m_2). The variable $D \in \{d_1, d_2\}$, representing the quality of the EMC design, may be either *good* (d_1) or *poor* (d_2). For simplicity, only two discrete states are considered for each of the node variables in Fig. 1, but larger state spaces could also be used. All of the node variables, their descriptions, and state spaces represented in Fig. 1 are detailed in Table I. The numerical probability values allocated to the state spaces in Fig. 1 are purely for the purposes of illustration and are not based on detailed analysis for any specific application.

The arrows in Fig. 1 are used to directly and qualitatively indicate the dependency relationships between the nodes. The dependencies between various attributes of the items could be either a causal relationship or a consequence effect. For example, the direction of the arrows emerging from nodes S and T , S towards node D is used to show a causal relationship, whereas the arrow from D to V represents a consequence effect. In this case, the use of a *established* technology can be a causal reason for anticipating a *good* EMC design and similarly, the consequent effect of having a *good* EMC design could be the achievement of a *pass* for immunity verification of the item.

III. BAYESIAN NETWORKS

A Bayesian network (BN) is a probabilistic graphical model, that is used extensively in fields like machine learning, artificial intelligence, medical diagnosis, computer vision etc. to provide a compact and natural representation of a set of probabilistic variables and their conditional dependencies.

A. Basic Terms

A BN is a *directed acyclic graph*, in which a set of nodes represent the random variables and a set of directed edges (represented as single-headed arrows) connect pairs of nodes to represent a *causal relationship* between those nodes.

The network is capable of encoding the probabilistic relationship between variables of a given model, such that the analyst is able to make an inference even in situations where data is missing or not known (a simple example of which is given in Section III.B). A BN approach can also be further used [6] for system risk management because it offers the following advantages:

TABLE I. ATTRIBUTES FOR NODE VARIABLES SHOWN IN FIG. 1.

Node Variables (X_n)	Description	State space (x_n)
Malfunction	likelihood of malfunction due to EMI	m_1 : unlikely, m_2 : likely
Design	EMC design quality	d_1 : good, d_2 : poor
Supplier	provenance of the item supplier	s_1 : reputable, s_2 : novice
Technology	novelty of the application or technology used to realize the function	t_1 : established t_2 : state-of-the-art
Verification	immunity verification test results	v_1 : pass, v_2 : fail

- provides ideal representation of combining prior knowledge and data;
- can be used to learn causal relationships, understand problem domain to predict the consequences of intervention;
- multiple states can be assigned to the state space of each node variable;
- enables extension of the model to influence diagrams [6], which can be used for decision making process with multiple choices;
- facilitates sensitivity analysis, where any consequent effect involving more than one causal factor can be assessed.

The arrow from D to M represents the influence of node variable D on node variable M . Hence, in this case D is the *parent node* and M is a *child node*. The dependency of node variable M on D is based on a reasonable prior belief, that is, an item with a *good* (d_1) EMC design is *unlikely* (m_1) to have malfunction due to EMI and similarly, for a *poor* (d_2) design quality a malfunction is *likely*. In many cases, such prior beliefs are not always completely true; hence every i^{th} state, x_i in the state space $x \in \{x_1, x_2, \dots, x_m\}$ associated with the random variable $X \in \{X_1, X_2, \dots, X_n\}$ is assigned with a prior conditional probability values.

B. Conditional Probability Distributions

In general, a *conditional probability distribution* (CPD) specifies a distribution over the states of a variable given each possible joint assignment with all of its parents' states. For nodes variables with no parent, it is simply a probability distribution that is conditioned on an empty set of the variables. The directed acyclic graphical network structure together with its CPDs constitutes a complete BN.

In Fig. 1, the CPD associated with each node variable is given adjacent to each node. In this case, the prior probability values given in the CPD tables are arbitrarily assigned for demonstration purpose. In practice, however, system engineers and domain experts would be able to assign the prior probabilities based on available statistical data and/or using data collected from specific simulations or experiments undertaken for that purpose. In other cases, when it is not possible to obtain prior probability values of a particular node variable, additional parent nodes that could act as causal factors/indicators could be used to assign the likelihood of the node variable.

In the example shown in Fig. 1, a poor EMC design will lead to higher probability of malfunction than a good design. However, in the absence of much information or expert knowledge it is not possible to determine the quality of the item EMC design. Hence, other non-technical attributes that could be available to the risk analyst at an early stage of the system, can be used as indicators of the likelihood of various states in state space of the variable D in Fig. 1. In this example the novelty of the technology (T) and the provenance of the item supplier (S) are proposed to fulfil this role. For an item judged to have a good (or poor) EMC design, the probability that the item will pass (or fail) the EMC test would then be anticipated as high (or low).

C. Model Structure

BN modelling for the likelihood estimation includes the right choice of entities/associated variables and structure of the graph. In cases where the structure of the BN is not known, sufficient data are required to learn the correlation between the node variables [8]. However, when there are no data immediately available, the model structure can be constructed in such a way that it reflects the causal order and dependencies (i.e. parent nodes reflect the causal factors and child nodes represent the consequent effects). The BN model shown in Fig. 1 is an example, where the structure was formed using a backward construction process [9], starting from the node M . Alternatively, one could also make use of hazard identification tools such as event trees (to determine consequences) and fault trees (for causes) [10].

IV. PRACTICAL APPLICATION ILLUSTRATIONS

Graphical models offer the advantage of improved visualization for complicated systems, and increased scope and granularity can be achieved by simply adding further relevant nodes and edges to the model. More importantly, PGMs can also be employed for estimating answers to probabilistic queries relating to the node variables of the model, as discussed in [2].

A. Bayesian Inference

In addition to the representation of causal orders, dependencies and prior beliefs, a BN can be further utilized for answering the probabilistic queries, as more evidence or observations become available to the analyst. This process is known as Bayesian inference. For example, a possible query Q_1 associated with Fig. 1 could be: "What is the probability of malfunction due to EMI, given evidence or observations that:

- 1) the item is implemented with established technology;
- 2) the item is supplied by a novice manufacturer;
- 3) the EMC immunity test result is a pass."

For Q_1 , the item design is considered as an unknown entity. To represent Q_1 formally, the conditional probability can be used:

$$P(Q_1) = P(m_2 | s_2, t_1, v_1) \quad (1)$$

The inference of Q_1 can be obtained by estimating the conditional probability [11], according to which (1) can be written as;

$$P(m_2 | s_2, t_1, v_1) = \frac{P(m_2, s_2, t_1, v_1)}{P(s_2, t_1, v_1)} \quad (2)$$

Further, using the *chain rule* [9], the numerator in (2) can be expanded as:

$$P(m_2, s_2, t_1, v_1) = \{P(d_1, m_2, s_2, t_1, v_1) + P(d_2, m_2, s_2, t_1, v_1)\} \quad (3)$$

Further expanding (3) yields:

$$P(m_2, s_2, t_1, v_1) = \{P(s_2) * P(t_1) * [P(m_2 | d_1) * P(v_1 | d_1) * P(d_1 | s_2, t_1) + P(m_2 | d_2) * P(v_1 | d_2) * P(d_2 | s_2, t_1)]\} \quad (4)$$

Using the numerical values detailed in Fig. 1, (4) can then be evaluated as:

$$P(m_2, s_2, t_1, v_1) = 0.01876 \quad (5)$$

Similarly, the denominator in (2) may be expanded as:

$$P(s_2, t_1, v_1) = \{P(d_1, m_1, s_2, t_1, v_1) + P(d_2, m_1, s_2, t_1, v_1) + P(d_1, m_2, s_2, t_1, v_1) + P(d_2, m_2, s_2, t_1, v_1)\} \quad (6)$$

Based on the numerical values detailed in Fig. 1, (7) can then be evaluated as:

$$P(s_2, t_1, v_1) = 0.0868 \quad (7)$$

Substituting the results of (5) and (7) into (2) finally yields the probability of the query Q_1 :

$$P(Q_1) = P(m_2 | s_2, t_1, v_1) = 0.22 \quad (8)$$

In a BN, probabilistic queries associated with any node variable can be calculated both deductively and inductively. For example, in Fig. 1 the provenance of the supplier (S) is a causal factor determining the quality of the item EMC design (D), whereas the likelihood of the verification results (V) being a pass or a fail is a consequent effect that depends on the quality of the EMC design.

Nevertheless, the probabilities for queries associated with both V and S (i.e. Q_{10} and Q_{14} of Table II) are calculated and listed in Table II, which also lists results for other sample queries for the model shown in Fig. 1. The formulation and calculation of queries for BN with many node variables becomes increasingly complex. However, this problem can be mitigated using the *conditional independence* properties of BN.

B. Conditional Independence and D-separation

A BN structure encodes the conditional independencies and the factorization of the distribution into local probability models. The independence properties asserted by the BN can

TABLE II. EXAMPLE BAYESIAN INFERENCE FOR RISK ANALYSIS USING PRIOR BELIEF

Node Variables (X_i)	Query Q_i given evidence	Probability $P(Q_i)$
Malfunction = m_2 (malfunction is likely)	$Q_1: (m_2 s_2, t_1, v_1)$	0.22
	$Q_2: (m_1 d_1, s_1, t_1, v_1)$	0.90
	$Q_3: (m_1 d_2, s_2, t_2, v_2)$	0.60
	$Q_4: (m_1 s_1, t_1, v_1)$	0.8342
	$Q_5: (m_1 s_2, t_2, v_2)$	0.6061
	$Q_6: (m_1 s_1)$	0.753
	$Q_7: (m_1 s_2)$	0.702
	$Q_8: (m_1 v_1)$	0.8049
	$Q_9: (m_1 v_2)$	0.6211
Verification = v_1 (EMC immunity test result is a fail)	$Q_{10}: (v_1 s_1, t_1)$	0.73
	$Q_{11}: (v_1 s_2, t_2)$	0.51
	$Q_{12}: (v_1 s_1)$	0.43
	$Q_{13}: (v_1 s_2)$	0.32
Supplier = s_1 (item is supplied by a reputable manufacturer)	$Q_{14}: (s_1 t_1, v_1)$	0.8248
	$Q_{15}: (s_1 t_1, v_2)$	0.7397

be used to reduce substantially the computation cost of the inference [6].

For example, the probability value of the possible query $Q_2: (m_1 | d_1, v_1, s_1, t_1)$ in Table II can be reduced to $P(m_1 | d_1)$. Given the state of variable D (i.e., $M \perp S, T, V | D$), the state taken by the variable M is independent of variables S, T , and V . Hence, $P(Q_2) = 0.9$ (the value is directly obtained from the CPD table given in Fig. 1).

The independence properties of distributions that factorize over the graph of a BN can be derived from the directed separation also known as *d-separation* [9]. For an intuitive understanding of the d-separation concept and the associated independence properties, the four possible trails between any two non-adjacent nodes of a BN are shown in Fig. 2 (which are derived from the example graph illustrated in Fig. 1).

In the *indirect causal* trail $S \rightarrow D \rightarrow M$ of Fig. 2(a), if the EMC design quality variable D is not known then the variable S corresponding to provenance of the supplier, influences the probability of malfunction M (see varying values of $P(Q_6)$ and $P(Q_7)$ in Table II, when D is not known), making the trail an *active trail*. However, if information on D is available, then variable M is no longer influenced by variable S , and the trail becomes an *inactive trail*. The local independence for the network in Fig. 2(a) is formalized as " $S \perp M | D$ ".

Although converse, the *indirect evidential* trail $M \leftarrow D \leftarrow S$ in Fig. 2(b) has a symmetrical notion of the independence properties to indirect causal trails, hence, $M \perp S | D$. Similarly, for the *common cause* case, $M \leftarrow D \rightarrow V$ in Fig. 2(c), if information on a variable associated with a parent node having multiple child nodes is not known, then the child nodes are independent of each other ($M \perp V | D$) and the trail is considered an active trail (inactive otherwise).

Unlike the cases discussed above, for structures of the *common effect* type (also called v-structures), the parent nodes of a child are independent only when the state of the child variable or any of its descendants is observed. Hence, in Fig. 2(d) the trail $S \rightarrow D \leftarrow T$ becomes active if, and only if, the state of variable D is observed (or inactive if unobserved).

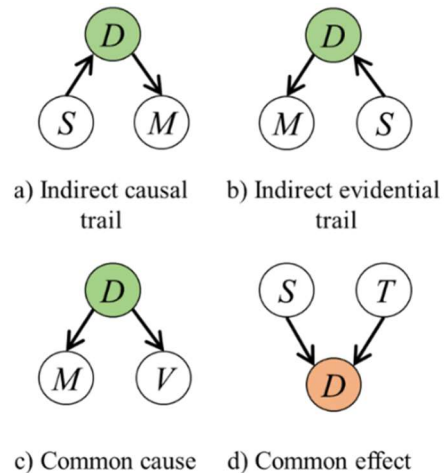


Fig. 2. Four possible two-edge inactive trails taken from Fig. 1 to explain the d-separation concept. States of green nodes are unknown and orange nodes are known, making each of the trails active.

Adjacent nodes with *direct cause/evidence trails* are never independent of each other. Interestingly, in the current rule-based EM approach, the EMC verification test results are considered as the *direct evidence* of the quality of the EMC design of the item. This direct evidence trail can be denoted as $V \leftarrow D$.

There are also readily available software tools, which can calculate the probability of such queries very quickly. One such example is MSBNx [12], a Windows application that can be used for creating, assessing, and evaluating BNs.

V. CONCLUSION

A risk-based EM threat management approach should be initiated during the concept phase of the system development, such that possible threats leading to undesirable consequences can be identified and their relative risk levels assessed from the outset of development. Methods such as BN can be used as an effective tool for estimating the likelihood of undesirable EM threats, such as EMI interaction with mission critical functions.

In the early stages of development, limitations such as lack of data and insufficient system knowledge can be handled using the properties of BN, such as causal mapping, utilizing domain expertise and prior knowledge. This paper has illustrated the use of non-technical information as possible proxies for detailed technical information at a stage where such data is not available. Furthermore, such models also provide the possibility to include more detailed data into the model as and when it becomes available. Thus, the EM risk model can be progressively refined throughout the development process.

Combining analysis of both technical and non-technical aspects of a complex system will not only provide increased confidence and understanding of risks at an early stage, but also facilitate broadening of the scope of the system EM risk analysis, thereby permitting the alignment of related system level attributes such as EMC and human exposure to EM fields, as well as functional safety and cybersecurity.

REFERENCES

- [1] D. Pissort and K. Armstrong. "Why is the IEEE Developing a Standard on Managing Risks Due to EM Disturbances?" Proceedings of the IEEE 2016 International Symposium on EMC, Ottawa, Canada, pp. 78 – 83, July 2016.
- [2] L. Devaraj, A.R. Ruddle and A.P. Duffy, "System-level risk analysis for immunity in automotive functional safety analyses", Proc. EMC Europe 2020, Rome, Italy, 2020.
- [3] BS EN IEC 31010:2019: "Risk management – Risk assessment techniques", British Standards Institute, Aug. 2019.
- [4] IEEE 1848:2020, "IEEE Approved Draft Techniques & Measures to Manage Functional Safety and Other Risks with Regard to Electromagnetic Disturbances", IEEE, 4th Jun. 2020.
- [5] L. Devaraj, A. R. Ruddle and A. P. Duffy, "Electromagnetic Risk Analysis for EMI Impact on Functional Safety with Probabilistic Graphical Models and Fuzzy Logic," IEEE Letters on Electromagnetic Compatibility Practice and Applications, vol. 2, no. 4, pp. 96–100, Dec. 2020.
- [6] D. Heckerman, "A tutorial on learning with Bayesian Networks", Technical Report, MS-TR-95-06, Microsoft Research, Redmond, WA 98052, 1995.
- [7] F. Taroni, A. Biedermann, S. Bozza, P. Garbolino, and C. Aitken, Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science, New York: John Wiley & Sons Inc., pp. 45–84 2006.
- [8] A. Freno and E. Trentin, Hybrid Random Fields: A Scalable Approach to Structure and Parameter Learning in Probabilistic Graphical Models, Intelligent Systems Reference Library Vol. 15, Springer-Verlag, Berlin, ISBN 978-3-642-20307-7, pp. 28–33, 2011.
- [9] D. Koller and N. Friedman, Probabilistic Graphical Models: Principles and Techniques, MIT Press, London; Cambridge, Massachusetts, 2009.
- [10] S. Rastayesh et al, "A system engineering approach using FMEA and Bayesian Network for Risk Analysis—A case study," Sustainability, vol. 12, no. 1, pp. 77, Dec. 2019.
- [11] R. E. Neapolitan, Learning Bayesian Networks, Prentice Hall; ISBN 9780123704771, pp. 27–29, 2003.
- [12] C. M. Kadie, D. Hovel and E. Horvitz, "MSBNx: a component-centric toolkit for modeling and inference with Bayesian Networks", Technical Report MSR-TR-2001-67 Microsoft Research, Redmond, WA 98052, 2001.