Article

Not peer-reviewed version

Research on Safety Risk Transfer in Subway Construction Based on Text Mining and Complex Networks

Kunpeng Wu , Jianshe Zhang , Yanlong Huang * , Hui Wang , Hujun Li , Huihua Chen

Posted Date: 18 September 2023

doi: 10.20944/preprints202309.1106.v1

Keywords: Text mining; apriori algorithm; complex network model; subway construction; risk transfer

Article

Research on Safety Risk Transfer in Subway Construction Based on Text Mining and Complex Networks

Kunpeng Wu [1],Jianshe Zhang [2], Yanlong Huang [2,*],Hui Wang [2], Hujun Li [2] and Huihua Chen [3]

1 Zhengzhou Metro Group Co. Ltd, Zhengzhou 450000, China; 17537101157@163.com (K.W.).
2 Henan Polytechnic University, Jiaozuo 454000, China; jianshezhang@hpu.edu.cn(J.Z.); 212108020025@home.hpu.edu.cn (Y. H.); wanghui9962@hpu.edu.cn (H.W.); lihujunem@hpu.edu.cn (H.L.).
3 Central South University,Changsha 410083, China; chh@csu.edu.cn (H.C.).
* Correspondence: 212108020025@home.hpu.edu.cn

Abstract: Subway construction is often in a complex natural and human-machine operating environment, and that complicated setting leads to subway construction more prone to safety accidents, which can cause substantial casualties and monetary losses. Thus, it is necessary to investigate the safety risks of subway construction. The existing literature on the identification and assessment of subway construction safety risks(SCSR) is susceptible to the influence of subjective factors. Moreover, although existing studies have explored the interrelationships between different risks, these studies usually analyze the interrelationships of single risks, lack the study of risk chain transfer relationships, and fail to find out the key path of risk transfer. Therefore, this paper innovatively combines text mining, association rules and complex networks to deep mine subway construction safety incident reports and explore risk transfer process. Firstly, it uses text mining technology to identify subway construction safety risk; Then, association rules are introduced to explore the causal relationships among safety risk; Finally, the key safety risk and important transfer paths of subway construction safety accidents (SCSA) are obtained based on the complex network model. Research results show that (a) improper safety management, unimplemented safety subject responsibilities, violation of operation rules, non-perfect safety responsibilities system and insufficient safety education and training are the key safety risk in SCSA; (b) two shorter key risk transfer paths in the subway construction safety network can be obtained: insufficient safety education and training→lower safety awareness→violation of operation rules→safety accidents; insufficient safety checks or hidden trouble investigations→violation of operation rules→safety accidents; (c) in the process of risk transfer, the risk can be controlled by controlling the key nodes or cutting off the transfer path. The results of the study provide new ideas and methods for SCSR identification and influence element mining, which help safety managers propose accurate subway construction safety risk control measures.

Keywords: text mining; apriori algorithm; complex network model; subway construction; risk transfer

1. Introduction

The Chinese government issued the 14th Five-year Development Plan for Modern Comprehensive Transportation System to promote the construction of highways, railways, transportation systems, etc. [1]. Aligned to the implementation of the new infrastructure construction policy, China has accelerated the development of the rail transportation industry. According to the statistical report released by the Ministry of Transportation and Communications in March 2023, a total of 292 urban rail transit lines have been opened and operated in 54 cities, with an operating mileage of 9652.6 kilometers [2], indicating that the subway and rail construction is in a rapid development stage. However, due to the complex geological environment, high technical requirements and many uncertain factors, SCSA are prone to occur. For example, the Hangzhou subway Line 1 collapsed during construction, resulting in 21 deaths, 24 injuries and a direct economic loss of 49.61 million RMB [3–5]. The Guangzhou subway Line 11 collapsed, resulting in 3 deaths and direct economic loss of 20.047 million RMB [6]. Therefore, it is an urgent problem to investigate the safety risks of subway construction and examine the relationship between the safety risks accurately in order to improve the level of accident prevention and control.

Large scale of previous scholars have carried out investigations on the SCSR, and these studies mainly focus on safety risk identification, safety risk relationships analysis and safety risk assessment. As for the identification of safety risks, most of the researchers followed the framework of "personnel-equipment-material-method-environment". Such as, Yan et al. [7] categorized the safety risk of subway construction into personnel-type, machine-type, material-type, method-type and environment-type risks. As for the safety risk relationships analysis, previous literature has used various approaches to analyze the associations between identified safety risks, including system dynamics models (SD) [8,9], structural equations models (SEM) [10], decision-making trial and evaluation laboratory (DEMATEL) [11,12], Bayesian network modeling[13–15], etc. For example, Wu et al. [16] established an evaluation model to analyze the correlation between safety risks during subway station construction. In the area of safety risk assessment, coupling theory[17], cloud models[18], deep learning[19], optimization algorithms[20] , etc. have been widely used. For example, Pan et al.[17] established a measure for safety risk coupling of shield tunnel construction based on coupling degree theory; Feng et al. [20]constructed a safety assessment model by using hybrid particle swarm optimization neural network. However, the existing literature on the identification and assessment of SCSR is susceptible to the influence of subjective factors, and the results are highly influenced by human influence. Moreover, although existing studies have explored the interrelationships between different risks, these studies usually analyze the interrelationships of single risks, lack the study of risk chain transfer relationships, and fail to find out the key path of risk transfer, resulting in a lack of targeted accident prevention and control measures. How to clarify the interactions between SCSR and analyze the risk transfer mechanism has become an urgent scientific issue for current subway construction projects.

In order to solve the above problems, this paper first uses text mining to identify SCSR based on construction safety accidents from 2005 to 2023. Then, association rules are introduced to examine the causal relationships among safety risks. Finally, by using complex network theory, the nodes significance in the safety risk network (SRN) are measured by degree centrality, closeness centrality, and betweenness centrality, and the overall feature attribute of the risk network is evaluated by selecting the clustering coefficient, average path length, and network density to find out the key safety risk and the critical safety transfer paths of subway construction accidents. So that managers can take effective measures on key safety risks and important transfer paths to reduce safety risks during subway construction.

2. Literature review

2.1. Safety Risk Identification of Subway Construction

Most scholars identify SCSR from different perspectives and the research presents different viewpoints. Wang et al. [21] identified 42 safety risks related to personnel risks based on literature research, questionnaire survey and expert interview; Pan et al. [22] pointed out the key safety risks affecting the safety of subway shield tunnel construction are construction technology, engineering materials, construction equipment, personnel, and support capabilities; Yu et al.[23] identified safety risks in Chinese subway construction, including safety attitude, construction site safety and government regulation. In addition, Zhang et al. [24] classified subway safety risks into human, machine, management, material ,and environmental risks based on accident causation theory and literature review; Qie and Yan [25] categorized the identified five types of SCSR, including human-type safety risk, equipment-type safety risk, environmental-type safety risk, management-type safety risk, and safety culture-type, and concluded that the human-type risks are the key risks causing high-risk accidents; Fang et al. [26] classified the safety risk of subway tunnel construction into personnel, machinery, materials, management, and environment based on the N-K model and concluded that the human risk and the management risk are the important causes of accidents.

 doi:10.20944/preprints202309.1106.v1

2.2. Safety Risk Assessment of Subway Construction

Some scholars use fuzzy set theory and quantitative analysis to assess SCSR. For example, On the basis of fuzzy set theory and the fuzzy comprehensive evaluation method, Wu et al. [27]introduced the analytic network process (ANP) to construct a comprehensive risk assessment model for subways. Luo et al. [28] constructed a comprehensive risk assessment model for construction safety of prefabricated subway stations by using the structure entropy weight method, matter element theory and evidential reasoning. In addition, Liu et al. [3] combined qualitative analysis with quantitative analysis, and used set pair analysis (SPA) method to evaluate the construction safety of subway tunnels. With the booming development of machine learning and artificial intelligence, some researchers have applied machine learning techniques to safety risk assessment, such as neural networks [29,30], random forests [31,32], Bayesian networks [33,34], support vector machines [35,36], etc. Zhang et al. [37] proposed a method for assessing the safety of tunnels based on based on case-based reasoning, advanced geological prediction, and rough set theory. Wen et al. [38] established a fuzzy Bayesian network-based model for analyzing the risk of tunnel water breakout, and He et al. [39] used Bayesian networks for risk assessment of deformation in large tunnels.

2.3. Safety Risk Relationship Analysis of Subway Construction

In recent years, scholars have been keen to study the relationships between safety risks, which are mainly categorized into causal and coupling relationships research. In terms of causality research, Jiang et al. [40] combined system dynamics (SD), error back propagation (BP) neural networks, and mean influence value (MIV) algorithm to examine the causality and influence function among the shield construction safety risks. Eybpoosh et al. [41] used structural equation modeling (SEM) to determine causal relationships between different safety risks. Zhou et al. [42] developed a SCSRN network that combines causality with a variety of accidents at subway construction sites. In terms of coupling relationships. Yan et al. [43] explained the coupling relationship between the risk factors affecting the construction of subway stations by constructing an interaction matrix. Hou et al. [44] applied the N-K model to get the key risk coupling ways affecting the vulnerability of the system and concluded that the vulnerability of the subway construction safety system is greater when the personnel factors, the management factors, and the environmental factors are fully coupled. Based on the data collected and analyzed from the questionnaire survey, Liu et al. [45] identified a total of 24 critical safety factors for subway construction and used explanatory structural modeling to determine their interrelationships.

As an emerging science, complex networks provide a wealth of theories and methods for identifying key risks and analyzing the complex relationships among them [46]. Currently, complex network theory has been applied to accident analysis and used for safety risk analysis in natural disasters [47], construction [48,49], mining [50], and railroads [51–54]. For example, Chen et al. [55] used network theory to explore the risk characteristics of hybrid bridge and tunnel construction.

Based on the above literature review, it can be found that the current research on SCSR has the following shortcomings: Scholars have conducted more research on SCSR, but because they usually analyze the interrelationships of single risks, it is difficult to reflect the occurrence and development of SCSA. In addition, the previous identification of safety risks was mostly through literature analysis and expert discussion, the identification of safety risks has a lot of subjectivity. Therefore, this paper uses text mining to identify SCSR and employs association rules and complex network modeling to explore the subway construction risk transfer relationship.

3. Research methodology

3.1. Text Mining

The use of modern information technology to collect, mine, and analyze data can reduce the probability of safety accidents [41]. And the data mining technology can in-deep analyze the intrinsic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 September 2023 doi:10.20944/preprints202309.1106.v1

correlation and value of the accident data [56]. Text mining technology, as a method of data mining, has become a research hotspot, and currently this technology has been applied to biomedical [57,58], agricultural [59], educational [60], and engineering [61] fields. The analysis of accident reports using text mining techniques has been widely used in the study of accident causes as it can significantly improve the accuracy of accident predictions[62]. For example, Xu et al.[63] used textual features and text mining methods to predict the cause of accidents; Li et al. [64] utilized text mining, association rule mining and Bayesian networks to mine the textual data of coal mine safety accident cases.

3.2. Association Rules

Association rules were initially proposed by Agrawal et al. [65] in 1993, association rules have become the most widely used methods in data mining. Association rules do not define dependent and independent variables, they reflect the relationship between two factors. Currently, the Apriori algorithm is the most classical frequent itemset generation algorithm, which finds frequent itemsets by calculating *support* and *confidence* to find association rules [66]. *Support*, *confidence* and *lift* are important parameters of association rules. *Support* reflects the importance of an association rule and indicates the frequency of an itemset in all data sets, and both *confidence* and *lift* are used to measure the correlation between itemsets and the reliability of an association rule [67].

(1) *Support*

Support indicates the probability of the simultaneous occurrence of itemset I_1 and itemset I_2 in all datasets and can be expressed by equation (1). If the *support* of the itemsets I_1 and I_2 is low, the rule occurs less frequently and is not generalized.

$$support(I_1 \Rightarrow I_2) = P(I_1 \cap I_2) \tag{1}$$

(2) *Confidence*

The probability of occurrence of itemset I_2 in all data sets where itemset I_1 occurs can be expressed by equation (2). If the *confidence* level is higher,, the more likely it is that itemset I_2 occurs when itemset I_1 exists.

$$confidence(I_1 \Rightarrow I_2) = \frac{P(I_1 \cap I_2)}{P(I_1)} = \frac{support(I_1 \Rightarrow I_2)}{support(I_1)} \tag{2}$$

(3) *Lift*

The lifting reflects the correlation between itemset I_1 and itemset I_2. When the *lift* is greater than 1, the stronger the positive correlation, when the *lift* is less than 1, the stronger the negative correlation, and when the *lift* is equal to 1, the itemset I_1 and itemset I_2 are not correlated, and the *lift* can be expressed by formula (3).

$$lift(I_1 \Rightarrow I_2) = \frac{P(I_2 / I_1)}{P(I_1)} = \frac{confidence(I_1 \Rightarrow I_2)}{support(I_2)} \tag{3}$$

A "good" association rule should have high *support* and *confidence*. Therefore, the *support* of the itemset I_1 should be greater than the *minimum support*, i.e. $support(I_1) \geq min-support$; When the association rule $I_1 \Rightarrow I_2$ satisfies *minimum support* and *minimum confidence*, it is said that $I_1 \Rightarrow I_2$ is a strong association rule, i.e., itemset I_1 is strongly associated with itemset I_2 when $support(I_1) \geq min-support$ and $confidence(I_1 \Rightarrow I_2) \geq min-confidence$.

3.3. Complex Networks

Any complex system containing a large number of units (or subsystems) can be examined as a complex network when its constituent units are expressed by nodes and interactions between units are expressed by edges [68]. The steps for determining risk network relationships based on accident reports are as follows:

Step1: The nodes of a risk network contain accident types and risk points. It is assumed that o accident types and t risk points are extracted from incident reports, the set of network nodes

 doi:10.20944/preprints202309.1106.v1

$$S=\left\{S_1,\ldots,S_i,\ldots,S_o,S_{o+1},\ldots,S_{o+j},\ldots,S_{o+t}\right\} \tag{4}$$

Step2: The node-to-node relationships constitute the edges of the risk network, and if node *i* has an effect on node *j*, it forms an edge as in Figure 1.

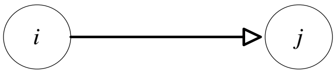

Figure 1. An edge from *i* to *j* in a risk network.

Step3: According to the causal relationship of association rule mining, if node *i* has influence on node *j*, then A_{ij}= 1; if node *i* has no influence on itself, then A_{ij} = 0. This can be expressed by the following equation:

$$A_{ij}=\begin{cases}1, & i\Rightarrow j\\ 0, & i\not\Rightarrow j \quad or \quad i=j\end{cases} \tag{5}$$

Step4: The adjacency matrix of the risk network $A=\left(A_{ij}\right)_{(m+n)\times(m+n)}$ is constructed based on the relationship between the nodes. Finally, the SCSR network model is established based on the adjacency matrix.

The network topology indicators can be calculated as follows.

(1) Network density

A higher network density indicates that the nodes are more connected to each other and can be calculated by the following equation:

$$D_i=\frac{\alpha}{\beta(\beta-1)} \tag{6}$$

Where α is the number of relationships present in the risk network, β is the number of network nodes, and $\beta(\beta\text{-}1)$ is the possible maximum value of the number of relationships.

(2) Clustering coefficient

The clustering coefficient of a node is positively correlated with the degree of connection between nodes and surrounding nodes. And the clustering coefficient can be calculated by using equation (7).

$$C_{\mathrm{i}}=\frac{2E_i}{k_i(k_i-1)} \tag{7}$$

In the formula (7), k_i denotes that the network node *i* has k_i edges connected to other nodes and E_i is the quantity of edges that exist with node *i*.

(3) Average path length

The shorter the average path of the network, the fewer intermediate nodes there are for information or energy to travel from one node to another. The average path length can be calculated as shown in equation (8).

$$L=\frac{1}{\frac{1}{2}\beta(\beta-1)}\sum_{i\geq 1}d_{ij} \tag{8}$$

Where d_{ij} is the quantity of edges on the shortest path between any two nodes *i* and *j* in the network and β is the quantity of network nodes.

(4) Degree

Degree is the basic indicator for the importance evaluation of the network nodes, and the degree value of a node is directly proportional to its importance. The degree value of a node consists of in-degree and out-degree, in-degree is the number of relations pointing to the node, and out-degree is the quantity of relations pointing from the node. A higher out-degree value indicates that the node

 doi:10.20944/preprints202309.1106.v1

influences other nodes to a higher degree, and a higher in-degree value indicates that the node is susceptible to the influence of other nodes. The degree calculation formula can be shown in equation (9).

$$C_{RDi} = \frac{X_{output} + X_{input}}{2(n-1)} \tag{9}$$

In the formula, X_{output} is the out-degree value of node X; X_{input} is the in-degree value of node X; n is the number of network nodes.

(5) Closeness centrality

In a risk network, closeness centrality reflects the distance from one risk node to other risk nodes, the larger the value of closeness centrality of a node means the closer it is to other nodes. Closeness centrality includes incloseness centrality and outcloseness centrality, and the calculation formulas can be shown in equations (10) and (11).

$$C_i^{in} = \frac{n-1}{\sum_{j=1}^{n-1} d_{ji}} \tag{10}$$

$$C_i^{out} = \frac{n'-1}{\sum_{j=1}^{n'-1} d_{ij}} \tag{11}$$

In the formula, d_{ij} is the shortest distance between node i and j; d_{ji} is the shortest distance between node j and i; n is the quantity of nodes i that can be reached; and n' is the quantity of nodes that node i can reach.

(6) Betweenness centrality

The greater the intermediate centrality, the greater the ability of the node to bridge other nodes. The formula for betweenness centrality is shown in equations (12) and (13).

$$C_{ABi} = \sum_{j}^{\beta}\sum_{k}^{\beta} b_{jk}(i) \tag{12}$$

$$b_{jk}(i) = \frac{g_{jk}(i)}{g_{jk}} \tag{13}$$

In the formula, $g_{jk}(i)$ is the quantity of nodes i on the shortest path between points j and k; g_{jk} is the quantity of shortest paths between node j and node k ; and β is the quantity of nodes.

4. Study on safety risk transfer in subway construction

4.1. Data Sources

Accident analysis can identify the causes of accidents and the frequency, probability of accidents, as well as the path of accidents[69], and the analysis of the causes and path of accidents can be targeted to the construction of safety management systems and safety training [70]. Safety accident investigation reports can be used for accident analysis, as they contain detailed descriptions on the accident time, accident process, accident direct or indirect causes, and accident prevention measures. In this paper, 101 subway construction safety accident reports were collected from the China's Ministry of Emergency Management and local emergency management bureaus from 2000 to 2023, including 40 cases of collapse accidents, 11 cases of high fall accidents, 9 cases of object attack accidents, 9 cases of vehicular injury accidents, 7 cases of mechanical injury accidents, 5 cases of explosions, 4 cases of electrocution accidents, and 16 other cases of other accidents (drilling through tunnels, shield machine flooding, fire, poisoning, etc.). Statistics on the classification of accident reports are shown in Figure 2.

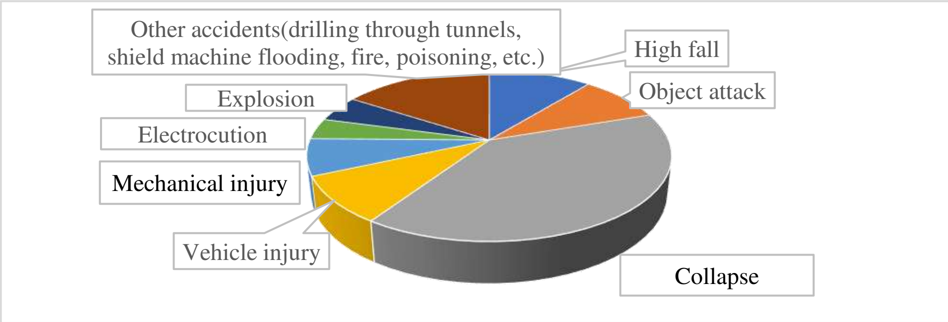

Figure 2. Statistics on the classification of subway construction accident reports.

4.2. Identification of Safety Risks in Subway Construction Based on Text Mining

Firstly, the direct and indirect causes of accidents in accident investigation reports are selected as text mining objects, and the text is formatted and numbered to construct the corpus. Secondly, Python was chosen as the program language for text mining, and the Chinese word segmentation module of Jieba is used to segment the corpus. In text mining, attention should be paid to details and precision, especially the recognition and merging of professional terms. In this paper, a self-defined subway construction safety professional thesaurus and stop word list are constructed, and the Jieba module is utilized to load and update the thesaurus, so as to avoid the problem that the word segmentation module cannot identify the professional vocabulary. Among them, the stop word list selects the "Harbin Institute of Technology stop word list" and self-defined stop words.

Synonyms summarization is also an effective way to improve the accuracy of text mining. We merged synonyms to avoid words with the same meaning being misjudged as different words. For example, "weak safety awareness" and "lack of awareness" were merged into "lower safety awareness"; "safety training and education" and "staff training" were merged into "insufficient safety education and training ". The less frequent accident types were merged into the "other accident" type. Finally, 29 safety risks and 8 accident types are obtained, as shown in Table 1.

Table 1. Safety Risk and Accident Types in Subway Construction.

Types	Number	Name
Accident types	A1	High fall
	A2	Object attack
	A3	Collapse
	A4	Vehicle injury
	A5	Mechanical injury
	A6	Electrocution
	A7	Explosion
	A8	Other accidents
Hunan risks	H1	Violation of operation rules
	H2	Lower safety awareness
	H3	Workers' operation error
	H4	Workers' lower capacity
	H5	Inadequate risk perception
Material risks	EM1	Equipment failure
	EM2	Non-standard construction materials
	EM3	Abnormal driving parameters
	EM4	Inadequate inspection and maintenance of mechanical equipment
	EM5	Materials damage
Environmental risks	E1	Complex geological conditions
	E2	Complex surrounding pipelines

 doi:10.20944/preprints202309.1106.v1

	E3	Abundant groundwater
	E4	Climatic conditions
	E5	Heavy traffic over the tunnel
	E6	Poor construction environment
Technical risks	T1	Improper construction methods
	T2	Misalignment between design and construction
	T3	Design defects
	T4	Insufficient advance support
	T5	Insufficient geological exploration
Management risks	M1	Insufficient safety checks or hidden trouble investigations
	M2	Improper safety management
	M3	Unreasonable personnel arrangement and division of labor
	M4	Insufficient technical disclosure or construction scheme
	M5	Insufficient safety measures
	M6	Insufficient safety education and training
	M7	Unimplemented safety subject responsibilities
	M8	Non-perfect safety responsibilities system

4.3. Causality Mining for Subway Construction Based on Apriori Algorithm

The safety risks in each safety investigation report were treated as an itemset. According to the Apriori algorithm, if the threshold is set too low, a large number of irrelevant association rules will be generated, which will affect the calculation results. If the threshold is set too high, the data will have high reliability, but some useful association rules will be omitted. For the collected data, the number of association rules generated by the application of data mining at different parameter thresholds was statistically analyzed. Therefore, in order to improve the reliability and credibility of the data, the minimum *support* was set to 6% and the *lift* was set to 1.0. A total of 1258 strong association rules were mined. According to formulas (1)–(3), some association rules are shown in Table 2.

Table 2. Association rules of safety accidents in subway construction.

Antecedent	consequent	*Support*(%)	*Confifence*(%)	*Lift*
M4	T1	7.92	61.54	2.83
M6	H2	28.71	80.56	1.43
H1	A1	10.89	77.78	1.42
M8	H1	10.89	91.67	1.30
M6	H5	17.82	69.23	1.23
M5	H1	16.83	85.00	1.21
H1	T1	17.82	81.80	1.16
H1	A8	12.87	81.69	1.16
H2	H1	28.71	80.56	1.15
H2、M1	A1	6.9	61.11	3.57
H1、H2	A1	10.89	62.07	3.48
E1、M7	A3	8.91	81.82	2.07
M4、H1	T1	6.9	77.78	3.57
M1、M4	T1	6.9	70.00	3.21
M5、M6	H1、H2	7.9	80.00	2.79

4.4. Construction and Analysis of SRN in Subway Construction

Complex network is an abstract modeling method based on graph theory for complex systems containing a large number of elements. Generally, the elements are regarded as nodes of the network, and the connections between the elements are regarded as edges. According to the formula (4), the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 September 2023 doi:10.20944/preprints202309.1106.v1

29 safety risks and 8 accident types in subway construction are selected as the nodes of risk network. The lift can reflect the strength of the correlation between different risks. For example, if the lift between I_1 and I_2 is 2, then when A occurs, the probability that B occurs doubles. Therefore, in all strong association rules, the lift between the antecedent item and the result item is used as the weight of the edge. In this paper, the visualization software Ucinet 6.6 is used to construct the SRN topology diagram of subway construction, as presented in Figure 3.

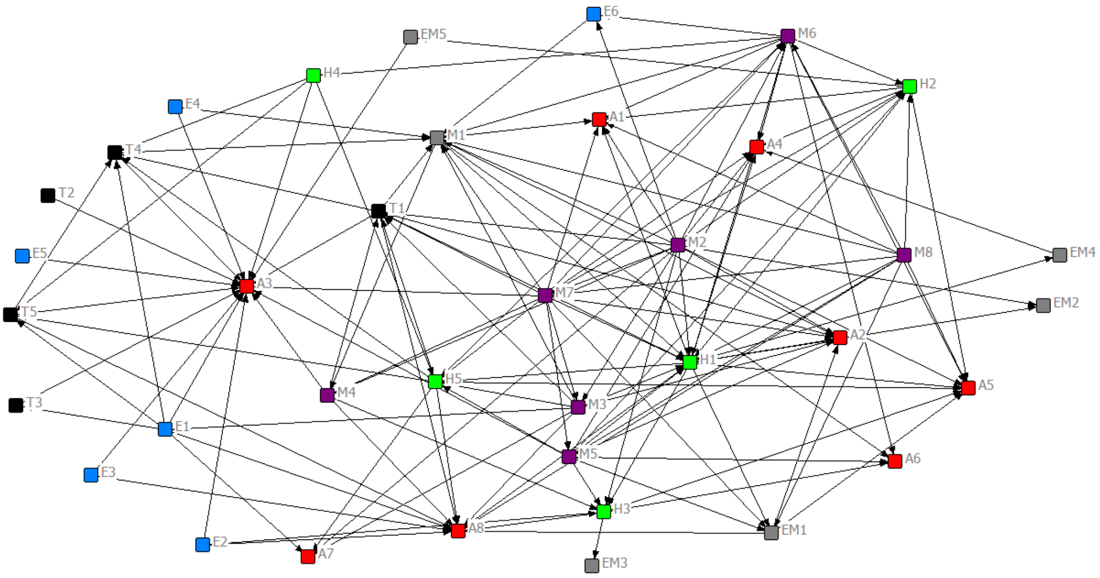

Figure 3. Topological diagram of SRN in subway construction.

4.4.1. Risk network overall feature attribute analysis

The calculation of overall feature attribute indicators is shown in Table 3. The network density is 0.207, indicating that the network is relatively compact. Safety risk transfer mainly depends on key nodes, which can be blocked by controlling key nodes. The average path length is 2.083. The shorter path length indicates that the safety risk transfer faster. In the SRN of subway construction, no more than 3 nodes may lead to safety accidents, such as node M6 (Insufficient safety education and training) and node H1 (Violation of operation rules), which can form a complete accident chain. The diameter of the SCSR network is 4, which means that there are 2 nodes between the two maximum distance nodes in the network, and the safety risk transfer speed is fast. The clustering coefficient is 0.280. Therefore, most nodes in the subway construction accident network are not directly connected but connected through key nodes, which indicates that key nodes play an important role in risk transfer. Managers should cut off the links between nodes to effectively avoid accidents.

Table 3. Overall Feature Attribute Indicators.

Network density	0.207	Average path length	2.083
Network diameter	4.000	Clustering coefficient	0.280

4.4.2. Risk network node analysis

(1) Degree

Based on the calculated degree values, the degrees of the top 20 nodes are plotted into a distribution as shown in Figure 4. Among the risk nodes, the out-degree values of nodes M2 (Improper safety management), M7 (Unimplemented safety subject responsibilities), H1 (Violation of operation rules), M8 (Non-perfect safety responsibilities system) and M6 (Insufficient safety education and training) are relatively large, which indicates that these node risks are important risks in affecting the other node risks. The out-degree value of node M2 is even more than 16, which indicates that the node M2 may result in the emergence of 16 safety risks, and therefore should be

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 September 2023 doi:10.20944/preprints202309.1106.v1

strictly controlled the occurrence of M2. On the contrary, nodes M1 (Insufficient safety checks or hidden trouble investigations), H5 (Inadequate risk perception) and T1 (Improper construction methods) have a high in-degree, indicating that these nodes are more susceptible to the influence of other nodes.

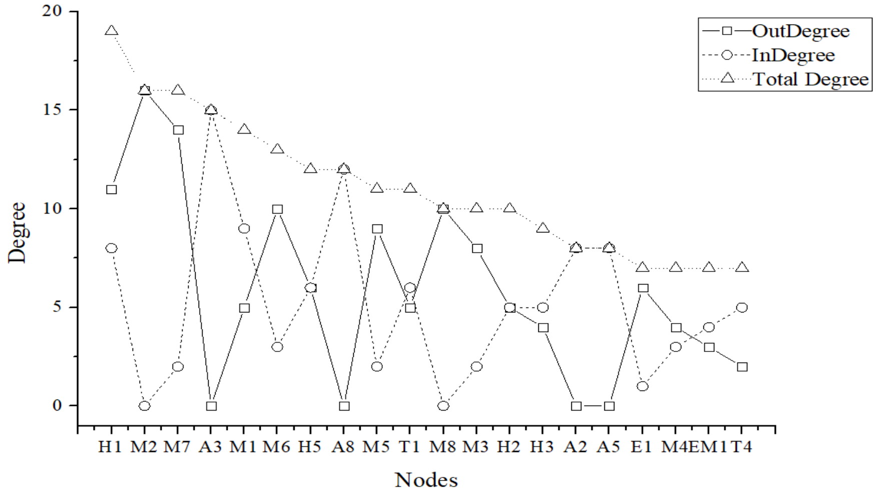

Figure 4. In-degree, out-degree and total degree of the top 20 nodes in total degree value.

In addition, node H1 has higher out-degree and in-degree values, indicating that H1 is easy to lead to the occurrence of other risks and easy to be affected by other risks. And H1 is the most complex risk in the high incidence of SCSA and transfer paths. The node A3 (collapse) has the largest in-degree value, which indicates that accidents in node A3 are most likely to occur under the effect of many safety risk. The degree value of the average node in the network is 3.73, indicating any safety risks in networks averagely connect 4 other safety risks. This means that a change in the state of one safety risk may change the state of more than 4 other safety risks during the transfer of a subway construction accident.

(2) Closeness centrality

In the risk network, the closeness centrality reflects the degree of closeness between nodes, the greater the closeness centrality of a node means the closer it is to other nodes and the more important that node is in the network. The calculation results of incloseness centrality and outcloseness centrality are shown in Figure 5. From the perspective of incloseness centrality, A3 (collapse) and A8 (other accidents), have the greatest consequences and are more likely to occur. In addition, among the safety risk nodes, EM3 (abnormal driving parameters), H3 (workers' operation error), EM1 (equipment failure), and EM2 (non-standard construction materials) have relatively large incloseness centrality, which indicates that these nodes are more closely related to and susceptible to the influence of other safety risks. Outcloseness centrality denotes the influence degree of a node has on other nodes. Safety risk nodes such as M2, M8, M7, and M6 have large outcloseness centrality, indicating that they are important risks affecting other nodes. The management and supervision of these important nodes need to be strengthened so as to improve the stability of the network.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 September 2023 doi:10.20944/preprints202309.1106.v1

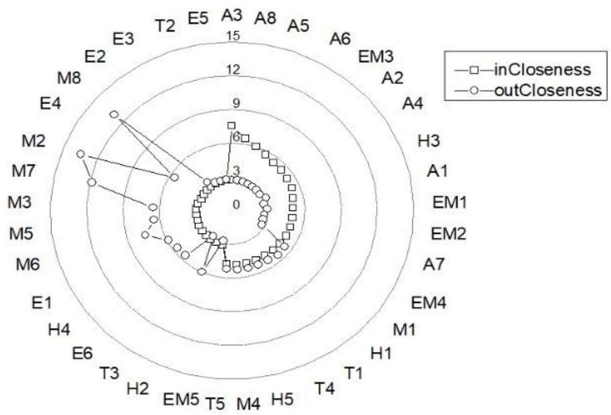

Figure 5. Incloseness centrality and Outcloseness centrality Values.

(3) Betweenness centrality

Some nodes are excluded because they have betweenness centrality value of 0, which means they do not act as a "bridge" in the safety risk transferring. The betweenness centrality of a node reflects its ability to transmit information, and the nodes are sorted according to the calculation results shown in Table 4. Among them, M1 (insufficient safety checks or hidden trouble investigations) has the largest betweenness centrality, indicating that this node is the most important bridge in transferring risk. If the safety inspection or hidden danger investigation is not carefully conducted, onsite managers' and workers' non-standard behaviors may not be timely detected, and thus safety accident may be quickly formed based on this path.

Table 4. Betweenness centrality Values.

Rank	Safety risk node	Betweenness centrality	Rank	Safety risk node	Betweenness centrality
1	M1	135.200	2	H1	126.602
3	T4	49.000	4	H5	37.821
5	T1	31.986	6	H3	25.983
7	M4	18.367	8	M7	8.286
9	M6	8.119	10	M3	7.650
11	H2	6.100	12	E1	5.500
13	M5	4.026	14	H4	2.510
15	EM1	1.017	16	EM5	1.000
17	T5	0.833	-	-	-

These nodes with greater betweenness centrality can connect other nodes that seem less intrinsically connected more closely. They establish risk pathways between nodes that are less connected, causing safety risks interactions and risk propagation more efficient. For example, the H1 (violation of operation rules) node plays an important connecting role between its antecedent risks, for instance, M6 (insufficient safety education and training) and the result risk H3 (workers' operation error). When insufficient safety education and training, violations of operation rules, and workers' operation errors co-occur at the same time, the probability of safety accidents is greatly increased. Thus, strengthening the safety education and training and the safety checks or hidden trouble investigations can effectively cut off the risk transfer path and reduce the possibility of accidents.

Based on the above study's overall feature attributes, such as network density, average path length, network diameter, clustering coefficients, etc., it can be found that in the SCSR network, the risk transfer relies on the key risk nodes, and the speed of the safety risks transfer is faster. The analysis of node degree, closeness centrality, and betweenness centrality can be used to obtain the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 September 2023 doi:10.20944/preprints202309.1106.v1

key safety risk and shorter critical risk transfer paths. The larger the node out-degree, the greater the impact of the safety risks on other safety risks, the larger the node in-degree, indicating that the risk is more susceptible to the impact of other risks, and the node in the betweenness centrality of the process of risk transfer plays an important role in the "bridge". It is the betweenness centrality of these risks caused the complex safety risk transfer in the risk networks. Lower safety awareness and violations of operation rules have high out-degrees and high betweenness centrality, indicating that they play a key "bridge" role in safety risk transfer. Insufficient safety education and training and insufficient safety checks or hidden trouble investigations have high out-degrees, and are the key risks in the risk transfer path. Therefore, according to the directed relationship among the risks, we can draw two key safety risk transfer paths, and they are insufficient safety education and training →lower safety awareness →violation of operation rules →safety accidents; insufficient safety checks or hidden trouble investigations → violation of operation rules → safety accidents.

5. Discussion and management implication

This paper identifies the safety risks and accident types of subway construction based on text mining algorithm. The safety risks include 5 first-level safety risks and 29 second-level safety risks, including human risk, material risk, environmental risk, technical risk and management risk. The accident types include collapse, high fall, object attack, vehicle injury, mechanical injury, explosion, electrocution and other accidents (drilling through tunnels, shield machine flooding, fire, poisoning, etc.). Same taxonomy of safety risks can be found in the existing literature. In the identification of bridge construction safety risks, most scholars applied the framework of "human- material - environmental - technical - management " or its deviants to examine the structures or lists of safety risk [71–74]. Compared with the safety risks identified in the past, the safety risks identified in this paper are less, because they are mined based on the accident investigation report, which is limited by the level of the accident investigation team, and some reasons may not be reflected in the accident investigation report.

This paper is based on the Apriori algorithm to calculate the causality of safety risks in the subway construction process. The Apriori algorithm is used to find the frequent itemset and use the frequent itemset to derive the association rules, and finally get the causal relationship [75]. Currently, the main methods to study the safety risks relationship are DEMATEL [76,77], SD [78,79], SEM [80,81] and ISM [82,83]. Compared with these methods, the Apriori algorithm has the following advantages: Firstly, the Apriori algorithm's computational speed is applicable to a wide range of large-scale datasets, and the algorithmic logic is clear and easy to understand. Secondly, the data of DEMATEL, SD, SEM, and ISM are from expert interviews, while the Apriori algorithm relies on accident investigation reports to avoid the influence of experts' subjectivity, making the examining results more accurate and reliable. In addition, some association rules with poor correlation are removed by setting the support and lift, and thus can make the results more reliable and salient.

Using the complex network model, it is found that the transfer of safety risks in subway construction mainly depends on key nodes (i.e., the key safety risks), and these key nodes include improper safety management, unimplemented safety subject responsibilities, violation of operation rules, non-perfect safety responsibilities system and insufficient safety education and training. Previous literature has also drew the consistent results [17,84]. For example, Li et al. [84] found that lower safety awareness, violations of operation rules, insufficient security checks and chaotic site management were the safety risks for SCSA. Interrelationships and interactions between risks are the root causes of safety risk accidents in metro construction, and the essence of risk transfer is the transfer path and process between risk nodes[85]. Based on 101 subway construction safety accident reports, this paper obtained two shorter key risk transfer paths in the subway construction safety network: insufficient safety education and training→lower safety awareness→violation of operation rules→safety accidents; insufficient safety checks or hidden trouble investigations→violation of operation rules→safety accidents, This is consistent with the risk transfer scenarios of real security incidents. For the safety risk relationship, most measures are to control the key nodes or cut off the connection between the key nodes[86].

Based on the research results, the following safety management measures can be proposed to manage safety at the site. (a) The manager should carefully analyze and identify the possible safety risks, and design safety risk measures based on the antecedent factors of these safety risks. (b) Strengthen the safety checks or hidden trouble investigations, and timely stop illegal activities and eliminate potential safety hazards. The manager should conduct comprehensive safety checks or hidden trouble investigations on a regular basis, and timely solve the problems to prevent the occurrence of potential safety accidents. (c) Managers should establish the safety responsibilities structure, clarify the responsibilities and tasks of each safety subject, and effectively implement safety measures to improve the effectiveness of safety management. (d) Managers should provide comprehensive safety education and training for workers, improve their safety awareness and accounting attention, let them understand safety rules and regulations and operating procedures, and form good safety habits in daily operations.

6. Results

This paper identifies SCSR based on text mining algorithms, and uses Apriori algorithm to examine the causal relationship between safety risks in the accident investigation report, and finally uses the complex network model to identify the key safety risks of subway construction and determine the shorter critical risk transfer path. The main research conclusions are as follows:

(1) The safety risks and accident types of subway construction are identified, including 5 types of first-level safety risks and 29 second-level safety risks. The first-level safety risks include human risk, material risk, environmental risk, technical risk and management risk. The accident types include collapse, high fall, object attack, vehicle injury, mechanical injury, explosion, electrocution and other accidents (drilling through tunnels, shield machine flooding, fire, poisoning, etc.).

(2) Improper safety management, unimplemented safety subject responsibilities, violation of operation rules, non-perfect safety responsibilities system and insufficient safety education and training are the key safety risks in SCSA. Two shorter key risk transfer paths in the subway construction safety network can be obtained: insufficient safety education and training→lower safety awareness→violation of operation rules→safety accidents; insufficient safety checks or hidden trouble investigations→violation of operation rules→safety accidents; In the process of risk transfer, the risk can be controlled by controlling the key nodes or cutting off the transfer path.

(3) The paper used complex network model to explore the safety risk transfer relationship of subway construction and came up with the key risk transfer nodes of subway construction and two shorter risk transfer paths. Studying risk transfer relationships in other engineering fields to validate the plausibility of the results of this study could be the next step in the research.

Author Contributions: Conceptualization and formal analysis, K.W.; investigation and original draft preparation, J.Z.; methodology, review and editing, Y.H.; project administration,H.W.; validation, H.L.&H.C.

Funding: This research was funded by the National Natural Science Foundation of China (grand number U1810203, 52178388), the Zhengzhou Metro Group Co. Ltd (grand number H22-541); the Fundamental Research Funds for the Univer-sities of Henan Province (grand number NSFRF230426), and the Doctoral fund of Henan Polytechnic University (grand number 760707/038).

Data Availability Statement: The data is available by querying the authors.

Acknowledgments: The research would like to thank Henan Polytechnic University for providing research facilities.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Central People's Government of the People's Republic of China. Circular of the state council on printing and distributing the development plan for modern comprehensive transportation system during the 14th five-year plan. **2020**. Available online: https://www.gov.cn/zhengce/content/2022-01/18/content_5669049.htm (accessed on 2 August 2023).

2. Ministry of Transport of the People's Republic of China. Quick report of urban rail transit operation data in march 2023. **2023**. Available online: https://www.mot.gov.cn/fenxigongbao/yunlifenxi/202304/t20230407_3790012.html. (accessed on 2 August 2023).
3. Liu, P. ; Wang, Y. ; Han, T. ; Xu, J. ; Li, Q. Safety evaluation of subway tunnel construction under extreme rainfall weather conditions based on combination weighting-set pair analysis model. Sustainability 2022, 14:9886.
4. Zhou, Z. ; Irizarry, J. Integrated framework of modified accident energy release model and network theory to explore the full complexity of the hangzhou subway construction collapse. Journal of Management in Engineering 2016, 32.
5. Zhou, Z. ; Liu, S. ; Qi, H. Mitigating subway construction collapse risk using bayesian network modeling. Automation in Construction 2022, 143.
6. Fang, M. ; Zhang, Y. ; Zhu, M.; Chen, S. Cause mechanism of metro collapse accident based on risk coupling. International Journal of Environmental Research and Public Health 2022, 19.
7. Yan, H.; Gao, C. ; Elzarka, H. ; Mostafa, K. ; Tang, W. Risk assessment for construction of urban rail transit projects. Safety Science 2019, 118:583-94.
8. Chai, G. ; Xu, Y. System dynamics analysis of security risk of subway construction from the perspective of schedule Science and Technology Management Research 2016, 36:85-90.
9. Wang, J. ; Chen, X. ; Wu, H. Simulation on safety capability of metro shield constructors based on sd model. Journal of Safety Science and Technology 2020, 16:108-14.
10. Li, X. ; Liao, F. ; Wang, C. ; Alashwal, A. Managing safety hazards in metro subway projects under complex environmental conditions. Asce-Asme Journal of Risk and Uncertainty in Engineering Systems Part a-Civil Engineering 2022, 8.
11. Li, H. ; Chen, H. ; Cheng, B. ; Hu, X. ; Cai, X. Study on formation model of subway construction safety climate based on fuzzy ism-dematel. Journal of Railway Science and Engineering 2021, 18:2200-8.
12. Fan, Y. ; Wang, R. Application of matter-element extension method in safety risk assessment of subway shield construction. Journal of Safety and Environment 2023, 23:1779-90.
13. Lu, X. ; Xu, C. ; Hou, B. ; Du, X. ; Li, L. Risk assessment of metro construction based on dynamic bayesian network. Chinese Journal of Geotechnical Engineering 2022, 44:492-501.
14. Wu, X. ; Ding, B. ; Zhang, L. ; Chen, Y. ; Xue, L. ; Song, R. Research on risk management of subway construction based on bayesian network. China Safety Science Journal 2014, 24:84-9.
15. Zhang, L. ; Liu, S. ; Li, K. ; Xu, J. ; Wang, S. ; Li, Q., et al. Prediction of shield construction risks in subway tunnelling based on fault tree and bayesian network. Modern Tunneling Technology 2021, 58:21-9+55.
16. Wu, B. ; Cai, Q. ; Liu, C. ; Huang, W. ; Xie, Y. Multi-scale evaluation model and application of safety risk in urban subway station construction. Journal of Safety and Environment 2023, 23:633-41.
17. Pan, H.; Gou, J. ; Wan, Z.; Ren, C. ; Chen, M. ; Gou, T.; Luo, Z. Research on coupling degree model of safety risk system for tunnel construction in subway shield zone. Mathematical Problems in Engineering 2019, 2019.
18. Guo, Q.; Hao, Q.; Wang ,Y.; Wang, J. Subway system resilience evaluation in based on anp-extension cloud model. Journal of System Simulation 2021, 33:943-50.
19. Fan, B. ; Dong , B. ; Wang, B. ; Li , M. ; Wu, S. ; Tong, R. Identification and application of unsafe behaviors of subway construction workers based on deep learning. China Safety Science Journal 2023, 33:41-7.
20. Feng, L. ; Zhang, L. Assessment of tunnel face stability subjected to an adjacent tunnel. Reliability Engineering & System Safety 2021, 205.
21. Wang, X.; Xia, N.; Zhang, Z.; Wu, C.; Liu, B. Human safety risks and their interactions in china's subways: Stakeholder perspectives. Journal of Management in Engineering 2017, 33.
22. Pan, H. ; Li , G. ; Chen, Y. ; Lwo, Z. ; Han, J. Critical factors influencing safety vulnerability of metro shield tunneling based on structural equation model importance-performance analysis. Tunnel Construction 2023:1-12.
23. Yu, Q.; Ding, L. ; Zhou, C. ; Luo, H. Analysis of factors influencing safety management for metro construction in china. Accident; analysis and prevention 2014, 68:131-8.
24. Zhang, L.; Wang, J.; Wu, H. ; Wu, M.; Guo, J.; Wang, S. Early warning of the construction safety risk of a subway station based on the lssvm optimized by qpso. Appl Sci-Basel 2022, 12.

25. Qie, Z. ; Yan, H. A causation analysis of chinese subway construction accidents based on fault tree analysis-bayesian network. Frontiers in psychology 2022, 13:887073.
26. Fang, J. ; Guo, P. ; Zhu, K. ; Chen, Z. Coupling evolution analysis of subway tunnel construction safety risk based on n-k model. China Safety Science Journal 2022, 32:1-9.
27. Wu, L.; Bai, H.; Yuan, C. ; Xu, C. Fanpce technique for risk assessment on subway station construction. Journal of Civil Engineering and Management 2019, 25:599-616.
28. Luo, Z.; Guo, J.; Han, J.; Wang, Y. Research on the construction safety risk assessment of prefabricated subway stations in china. Engineering Construction and Architectural Management 2022.
29. Guo, W. Safety risk assessment of tourism management system based on pso-bp neural network. Computational Intelligence and Neuroscience 2021, 2021.
30. Marsadek, M. ; Mohamed, A. Risk based security assessment of power system using generalized regression neural network with feature extraction. Journal of Central South University 2013, 20:466-79.
31. Chen, Y.; Zheng, W.; Li, W.; Huang, Y. Large group activity security risk assessment and risk early warning based on random forest algorithm. Pattern Recognition Letters 2021, 144:1-5.
32. Lu, J. ; Su, W. ; Jiang, M. ; Ji, Y. Severity prediction and risk assessment for non-traditional safety events in sea lanes based on a random forest approach. Ocean & Coastal Management 2022, 225.
33. Dong, Y.; Sun, B. ; Wang, G. Research on modeling method of power system network security risk assessment based on object-oriented bayesian network. Energy Reports 2021, 7:289-95.
34. Zhou, B.; Sun, B.; Zang, T.; Cai, Y.; Wu, J.; Luo, H. Security risk assessment approach for distribution network cyber physical systems considering cyber attack vulnerabilities. Entropy 2023, 25.
35. Jiang, R. ; Ma, Z.; Yang, J. An assessment model for cloud service security risk based on entropy and support vector machine. Concurrency and Computation-Practice & Experience 2021, 33.
36. Wang, X. ; Li, L. ; Sun, J. ; Jin, X. ; Sun, T. ; Bai , Y. Risk pre-warning of hazardous materials in cereal supply chain based on deep belief network-multiclass fuzzy support vector machine (dbn-mfsvm). Food Science 2020, 41:17-24.
37. Zhang, G.; Jiao, Y.; Chen, L.; Wang, H. ; Li, S. Analytical model for assessing collapse risk during mountain tunnel construction. Canadian Geotechnical Journal 2016, 53:326-42.
38. Wen, Z. ; Xia, Y.; Ji, Y.; Liu, Y.; Xiong, Z.; Lu, H. Study on risk control of water inrush in tunnel construction period considering uncertainty. Journal of Civil Engineering and Management 2019, 25:757-72.
39. Manchao, H. ; Sousa, R.; Muller, A. ; Vargas, E. ; Sousa, L.; Xin, C. Analysis of excessive deformations in tunnels for safety evaluation. Tunnelling and Underground Space Technology 2015, 45:190-202.
40. Jiang, X. ; Hu, W. ; Yuan, X. ; Sun, Z. ; Zhang, X. Research on bp-sd model for evolution of construction safety risk of subway tunnel. Journal of Safety Science and Technology 2017, 13:67-72.
41. Eybpoosh, M. ; Dikmen, I. ; Birgonul, M. T. Identification of risk paths in international construction projects using structural equation modeling. Journal of Construction Engineering and Management 2011, 137:1164-75.
42. Zhou, Z.; Irizarry, J. ; Guo, W. A network-based approach to modeling safety accidents and causations within the context of subway construction project management. Safety Science 2021, 139.
43. Yan, W. ; Men, X. ; Yang, F. Risk assessment of metro construction based on interaction matrix. Engineering Journal of Wuhan University 2019, 52:796-801.
44. Hou, G. ; Liu, W. ; Li, L. ; Ma, X. ; Mu, X. ; Liu, Y. Vulnerability analysis of the subway construction safety system with coupled multiple risk factors. China Civil Engineering Journal 2022, 55:111-9.
45. Liu, P. ; Li, Q. ; Bian, J. ; Song, L.; Xiahou, X. Using interpretative structural modeling to identify critical success factors for safety management in subway construction: A china study. International Journal of Environmental Research and Public Health 2018, 15.
46. Wang, W.; Wang, Y.; Wang, G.; Li, M. ; Jia, L. Identification of the critical accident causative factors in the urban rail transit system by complex network theory. Physica a-Statistical Mechanics and Its Applications 2023, 610.
47. Zheng, L. ; Wang, F. ; Zheng, X.; Liu, B. Discovering the relationship of disasters from big scholar and social media news datasets. International Journal of Digital Earth 2019, 12:1341-63.
48. Guo, S. ; Zhou, X.; Tang, B. ; Gong, P. Exploring the behavioral risk chains of accidents using complex network theory in the construction industry. Physica a-Statistical Mechanics and Its Applications 2020, 560.
49. Chen, F. ; Wang, H. ; Xu, G. ; Ji, H.; Ding, S.; Wei, Y. Data-driven safety enhancing strategies for risk networks in construction engineering. Reliability Engineering & System Safety 2020, 197.

50. Qiu, Z.; Liu, Q. ; Li, X.; Zhang, J.; Zhang, Y. Construction and analysis of a coal mine accident causation network based on text mining. Process Safety and Environmental Protection 2021, 153:320-8.
51. Zhou, J. ; Xu, W.; Guo, X. ; Ding, J. A method for modeling and analysis of directed weighted accident causation network (dwacn). Physica a-Statistical Mechanics and Its Applications 2015, 437:263-77.
52. Li, T. ; Rong, L. A comprehensive method for the robustness assessment of high-speed rail network with operation data: A case in china. Transportation Research Part a-Policy and Practice 2020, 132:666-81.
53. Li, T. ; Rong, L.; Yan, K. S. Vulnerability analysis and critical area identification of public transport system: A case of high-speed rail and air transport coupling system in china. Transportation Research Part a-Policy and Practice 2019, 127:55-70.
54. Lam, C.; Tai, K. Network topological approach to modeling accident causations and characteristics: Analysis of railway incidents in japan. Reliability Engineering & System Safety 2020, 193.
55. Chen, F.; Ji, H.; Wei, Y. Using network theory to explore the risk characteristics of bridge-tunnel hybrid construction. Ieee Access 2019, 7:116038-46.
56. Nikiforos, M. N. ; Voutos, Y. ; Drougani, A. ; Mylonas, P. ; Kermanidis, K. L. The modern greek language on the social web: A survey of data sets and mining applications. Data 2021, 6:52.
57. Zhu, F. ; Patumcharoenpol, P. ; Zhang, C. ; Yang, Y. ; Chan, J. ; Meechai, A.; Vongsangnak, W.; Shen, B. Biomedical text mining and its applications in cancer research. Journal of Biomedical Informatics 2013, 46:200-11.
58. Singhal, A. ; Leaman, R. ; Catlett, N. ; Lemberger, T. ; McEntyre, J. ; Polson, S.; Xenarios, I.;Arighi, C.; Lu, Z. Pressing needs of biomedical text mining in biocuration and beyond: Opportunities and challenges. Database-the Journal of Biological Databases and Curation 2016.
59. Drury, B. ; Roche, M. A survey of the applications of text mining for agriculture. Computers and Electronics in Agriculture 2019, 163.
60. Ferreira-Mello, R. ; Andre, M. ; Pinheiro, A. ; Costa, E. ; Romero, C. Text mining in education. Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery 2019, 9.
61. Marzouk, M. ; Enaba, M. Text analytics to analyze and monitor construction project contract and correspondence. Automation in Construction 2019, 98:265-74.
62. Huang, Y.; Zhang, Z. ; Tao, Y. ; Hu, H. Quantitative risk assessment of railway intrusions with text mining and fuzzy rule-based bow-tie model. Advanced Engineering Informatics 2022, 54.
63. Xu, H. ; Liu, Y. ; Shu, C. ; Bai, M.; Motalifu, M. ; He, Z.; Wu, S. ; Zhou, P. ; Li, B. Cause analysis of hot work accidents based on text mining and deep learning. Journal of Loss Prevention in the Process Industries 2022, 76.
64. Li, S. ; You, M.; Li, D. ; Liu, J. Identifying coal mine safety production risk factors by employing text mining and bayesian network techniques. Process Safety and Environmental Protection 2022, 162:1067-81.
65. Agrawal, R. ; Imieliński, T. ; Swami, A. Mining association rules between sets of items in large databases. Proceedings of the 1993 ACM SIGMOD international conference on Management of data; Washington, D.C., USA: Association for Computing Machinery (1993).
66. Telikani, A. ; Gandomi, A. H. ; Shahbahrami, A. A survey of evolutionary computation for association rule mining. Information Sciences 2020, 524:318-52.
67. Czibula, G. ; Czibula, I. G. ; Miholca, D. L. ; Crivei, L. M. A novel concurrent relational association rule mining approach. Expert Systems with Applications 2019, 125:142-56.
68. DAI, R. Research on system science and system complexity. Journal of System Simulation 2002:1411-6.
69. Chen, Y. ; Feng, W. ; Jiang, Z. ; Duan, L. ; Cheng, S. An accident causation model based on safety information cognition and its application. Reliability Engineering & System Safety 2021, 207:107363.
70. Tetzlaff, E. J. ; Goggins, K. A. ; Pegoraro, A. L. ; Dorman, S. C. ; Pakalnis, V. ; Eger, T. R. Safety culture: A retrospective analysis of occupational health and safety mining reports. Safety and Health at Work 2021, 12:201-8.
71. Li, Q.; Zhou, J.; Feng, J. Safety risk assessment of highway bridge construction based on cloud entropy power method. Appl Sci-Basel 2022, 12.
72. Ji, T. ; Liu, J. ; Li, Q. Safety risk evaluation of large and complex bridges during construction based on the delphi-improved fahp-factor analysis method. Advances in Civil Engineering 2022, 2022.
73. Chen, F. ; Wu, X. ; Wei, Y. Causation analysis for bridge-tunnel hybrid construction accident based on fism-dematel. IFAC-PapersOnLine 2022, 55:1429-34.

74. Li, G. ; Ran, R.; Fang, J. ; Peng, H. ; Wang, S. Early warning for the construction safety risk of bridge projects using a rs-ssa-lssvm model. Advances in Civil Engineering 2021, 2021.
75. Chen, D.; Pei, Y.; Xia, Q. Research on human factors cause chain of ship accidents based on multidimensional association rules. Ocean Engineering 2020, 218.
76. Yang, Y.; Wang, Y. ; Easa, S. M. ; Yan, X. B. Factors affecting road tunnel construction accidents in china based on grounded theory and dematel. International Journal of Environmental Research and Public Health 2022, 19.
77. Xu, K. ; Li, S. ; Liu, J. ; Lu, C. ; Xue, G. ; Xu, Z.; He, C. Evaluation cloud model of spontaneous combustion fire risk in coal mines by fusing interval gray number and dematel. Sustainability 2022, 14.
78. Xu, N. ; Liu, Q. ; Ma, L. ; Deng, Y.; Chang, H. ; Ni, G., et al. A hybrid approach for dynamic simulation of safety risks in mega construction projects. Advances in Civil Engineering 2020, 2020.
79. Zhang, Y.; Xing, X.; Antwi-Afari, M. F. ; Wu, M. Safety risk estimation of construction project based on energy transfer model and system dynamics: A case study of collapse accident in china. International Journal of Environmental Research and Public Health 2022, 19.
80. Li, X.; Wang, C. ; Kassem, M. A. ; Alhajlah, H. H. ; Bimenyimana, S. Evaluation method for quality risks of safety in prefabricated building construction using sem-sdm approach. International Journal of Environmental Research and Public Health 2022, 19.
81. Yang, Y.; Wang, Y. ; Easa, S. M. ; Yan, X. Risk factors influencing tunnel construction safety: Structural equation model approach. Heliyon 2023, 9.
82. Du, H. ; Zhao, Y. Research on influencing factors for risk of security inspection system in civil transportation airport. Journal of Safety Science and Technology 2020, 16:37-43.
83. Gu, W. ; Wang, S. Collapse risk analysis for the loess-fea tured tunnels in railway construction based on ism and fuzzy fault tree method. Journal of Safety and Environment 2016, 16:31-6.
84. Li, J. ; Wang, J. ; Xu, N. ; Zhou, Z. Analysis of safety risk factors for metro construction based on text mining method. Tunnel Construction 2017, 37:160-6.
85. Huangfu, Y. ; Xu, J. ; Zhang, Y. ; Huang, D. ; Chang, J. Research on the risk transmission mechanism of international construction projects based on complex network. PloS one 2023, 18:e0285497.
86. Huang, X. ; Jin, J. ; Lin, Z. ; Chen, L. ; Liu, J. Complex network-based analysis on fire disaster chain of urban deep underground space and intelligent disaster mitigation measures. Journal of Basic Science and Engineering 2021, 5:1280-91.