Physica A 467 (2017) 226–239

Contents lists available at ScienceDirect

Physica A

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

A hierarchical network modeling method for railway tunnels safety assessment

Jin Zhou [a], Weixiang Xu [a,*], Xin Guo [a,b], Xumin Liu [c]

[a] School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, PR China
[b] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, PR China
[c] College of Information Engineering, Capital Normal University, Beijing 100048, PR China

HIGHLIGHTS

- A network modeling method is proposed to evaluate tunnels safety situation.
- Apriori algorithm is improved to rapidly and efficiently mine frequent patterns.
- Approaches are proposed to represent frequent patterns in networks.
- The method can reveal key defects and critical potential risks of defects.

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 1 August 2016
Available online 12 October 2016

Keywords:
Network theory
Railway tunnels safety assessment
Data mining
Apriori algorithm
Fuzzy set theory

ABSTRACT

Using network theory to model risk-related knowledge on accidents is regarded as potential very helpful in risk management. A large amount of defects detection data for railway tunnels is collected in autumn every year in China. It is extremely important to discover the regularities knowledge in database. In this paper, based on network theories and by using data mining techniques, a new method is proposed for mining risk-related regularities to support risk management in railway tunnel projects. A hierarchical network (HN) model which takes into account the tunnel structures, tunnel defects, potential failures and accidents is established. An improved Apriori algorithm is designed to rapidly and effectively mine correlations between tunnel structures and tunnel defects. Then an algorithm is presented in order to mine the risk-related regularities table (RRT) from the frequent patterns. At last, a safety assessment method is proposed by consideration of actual defects and possible risks of defects gained from the RRT. This method cannot only generate the quantitative risk results but also reveal the key defects and critical risks of defects. This paper is further development on accident causation network modeling methods which can provide guidance for specific maintenance measure.

1. Introduction

In China, railway tunnels safety situations get more and more attention in railway management. Due to a lot of railway tunnels suffer serious or minor defects, it may lead to unsafe circumstance of train running even could result in accidents [1]. In order to prevent such accidents, management took great efforts to collect defects information of tunnels every year.

* Corresponding author.
E-mail address: wxxu317@126.com (W. Xu).

http://dx.doi.org/10.1016/j.physa.2016.10.026
0378-4371/

They want to know the safety situations of tunnels and potential risks of these tunnels, which can support them to make appropriate decisions to prevent accidents.

In order to reduce the risk of accidents in complex systems, innumerable researches have introduced various safety assessment methods into safety management [2,3]. One of the most suitable methods to model complex systems is systemic accident causation models. Various types of systemic accident models have been introduced by recent studies for qualitative analysis, such as Systems-Theoretic Accident Model and Process (STAMP) [4], AcciMap technology [5], Human Factors Analysis and Classification System (HFACS) [6], etc. This kind of models considers the performance of the system as a whole, which makes the analysis more comprehensive and reasonable. Usually, these models divide the whole system into several subsystems or several hierarchical levels [7,8]. It can provide a clear understanding of the procedure of risks evaluation.

However, the systemic accident causation models can only obtain qualitative analysis but not quantitative results [9]. In order to overcome this limitation, a lot of researchers introduced network theory which is a powerful tool to model complex systems [9–15]. Ma et al. analyzed the causation of a railway accident and established a network to describe the complex interactions in complex systems [12]. Zhou et al. proposed a railway faults spreading model based on dynamic of complex network in order to simulate the dynamic process of failures interactions in networks [13]. Zhou et al. established a subway construction accident network (SCAN) and analyzed the complexity of SCAN [14]. Zhou et al. proposed a method for modeling and analysis of directed weighted accident causation network (DWACN) [15]. No matter which kind of modeling methods, the key issues of network establishment can be summarized as [15]: (1) how to identify the causal factors which are denoted by the nodes in networks; and (2) how to identify the relationships among factors which are represented by the edges. The causal factors are usually elicited from the items in the expert experience, database or accident reports. Then according to the expert investigations, valuable items or factors are selected as the nodes. Whether there are relationships among these causal factors can be identified via expert investigations, event chains and accident chains.

Most of the past studies did not take into consideration of the weights of edges. However, Zhou et al. pointed out that the weights of edges have an influence on the analysis results [15]. In order to quantitative describe the relations among items in database, data mining techniques provide useful tools to mine regularities and knowledge in database [16]. Data mining is the process of selection, exploration, and modeling of large quantities of data to discover regularities or relations that are at first unknown with the aim of obtaining clear and useful results for the owner of the database [17]. Among various types of data mining techniques, association rules are the most basic tools to discover the frequent patterns, associations and correlations [18]. Frequent patterns are aimed at quantitatively demonstrating relations according to various combinations of the set of discrete variables. Most of the studies by using association rules are aimed at mining the frequent patterns [19–21]. It still needs further study on how to translate the frequent patterns into networks to support knowledge representation. Thus, in this paper, the complex relationships among nodes in network are represented by risk-related regularities table (RRT). Lots of defects detection data for railway tunnels are collected in autumn every year in China. Whether associations exist among these data? If so, how can we utilize the underlying information to assist tunnels safety assessment? Hence, focusing on these issues, this paper will propose a hierarchical network modeling method for risk management in railway tunnels projects.

This paper is organized as follows. The methodologies which are applied to our research are described in Section 2. Section 3 shows the data, data analysis and processing. The methods of establishing the network, generating the RRT and evaluating the safety situations are proposed in Section 4. The relevant results and analysis are given in Section 5. At last, conclusions are summarized in Section 6.

2. Methodology

2.1. Network theory

Network theory provides a powerful tool to analyze complex systems. A network consists of nodes which denote components and edges which represent the complex interrelationships among these nodes in a system [9]. A directed weighted network with N nodes can be represented mathematically by an $N \times N$ adjacency matrix $\boldsymbol{AM}$ with elements [22].

$$AM_{cm} = \begin{cases} b_{cm} \cdot x_{cm} & \text{if node } c \text{ points to node } m, \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where b_{cm} takes the value 1 if node c points to node m and 0 otherwise. And x_{cm} specifies the weight on the edge if node c points to node m ($x_{cm} = 0$ otherwise).

The node strength of s_c of a node c is the weight of neighbor nodes incident with the node. In directed networks, the strength of the node has two components: the number of outgoing weight of node c: $s_c^{out} = \sum_m x_{cm}$, and the number of incoming weight of node c: $s_c^{in} = \sum_m x_{mc}$. The total weight is then defined as $s_c = s_c^{out} + s_c^{in}$.

2.2. Association rule

Association rule is one of the most fundamental parts in data mining methodologies. Han et al. [16] provide the description of association rules as the following contents. Let $Z = \{Q_1, Q_2, \ldots, Q_m\}$ be an itemset. Let O be a set of database