Safety Science 102 (2018) 211–225

Contents lists available at ScienceDirect

Safety Science

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

Addressing the epistemic uncertainty in maritime accidents modelling using Bayesian network with interval probabilities

Guizhen Zhang[a], Vinh V. Thai[b,*]

[a] DHI-NTU, Nanyang Environment and Water Research Institute, Interdisciplinary Graduate School, Nanyang Technological University, 1 Clean Tech Loop, Singapore 637141, Singapore

[b] School of Business IT & Logistics, RMIT University, Melbourne, VIC 3000, Australia

ARTICLE INFO

Keywords:
Interval probabilities
Bayesian network
Maritime accidents
Experts' elicitation
Epistemic uncertainty

ABSTRACT

Bayesian Network (BN) is often criticized for demanding a large number of crisp/exact/precise conditional probability numbers which, due to the lack of statistics, have to be obtained through experts' judgment. These exact probability numbers provided by the experts often carry a high level of epistemic uncertainty due to the incompleteness of human knowledge, not to mention the hardness in obtaining them in the first place. The existence of uncertainty in risk modelling was well recognized but seldom discussed. This paper explores the extension of BN with interval probabilities to the modelling of maritime accidents, which allows for the quantification of the epistemic uncertainty. Ship collision is chosen for case study for the strategic importance of navigational safety. The user friendly linguistic terms defined with interval scales were used for elicitation of interval conditional probabilities from industry experts. Inferences were made directly with the interval probabilities with the GL2U algorithm. Meanwhile, the interval probabilities were converted into point probabilities and computed with the traditional BN method for comparison, which were all shown to be within the ranges of the calculated posterior intervals probability. Results with inputs from different experts reveal discrepancies, which in turn verify the existence of uncertainty in risk modelling. A discussion was also provided on how the uncertainty in risk assessment propagates to the decision making process and influences the ranking of potential risk control options.

1. Introduction

1.1. Maritime accidents and BN

Maritime accidents have continued to occur, which threaten the safety of seafarers at sea, the economic performance of shipping companies and the environment. Therefore, understanding why and how accidents happen is of great importance for future safety management. Since accidents cannot be completely avoided, the reasonable goal is to control the accident risk to a desired level. Risk assessment is essential for this purpose. By performing risk analysis, we can evaluate the safety level of the current system as well as identifying the most critical issues. Some risk assessment methods also enable the evaluation of risk control options and thus ascertaining the most cost-effective way for reducing the risk level. Many risk analysis methods have been developed in the past few years, including Hazard and Operability Studies (HAZOP), Failure Mode and Effects Analysis (FMEA), Event Tree Analysis (ETA), Fault Tree Analysis (FTA) and Bayesian Belief Network (BN). Each of these risk analysis tools has its unique characteristics and fits different purposes.

BN is becoming an increasingly popular methodology for risk analysis of the maritime transportation system in recent years due to its capability to model causal interdependence, to incorporate of experts' knowledge when statistical data does not exist, to make dynamic updates when new observation is made, and to include human and organizational factors. Helle et al. (2011), Lehikoinen et al. (2015), Montewka et al. (2014), Banda et al. (2016) and Zhang et al. (2013) are a few examples of BN applications in the maritime risk analysis field. A more detailed review of the literature on maritime accidents risk prediction based on Bayesian Network can be referred to Goerlandt and Montewka (2015b) as well as Zhang and Thai (2016). BN was also

Abbreviations: AHP, Analytic Hierarchy Process; AIS, Automatic Identification System; ATSB, Australian Transport Safety Bureau; BN, Bayesian Network; BRM, Bridge Resource Management; DNV, Det Norske Veritas; ETA, Event Tree Analysis; FMEA, Failure Mode Effects Analysis; FSA, Formal Safety Assessment; FTA, Fault Tree Analysis; GCAF, Gross Cost of Averting a Fatality; GL2U, Generalized Loopy 2-Updating; HAZOP, Hazard and Operability Analysis; HFACS, Human Factors Analysis and Classification System; IMO,, International Maritime Organization; IPE, Iterated Partial Evaluation; IPT, Interval Probability Theory; MAIB, Marine Accident Investigation Branch; NCAF, Net Cost of Averting a Fatality; OOW, Officer on Watch; RCO, Risk Control Options; SLP, Support Logic Programming; STRAITREP, The Mandatory Ship Reporting System in the Straits of Malacca and Singapore; SV2U, Structured Variational 2U Method; TSB, Transportation Safety Board of Canada; VTS, Vessel Traffic Service

* Corresponding author.
E-mail addresses: GZHANG004@e.ntu.edu.sg (G. Zhang), vinh.thai@rmit.edu.au (V.V. Thai).

http://dx.doi.org/10.1016/j.ssci.2017.10.016
Received 29 September 2016; Received in revised form 24 August 2017; Accepted 18 October 2017

Table 1
Comparison of aleatory uncertainty and epistemic uncertainty.

Comparison	Aleatory uncertainty	Epistemic uncertainty
Source of the uncertainty	Variability of the underlying stochastic process	Incomplete knowledge of the system
Existence	Quantities	Model and parameter
Associated with	System	Analyst
Reducible	Not reducible	Reducible
Modelling by	Probability number/distribution	Alternative probability number/distribution
Equivalent terms	Variability, natural uncertainty, objective uncertainty, inherent variability, (basic) randomness, and type-A uncertainty	Subjective uncertainty, lack-of-knowledge or limited-knowledge uncertainty, ignorance, specification error, prediction error, and type-B uncertainty

recommended for risk assessment (Step 3 of the Formal Safety Assessment, FSA) to International Maritime Organization (IMO) (IMO, 2006). The application of BN includes three steps, i.e. BN structure development, parameterization, and inferences. Both the BN structure and parameters could be built manually, automatically or a combination of both (Neil et al., 2000; Kjrćulff and Madsen, 2013; Neil et al., 2000).

1.2. Uncertainty and BN modelling

The consideration of uncertainties is crucial for obtaining reliable results in risk analysis (Merz and Thieken, 2005). By sources, uncertainties could be broadly separated into two class: aleatory and epistemic uncertainty, the comparison of which could be found in Table 1.

Liu et al. (2003) reviewed some of the most important uncertainty reasoning approaches, including the Bayesian theory of probability, Dempster-Shafer theory of evidence, and fuzzy set theory. Each of these approaches views and handles uncertainties from different perspectives. Bayesian Theory has many good features such as strong theoretical root, less computational complexity compared with other approaches. It models aleatory uncertainty through probability but could not include epistemic uncertainty since each entity must be assigned with exact probability numbers.

Due to the lack of available statistical data, experts' opinion is an important source for the probability specification or parameterization in BN modelling, especially for applications to the maritime risk assessments. This, however, poses huge challenges for the reliability of the model as well as the involved domain experts. First, for probability elicitation, the experts are often asked about the conditional dependence between the model elements on top of their own expertise. Moreover, the requirement to elicit a large number of probability numbers adds to the workload of the experts. From the viewpoint of modelling, the involvement of experts will lead to epistemic uncertainty, due to the lack of knowledge about the system (Liu et al., 2003; Merrick et al., 2005; Fallet et al., 2011), which can sometimes be referred to as quantities which have fix values, but their exact value are unknown (Swiler et al., 2009).

The lack of systematic consideration of uncertainty in the applications of maritime transportation risk analysis was identified through a detailed review in Goerlandt and Montewka (2015b), even though the existence of uncertainties are recognized and accepted. One exception was Merrick et al. (2005) which used Bayesian approach to estimate the impacts of parameter uncertainties in the traffic simulation model (not the risk analysis model) for evaluating the ferry expansion alternatives. In the last few years, there have been more studies with focus on uncertainties in the maritime risk models. For example, Sormunen et al. (2014) showed through an extensive study that the uncertainties in accident and risk models can be significant. Goerlandt and Montewka (2015a) went one step further by introducing a framework where uncertainties are qualitatively assessed. However, so far, there are no studies on maritime risk analysis which quantitatively address the epistemic uncertainties. The objective of this paper is to provide a method to model the epistemic uncertainties related with the probability parameters in Bayesian Network models for maritime risk analysis.

To achieve the objective, this paper seeks to extend BN by including interval probabilities. Interval probability expresses imprecision in a more straightforward way. Fallet et al. (2011) concluded that the interval probability method best represents the experts' knowledge due to more appropriate semantics as compared to hard evidence, soft evidence, and total ignorance. In applications, obtaining interval probability parameters are much easier than getting point probabilities, especially when there is only little, incomplete or conflicting information available to assist experts' judgment (Guo and Tanaka, 2010). In other cases, when multiple experts are involved, each expert may indicate their own belief and if no consensus could be reached, the result are interval probabilities (Cozman, 2000). Considering these facts, the application of interval probability (with an upper bound and lower bound) in BN could bring more application value.

1.3. Organization of this paper

The rest of this paper is organized as follows. Section 2 defines interval probability, discusses its properties and summarizes the updating algorithm for BN with interval probabilities. Section 3 presents the application of BN with interval probabilities to ship collision causation probability modelling. The detailed elicitation process is also discussed in this section. Section 4 shows the inference result with the interval probabilities. An example of the influence on the evaluation of risk control options with interval probabilities is provided as well. Finally, Section 5 summarizes the paper.

2. Methodology

This paper extends the traditional BN to include interval probability parameters for maritime risk modelling. Inferences are made directly with interval parameters. The following subsections present the definition of credal network, interval probability and the relative properties.

2.1. Credal network and BN with interval probability parameters

BN with interval probabilities is a special type of credal network, which extends BN to deal with imprecision and uncertainty (Corani et al., 2012). A credal network over a set of random variables $\mathbf{X} = (X_1, \dots, X_k)$ is $\langle G, \{\mathbf{P_1}, \dots, \mathbf{P_m}\}\rangle$, where G is a directed acyclic graph whose nodes have one-to-one correspondence to the elements in $\mathbf{X}$, and $\langle G, \mathbf{P}_j\rangle$ is a BN over $\mathbf{X}$ for each $j = 1, \dots, m$ (Antonucci, 2008; Antonucci and Zaffalon, 2008). This definition indicates that a credal network could be regarded as a set of BNs, as illustrated in Fig. 1.

Fig. 1a is an example of the traditional BN where all the probabilities are exact numbers. Fig. 1b is a credal network with the same structure. The probabilities in the credal network are imprecise, being an interval or comparisons of the probabilities, which enables the representation of classificatory and comparative probability judgements (Piatti et al., 2010). For example, for node A, the judgement $1 < P(a)/P(\neg a) < 3$ means that the chance of $A = a$ is one to three times higher than the chance $A = \neg a$. For node B, the probability of $P(b|a)$ could be any value between 0.2 and 0.3. The BN in Fig. 1a is just one among many others that satisfy the probability conditions of the credal network in Fig. 1b. The inference with a credal network is the same with inferences with its vertices (Antonucci, 2008). However, the number of vertices is exponential to the input size except for the case of binary