Expert Systems with Applications 40 (2013) 4273–4282

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Decision support analysis for safety control in complex project environments based on Bayesian Networks

Limao Zhang [a,b], Xianguo Wu [a,*], Lieyun Ding [a], Miroslaw J. Skibniewski [b], Y. Yan [a]

[a] School of Civil Engineering & Mechanics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China
[b] Department of Civil & Environmental Engineering, University of Maryland, College Park, MD 20742-3021, USA

ARTICLE INFO

Keywords:
Decision support analysis
Safety control
Bayesian Networks (BNs)
Ground settlement
Complex environments

ABSTRACT

This paper presents a novel and systemic decision support model based on Bayesian Networks (BN) for safety control in dynamic complex project environments, which should go through the following three sections. At first, priori expert knowledge is integrated with training data in model design, aiming to improve the adaptability and practicability of model outcome. Then two indicators, *Model Bias* and *Model Accuracy*, are proposed to assess the effectiveness of BN in model validation, ensuring the model predictions are not significantly different from the actual observations. Finally we extend the safety control process to the entire life cycle of risk-prone events in model application, rather than restricted to pre-accident control, but during-construction continuous and post-accident control are included. Adapting its reasoning features, including forward reasoning, importance analysis and background reasoning, decision makers are provided with systematic and effective support for safety control in the overall work process. A frequent safety problem, ground settlement during Wuhan Changjiang Metro Shield Tunnel Construction (WCMSTC), is taken as a case study. Results demonstrate the feasibility of BN model, as well as its application potential. The proposed model can be used by practitioners in the industry as a decision support tool to increase the likelihood of a successful project in complex environments.

1. Introduction

In the past 10 years, metro construction has presented the powerful momentum for rapid development all over the world. Owing to various risk factors of uncertainty in complex environments, safety violations occur frequently in metro construction. On January 12, 2007, Pinheiros Station on Metro Line Four collapsed at Sao Paulo's Aquarium in Brazil, causing death of seven people (Ti, 2007). On July 6, 2010, a tunnel boundary accident also took place in Prague, Czech Republic, causing a 15-meter-wide sunken pit (Ti, 2010). On August 23, 2012, metro line leak caused chaos in Warsaw. Water flooded into the tunnel at the planned Powisle station. Fortunately nobody hurt, but traffic had been redirected, causing considerable transport problems in the already gridlocked city (Zagranicy, 2012). Also in China, the number of construction accidents shows a big rise during large-scale metro construction projects. On July 1, 2003, great quantities of sand swarmed into the tunnel in Shanghai Track Traffic Line Four. The tunnel was severely damaged and discarded as useless, causing a total direct economic loss exceeding 1000 million Yuan (Huang Tj, 2003). On November 15, 2008, 21 people were killed as a result of a road cave-in above a metro tunnel under construction in Hangzhou (Yb, 2008). As for Shenzhen Metro Construction, four accidents happened with five deaths in October 2009 (Song, 2009). Also, three people were killed and five were injured during collapse accidents from April to May in 2011. As you can see, metro construction is a complicated project with high risks, which would bring enormous hidden dangers for the public safety.

To avoid heavy casualties and property losses caused by the safety violations, innumerable studies have introduced risk-based analysis into safety control. Risk analysis can be divided into qualitative and quantitative risk analysis. The former includes fault tree analysis (FTA), comprehensive fuzzy evaluation method and check list; while the latter includes job risk analysis method (LEC) method, influence diagrams (de Klerk, 2001), neural network (Carr & Tah, 2001), support vector machines (Wang, Yuan, Chen, Yang, Ni, & Li, 2012) and decision trees (Vens, Struyf, Schietgat, Džeroski, & Blockeel, 2008). These risk-based analysis methods make a significant contribution to the safety control in complex engineering projects (del Caño & de la Cruz, 2002; Piniella & Fern A Ndez-Engo, 2009; Vinod, Bidhar, Kushwaha, Verma, & Srividya, 2003), but they are confined to static control management (Alaeddini & Dogan, 2011). Khakzad described FTA unsuitable for complex problems for its limitation in explicitly representing the dependencies of events, updating probabilities, and coping with uncertainties (Khakzad, Khan, & Amyotte, 2011). Yang et al. regarded LEC

* Corresponding author. Tel.: +1 301 221 8777.
E-mail address: wxg0220@126.com (X. Wu).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.11.022

unsuitable for complex dynamic environment, resulting from its insufficiency in timely diagnosing and dealing with various problems (Yang, Liu, Huang, & Wu, 2011). When associated parameters, such as geological, design and construction parameters are changed, these methods cannot accurately illustrate the updated features of dynamic environments as the construction progress evolves. Nor can the professional supports or suggestions be provided in real time as these parameters are updated.

Safety control is a dynamic management process in complex engineering environments, where the relative parameters would truly change along with the change in space and time. With the capacity of integrating prior knowledge and sample data, Bayesian Networks (BN) provide a powerful tool for knowledge representation and reasoning under the dynamic environments. Compared to neural networks, support vector machines, decision trees, and so forth, BN have shown superior for several high-level classification tasks such as datamining fault monitoring and bioinformatics (Hsu, 2004; Xu, 2012). At present, BN have been widely used in quality evaluation (Correa, Bielza, & Pamies Teixeira, 2009), dynamic management (Barrientos & Vargas, 1998), knowledge discovery (Lee & Abbott, 2003), as well as decision support (Lauria & Duchessi, 2006). As a consequence, BN is introduced as the resolving method for safety control in dynamic and complex engineering environments.

In recent years, relatively few researchers have adopted BN method to risk management in large engineering projects (Ordonez, 2007), mainly focusing on the application. Lee, Park, and Shin (2009) presented a large engineering project risk management procedure using BN and then identified the major risk items that affected project performance. Bayraktar and Hastak (2009) applied BN in the decision support system for safety maintenance in developing suitable contracting strategies among different project components in highway work zone projects. Sousa and Einstein (2012) presented a construction strategy decision model based on BN to systematically access and manage the risks associated with tunnel construction. However, there are primarily two problems in the current study related to BN application: (1) Excessive attention is given to the establishment of BN model, but the validation of an established BN model is rarely completed. As a matter of fact, the effectiveness of BN model is considered as an essential guarantee of the plausibility in its application; (2) BN models are mostly applied before an accident happens, also referred to as pre-accident control (Ledolter & Swersey, 2005; Li, Xu, Wang, & Song, 2009). Few studies have explored during-construction continuous control and post-accident control, which are two indispensable links for safety control in the overall process. Applying an accurate BN model that has already been validated to decision support analysis for safety control in complex project environments and provide professional technical support for decision makers in the overall process constitutes an entirely new domain. We first propose a decision support model for safety control based on BN, then make detailed expatiation of the design and validation process of the BN model. At last, adapting its reasoning features, including forward reasoning, importance analysis and background reasoning, decision makers could be provided with scientifically documented and effective support in regard to safety control. A typical metro construction hazard ground settlement such as the one in Wuhan Changjiang Metro Shield Tunnel Construction (WCMSTC) is taken up as a case study. Results demonstrate the feasibility of proposed method, as well as its application potential.

The structure of this paper is organized as follows. The fundamental theory and analyzing method of BN are introduced in Section 2, including model design, validation and application. Section 3 presents a BN model for safety control in shield tunnel construction, namely Ground Settlement Bayesian Network (GSBN). In Section 4, the proposed BN model, GSBN, is validated by two indicators: *Model Bias* and *Model Accuracy*. In Section 5, GSBN is applied for decision support analysis on safety control in the overall work process. Afterward, the conclusions are drawn in Section 6.

2. Research method

2.1. Model design

Bayesian Networks (BN), a combination of two different mathematical areas, graph theory and probability theory, consist of a directed acyclic graph (DAG) and an associated joint probability distribution (JPD) (Zhu & Deshmukh, 2003). A BN model with n nodes can be represented as $B\langle G, \Theta\rangle$, where G stands for a DAG with n nodes, and Θ stands for the JPD of the BN model. The nodes $\{X_1,\ldots,X_n\}$ in the graph are labeled by related random variables. The directed edges between nodes present the association relationship within the variables. DAG contains conditional independence assumptions. The relations represented by DAG allow the JPD to be specified locally by the conditional probability distribution of each node. Assuming $\pi(X_i)$ is the parent nodes of X_i in the DAG, the conditional probability distribution of X_i is denoted by $P(X_i|\pi(X_i))$. When X_i has no parents, the corresponding distribution is simply $P(X_i)$. The JPD of $P(X_1,\ldots,X_n)$ can then be written as Eq. (1).

$$P(X_1,\ldots,X_n) = \prod_{X_i\in\{X_1,\ldots,X_n\}} P(X_i|\pi(X_i)) \tag{1}$$

The design of a BN model involves two parts, structure learning and parameter learning. Structure learning aims to figure out the proper DAG, confirming the association relationship between nodes. Parameter learning aims to determine the conditional probability distribution of each node under the established BN structure (Cooper & Herskovits, 1992; Oni S Ko, Druzdzel, & Wasyluk, 2001). Normally, there are three approaches for BN design: (1) Depending entirely on expert knowledge for both structure and parameter learning; (2) Depending entirely on training data for both structure and parameter learning; (3) BN structure is mainly designed through expert prior knowledge, and training data is only used for parameter learning.

Due to the limitation and uncertainty of human interventions in knowledge acquisition, the outcome of the BN model designed by the first approach may have discrepancy with the real status. Bias discrepancies may exist between the predicted results and actual observations (Cano, Masegosa, & Moral, 2011). The second approach is a data-driven way, and the predicted results may have strong adaptability when simulating actual conditions. Nevertheless, this approach has demanding requirements on both quantity and quality of the training data. The biggest problem is that the BN structure obtained after data-driven parameter learning is difficult to understand (Khakzad et al., 2011). The third approach is a compromise, which can increase the learning efficiency when the relationship among variables is obvious. In fact, enormous amounts of historical information about safety control in the engineering is accumulated, such as standard specification, technical manuals and fault trees. The historical information can be served as effective prior knowledge for structure learning of BN model. On that basis, relatively limited training data is gathered from project sites as to meet the requirements in parameter learning. By this way, we can increase the adaptability of model outcome, and reduce the subjectivity of BN design at the same time.

2.2. Model validation

The conditional probability distribution of each node is developed independently in model design (Wang, Robertson, & Haines, 2009). Therefore, it is necessary to test and validate the established