Expert Systems with Applications 41 (2014) 1632–1646

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Bayesian network modeling of Port State Control inspection findings and ship accident involvement

Maria Hänninen [a,*], Pentti Kujala [b]

[a] Aalto University, School of Engineering, Department of Applied Mechanics, Marine Technology, Kotka Maritime Research Centre, Heikinkatu 7, FI-48100 KOTKA, Finland
[b] Aalto University, School of Engineering, Department of Applied Mechanics, Marine Technology, P.O. Box 15300, FI-00076 AALTO, Finland

ARTICLE INFO

Keywords:
Maritime traffic accidents
Bayesian networks
Port State Control
Hidden variables
Maritime traffic safety

ABSTRACT

The paper utilizes Port State Control inspection data for discovering interactions between the numbers of various types of deficiencies found on ships and between the deficiencies and ship's involvement in maritime traffic accidents and incidents. Bayesian network models for describing the dependencies of the inspection results, ship age, type, flag, accident involvement, and incidents reported by the Vessel Traffic Service are learned from the Finnish Port State Control data from 2009–2011, 2004–2010 Baltic Sea accident statistics and the reported Gulf of Finland Vessel Traffic Service incidents within 2004–2008. Two alternative Bayesian network algorithms are applied to the model construction. Further, additional models including a hidden variable which represents the complete system and its safety features and which links the accident and incident involvement and Port State Control findings are presented. Based on model-data fit comparisons and 10-fold cross-validation, a constraint-based learning algorithm NPC mainly outperforms the score-based algorithm repeated hill-climbing with random restarts. For the highest scoring models, mutual information and influence of evidence analyses are conducted in order to analyze which network variables and variable states are the most influential ones for determining the accident involvement. The analyses suggest that knowledge on ship type, the Port State Control inspection type and the number of structural conditions related deficiencies are among the ones providing the most information regarding accident involvement and the true state of the hidden system safety variable.

1. Introduction

Maritime traffic accidents have fortunately been rather rare events. For example, approximately 0.7 accidents per 1000 port ship calls have taken place in the Gulf of Finland, the densely trafficked but shallow Baltic Sea area between Finland, Russia and Estonia (Kujala, Hänninen, Arola, & Ylitalo, 2009; Kuronen, Lehikoinen, & Tapaninen, 2009). However, if occurring, a maritime traffic accident could have adverse outcomes resulting in the loss of lives or a severe damage to the environment (Helle, Lecklin, Jolma, & Kuikka, 2011).

The occurrence of accidents in a system can be seen as an indicator of the system safety, or rather the absence of safety. However, past accident frequency alone is not a very informative indicator on how accidents occur or what makes a system safe. Maritime traffic system, or even a single ship, is a complex socio-technical system. Since Heinrich (1931) presented the Domino theory, various frameworks and models of how and why accidents occur in such a system have been proposed, e.g. Perrow (1984), Reason (1990), Rasmussen (1997), Hollnagel (1998), Dekker (2002), Leveson (2004) and Hollnagel (2009); for a summary, see Qureshi (2007). The recent models approach the problem from a systemic viewpoint and state that the accidents are a result of variability in the system, a property which is inevitable, and even desired, as the variability also prevents many accidents (Hollnagel, 2004). Furthermore, according to the systemic accident models it is impossible and unadvisable to describe the cause-effect relationships of accidents within a complex system (Hollnagel, 2009; Dekker, 2011).

Instead of attempting to develop a causal model of a maritime accident occurrence, modeling any potential patterns between different system safety indicators, given that the indicators truly measure safety, could provide indirect albeit potentially useful information for the purposes of safety management. Further, as such a model contains factors describing measurable features, its validation might be less challenging. Still, one should consider the inevitable epistemic and aleatory uncertainty due to multiple reasons such as the chosen modeling approach itself, the necessary modeling assumptions and simplifications, incomplete knowledge of the system components and their dependencies and due to the system complexity and variation between the shipping companies, vessels, crew properties, and the dynamic circumstances, for instance. Also, some relevant factor might be unobservable and thus

---

* Corresponding author. Tel.: +358 50 410 6518.
E-mail address: maria.hanninen@aalto.fi (M. Hänninen).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.08.060

Table 1
The Bayesian network variables to be learned from the dataset.

Variable	Origin	Type
Alarms deficiencies	PSC data	0/1 or more
Cargo operations/equipment deficiencies	PSC data	0/1 or more
Crew certificates deficiencies	PSC data	0/1 or more
Dangerous goods deficiencies	PSC data	0/1 or more
Documents deficiencies	PSC data	0/1 or more
Emergency systems deficiencies	PSC data	0/1 or more
Fire safety deficiencies	PSC data	0/1 or more
the ISM code (The International Safety Management) deficiencies	PSC data	0/1 or more
Life saving appliances deficiencies	PSC data	0/1 or more
Living conditions deficiencies	PSC data	0/1 or more
Pollution prevention deficiencies	PSC data	0/1 or more
Propulsion and auxiliary machinery deficiencies	PSC data	0/1 or more
Radio communications deficiencies	PSC data	0/1 or more
Safety of navigation deficiencies	PSC data	0/1 or more
Ship certificates deficiencies	PSC data	0/1 or more
Structural conditions deficiencies	PSC data	0/1 or more
Water/weathertight conditions deficiencies	PSC data	0/1 or more
Working conditions deficiencies	PSC data	0/1 or more
Other deficiencies	PSC data	0/1 or more
Inspection type	PSC data	categorical
Ship type	PSC data	categorical
Age	PSC data	categorical (interval)
Flag	PSC data	categorical
Detention	PSC data	yes/no
Accident involvement	HELCOM data	0/1 or more accidents
VTS reported	VTS reports	0/1 or more incidents

no data is available on that. An attractive quantitative modeling technique that can present relatively complex, not necessarily causal dependencies and cope with at least some of these uncertainties and unobservable variables while also having a qualitative, graphical dimension, is Bayesian belief networks (Pearl, 1988).

Safety indicators are measurable features which can be utilized in describing the safety state of a system (e.g. Øien, Utne, & Herrera (2011)). If data is available, machine learning techniques could be employed to discover the dependencies between the relevant safety indicators – either with a purely data-driven perspective or by incorporating prior information using Bayesian techniques. Previous studies have applied machine learning to discovering various maritime safety-related information from data (Le Blanc & Rucks, 1996; Le Blanc, Hashemi, & Rucks, 2001; Grech, Horberry, & Smith, 2002; Tzannatos & Kokotos, 2009; Hashemi, Le Blanc, Rucks, & Shearry, 1995; Kokotos & Smirlis, 2005; Kelangath, Das, Quigley, & Hirdaris, 2011; Antao, Guedes Suares, Grande, & Trucco, 2008; Mascaro, Nicholso, & Korb, 2013). However, although some initial work has been published on utilizing machine learning in maritime traffic accident modeling (Kristiansen, 2010), their applications to probabilistically describe the dependencies of various factors and maritime accidents without a causal interpretation remain absent.

For maritime traffic safety controlling purposes, one set of potential safety indicators results from Port State Control inspections. Port State Control (PSC) inspection, conducted by the port state authority when foreign ships visit their port, checks the condition, the equipment, manning and the operation of foreign state vessels for verifying that the aforementioned aspects on board comply with international regulations (IMO, 2011). If the inspection reveals deficiencies that pose a safety hazard, the ship may be detained at the port until the deficiencies have been rectified. This paper continues the work conducted by the authors in Hänninen and Kujala (submitted for publication) and explores how Port State Control inspection findings on different areas of inspection are linked to each other and to maritime traffic accidents and incidents using Bayesian belief networks. Previously published studies utilizing PSC data (Mejia Jr, Cariou, & Wolff, 2010; Knapp & Franses, 2007; Li, Yin, Yang, & Wang, 2010; Soma, 2004) have not studied or modeled these connections in detail. Further, the paper introduces Bayesian network models where "accident causation mechanisms" or "system variability" is presented with one immeasurable so-called hidden variable which is linked to the safety indicators, i.e., the PSC deficiencies, incident and accident involvement. The resulting Bayesian network models can be used as an aid in safety management, for example when a certain number of PSC deficiencies of one – or multiple – type(s) have been observed and one wants to examine what this observation tells about the other deficiencies, how the proneness to accidents changes given various deficiency observations, or how the deficiencies and other model variables differ between ships which have been involved in accidents or incidents and the ones which have not.

The rest of the paper is organized as follows. The data and the methods used in the study are described in Section 2. Section 3

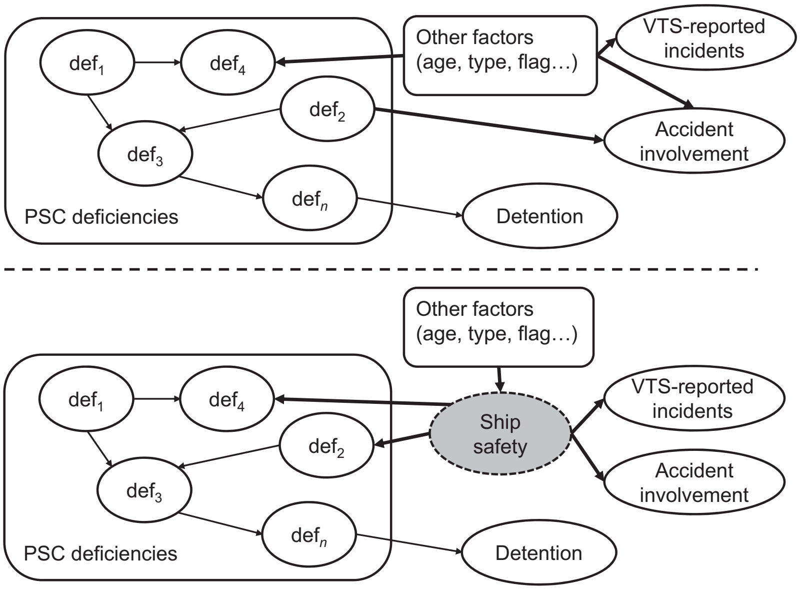

Fig. 1. A hypothetical simplified Bayesian network model learned from the PSC, VTS and accident data and the principle of adding the hidden *safety state* variable.