Analysis on Relationships of Safety Risk Factors in Metro Construction

Xu Na1,2,*, Wang Jianping1, Li Jie1 and Ni Guodong1,3
1 School of Mechanics and Civil Engineering, China University of Mining & Technology, Xuzhou 221000, China
2 State Key Laboratory for Geo-Engineering, Xuzhou 221000, China
3 A. James Clark School of Engineering, University of Maryland, College Park, MD, 20742, USA

Received 8 April 2016; Accepted 19 October 2016

Abstract
Accidents in metro construction occur frequently, causing significant economic loss, mass casualties, and adverse social effect. Thus, conducting effective analyses on safety risk factors in metro construction is urgent. Safety risk factors are probably related because of the complexity of metro construction, which is often disregarded in traditional models of risk analysis. A framework to explore the mutual influence among risk factors to enhance accuracy in risk assessment was presented. A list of 43 safety risk factors was extracted using text mining method from 100 investigation reports on metro construction accidents, which were obtained from the official website of China State Administration of Work Safety. A multi-layer interpretative structural model was proposed based on the analysis of the pairwise correlation of risk factors to determine the contextual interactions among risks. A case study of Shenzhen Metro Line 9 project was conducted to explore the effect of the proposed model. Results show that the safety status of metro construction is affected by risks in field work and closely related to the risks in geological investigation and design work. The risks in geology, design, safety systems, and personnel quality are the root causes of accident among all the risk factors. Root risks do not directly result in safety accidents but cause potential uncertainties to indirect risks, which magnify the likelihood of direct risks. Direct risks, such as material risks and work operation risks, directly cause accidents in metro construction. The study plays an important role in assisting decision makers in formulating different precaution strategies, and the proposed analysis procedure and method can be applied to other similar and complex projects.

Keywords: Metro Construction, Safety Risks, Relationships of Risk Factors, Text Mining, Interpretative Structural Modelling

1. Introduction
Metro construction is always connected with substantial risks because of unpredictable geological and hydrological conditions, complex construction equipment, and methods [1]. Risk management (RM) traditionally consists of four major steps, namely, risk identification, risk analysis (including quality analysis and quantity analysis), risk response, and risk control [2]. Although some project managers claim that they perform RM in project management, accidents still frequently occur because of the lack of effective identification and analysis of safety risk factors. Existing studies mostly focus on independent risks, which seldom exist in reality. The relationships of risk factors are not identified, and are thus unmanaged. Most risk factors are interrelated and exhibit complex relationships. For example, risk factors “weather” and “soil condition,” as well as risk factors “material availability” and “management quality,” are correlated [3]. If the relationships of risk factors were not considered, potential losses would be disregarded or underestimated and adverse consequences would occur. Therefore, an overall identification of safety risk factors in metro construction using text mining (TM) method is presented and a model is developed to analyze the hierarchical structure and relationships of risk factors based on interpretative structural modelling method. A case study on Shenzhen Metro Line 9 construction is conducted to explore the effect of the proposed model. This paper aims to reveal the overall risk factors and their relationships, which lead to accidents in metro construction, and increase the accuracy of risk evaluation and effectiveness of risk response.

2. State of the Art
The definitions of “risk communication,” “risk correlation,” and “risk relationship” were first studied in the field of medical risks and financial risks in the 1990s. In 1994, Ren [4] introduced the definition of “risk relationship” to the field of risk evaluation in construction projects and proposed four basic patterns, i.e., independence, dependence, parallel, and series. Recent studies focused on the interdependence of risks in analyzing and evaluating engineering project risks [5]. Vidal [6] pointed out that the limitation of the traditional approach of risk classification was that project risk interactions were not properly considered, and a clustering approach was proposed to group risks to maximize the project risk interaction rate inside clusters and minimize it outside clusters. Bu-Oammaz [7] established analytic network process (ANP) model and analyzed the degree of influence of 28 international engineering project risk factors using a pairwise comparison matrix. Luu et al. [8] conducted a quantitative analysis of construction risks in developing countries using Bayesian network (BN) and proposed 18 causal relationships among 166 risk factors. Even though CA, ANP, and BN reflect the pairwise correlation between
risk factors, these methods cannot present the hierarchical structure and the relationships of risk factors. Interpretable structural modelling (ISM) is a graphic presentation of the hierarchical arrangement concerning the relationship of one element to other elements. ISM was first proposed by J. Warfield in 1973 [9]. Lyer [10] identified 17 risks encountered during the development stage of public and private projects in India and used ISM to develop a hierarchical structure. Huerga et al. [11] applied cross-impact analysis and ISM to analyze the complex cascading effects in operational RM to determine the relationships between different risks and consequences, direct links, indirect links, and cascading effects. Therefore, the ISM method is especially suitable for analyzing the relationships of risk factors in infrastructure projects that have numerous variables, complex relations, and unclear structures [12].

The safety risks in metro construction and the relationships of risk factors have attracted the interest of several scholars. For example, Xie [13] combined decision-making trial and evaluation laboratory method with ISM to analyze the relations of the factors that cause collapse accidents in foundation pit engineering construction and established a five-layer hierarchical structure of a system to determine the causes of such accidents. Guo et al. [14] built an ISM model of risks in metro shield construction from three aspects, namely, construction technology, management, and environment. These research achievements laid the foundation for safety RM and control of metro construction. However, the occurrence of accidents in metro construction is mostly caused by inaccurate geological investigations and deficient designs according to the statistical analyses of safety accidents [15]. Thus, risk factors in the geological investigation and design stage will be considered during the process of identifying risks and analyzing the relationships of risk factors.

The remainder of this paper is organized as follows: Section 3 establishes the analysis procedure and method. Section 4 analyzes the relationships of safety risk factors and discusses the applicability of the model through a case study. Section 5 summarizes the conclusions.

3. Methodology

3.1 Risk identification

Expert judgment methods, such as Delphi and brainstorming, are commonly used in the investigation on project managers to determine risk factors [16], [17]. However, organizing a group of experts is difficult and the results are limited by the experiences of these experts. TM was extensively developed as data scientists focused on analyzing unstructured data [18]. TM is a powerful method to automatically extract important words and phrases within a set of documents identified during a screening process [19]. The data on accidents in metro construction are available and easy to obtain. According to the statistics, approximately five fatal accidents occur per 100 kilometres in metro construction [20]. Investigations of these accidents include the causes, liable divisions, and other information, which form a large database of safety risk factors in metro construction. A total of 100 investigation reports on construction accidents were collected from the official website of the China State Administration of Work Safety, and the causes of each accident were analyzed based on the TM method. The process includes four steps, as shown in Figure 1.

![Fig.1. Risk-identification process based on text mining method](image)

Step 1: Document collection: This task involved collecting 100 investigation reports of metro construction accidents from 2003 to 2015 in China, and establishing data folders in accordance with standard rules.

Step 2: Text preprocessing: Text cleaning was conducted by excluding useless words and terms, such as "according to" and "due to," to avoid identification deviation. The entire investigation report was disintegrated into several words and paragraphs with different meanings through text segmentation. Thus, the investigation reports were converted to a form that could be analyzed by TM tool.

Step 3: Text intermediate form: This step involved obtaining key words that indicate the causes of accidents through word frequency.

Step 4: TM results: After the above steps, a total of 43 risk factors were sorted and the results are shown in Table 1. The set of risk factors is defined as $S_i$, and the number of risk factors is $n$.

<table>
<thead>
<tr>
<th>ID</th>
<th>Risk factor set ($S_i$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_0$</td>
<td>Deficiency in report of geological investigation</td>
</tr>
<tr>
<td>$S_1$</td>
<td>Inappropriate planning of construction arrangement</td>
</tr>
<tr>
<td>$S_2$</td>
<td>Ineffective safety performance measures for construction</td>
</tr>
<tr>
<td>$S_3$</td>
<td>Incomplete management procedures for construction</td>
</tr>
<tr>
<td>$S_4$</td>
<td>Incomplete technical specifications for construction</td>
</tr>
<tr>
<td>$S_5$</td>
<td>Deviation of construction safety management</td>
</tr>
<tr>
<td>$S_6$</td>
<td>Unsuitable construction instructions</td>
</tr>
<tr>
<td>$S_7$</td>
<td>Tight construction schedule</td>
</tr>
<tr>
<td>$S_8$</td>
<td>Inadequate personnel quality</td>
</tr>
<tr>
<td>$S_9$</td>
<td>Insufficient personnel experience</td>
</tr>
<tr>
<td>$S_{10}$</td>
<td>Wearing safety device improperly</td>
</tr>
<tr>
<td>$S_{11}$</td>
<td>Inappropriate mechanical operation</td>
</tr>
<tr>
<td>$S_{12}$</td>
<td>Unexpected underground pipeline</td>
</tr>
<tr>
<td>$S_{13}$</td>
<td>Complex environment of construction site</td>
</tr>
<tr>
<td>$S_{14}$</td>
<td>Bad weather situation</td>
</tr>
<tr>
<td>$S_{15}$</td>
<td>Complex hydrologic condition</td>
</tr>
<tr>
<td>$S_{16}$</td>
<td>Complex geological condition</td>
</tr>
<tr>
<td>$S_{17}$</td>
<td>Unexpected situation of surrounding buildings</td>
</tr>
<tr>
<td>$S_{18}$</td>
<td>Unexpected situation of surrounding structures</td>
</tr>
<tr>
<td>$S_{19}$</td>
<td>Incomplete design drawings</td>
</tr>
<tr>
<td>$S_{20}$</td>
<td>Unsuitable construction targets</td>
</tr>
<tr>
<td>$S_{21}$</td>
<td>Irrational design calculations</td>
</tr>
<tr>
<td>$S_{22}$</td>
<td>Irrational design parameters</td>
</tr>
<tr>
<td>$S_{23}$</td>
<td>Variations of construction scope</td>
</tr>
<tr>
<td>$S_{24}$</td>
<td>Inadequate safety knowledge of workers</td>
</tr>
<tr>
<td>$S_{25}$</td>
<td>Poor safety culture</td>
</tr>
<tr>
<td>$S_{26}$</td>
<td>Incomplete safety systems</td>
</tr>
</tbody>
</table>

Table 1. Safety risk factor set of metro construction
### 3.2 Analysis on relationships of risk factors

The risk factors are not isolated from one another, and directly or indirectly influence one another [5]. ISM, which is a powerful modelling approach to establish a hierarchical structure, is used to analyze the relationships of risk factors in metro construction. The implementation steps are as follows:

Step 1: To establish adjacency matrix $A$. Based on a questionnaire survey, an $n \times n$ risk correlation judgment matrix $A$ is established using the pairwise comparison method, which is commonly used in analytical hierarchy process method. Matrix $A$ indicates whether the risk factors influence one another. Matrix element $a_{ij}$ indicates factor $S_i$ influences factor $S_j$ and matrix element $a_{ij} = 0$ if the elements do not directly influence one another. The maximum eigenvalue and eigenvector of this adjacency matrix must be solved and checked must be conducted to ensure consistency.

Step 2: To establish reachable matrix $M$. Equation 1 is used to solve reachable factors, where $I$ is the unit matrix. Reachable matrix $M$ is established by adopting Boolean algebra algorithm for squaring until all the results after a certain power are equal. Reachable matrix indicates the path that connects one element to another. It also is the cause and effect path between each element.

\[
(A + I)^k - 1 = (A + I)^k = (A + I)^k + 1 = M
\]

Step 3: To solve reachable set $R(S)\text{,}$ antecedent set $A(S)$, and common set $C(S)$. Equation 2 is used to solve reachable set $R(S)$ by using the element corresponding to the column whose value in row $i$ of reachable matrix $M$ is 1, which indicates the set of all the factors reachable from factor $S_i$. Equation 3 is used to solve antecedent set $A(S)$ by using the element corresponding to the column whose value in column $i$ of reachable matrix $M$ is 1, which indicates the set of all the factors reachable to factor $S_i$. A common set is the intersection set of reachable set $R(S)$ and antecedent set $A(S)$, whose details are indicated in Eq. 4.

\[
R(S) = \{S_i | M_{ij} = 1\}
\]

\[
A(S) = \{S_i | M_{ij} = 1\}
\]

\[
C(S) = R(S) \cap A(S)
\]

Step 4: To conduct region division and level division. Region division $\Pi_0 = \{P_i\}$ refers to the process of dividing sets of system elements into several mutually independent regions $P_i$. Judgment is conducted by using the elements in initial set $B(S)$ (see Eq. 5). These elements belong to different regions if $R(S_j) \cap R(S_i) = \emptyset$. Level division $\Pi_0 = \{L_i\}$ aims to determine the status of the level at which each element within this region is located. Equation 6 is used to conduct judgment. The elements of the same set $L_i$ are at the same level. Then, the rows and columns corresponding to elements of $L_i$ are deleted from reachable set $M$ to obtain matrix $M'$. Afterwards, steps 3-4 are conducted for $M'$ to determine the other regions in sequence until all regions are divided.

\[
B(S) = \{S_i | C(S) = A(S)\}
\]

\[
L_i = \{S_i | C(S) = R(S)\}
\]

Step 5: To draw a multi-layer hierarchical digraph to convert it to the interpretation structure model. The digraph that is arranged by levels with the connection relation is established according to region and level division. After the hierarchical relation is converted to the interpretative structural model, the hierarchical relation between risk factors can be analyzed.

Step 6: To interpret and verify the model based on theoretical knowledge and engineering experience. In case of inconformity, step 1 is repeated to modify and rebuild the model considering the definition of risk factors and their correlativity.

### 4. Result Analysis

#### 4.1 Case study

China developed a rail transit construction market with the largest scale and fastest growth in the world. In 2015, the total distance of metro construction in the country reached 2,530 km. Its operation distance is estimated to reach 7500 km in 2020. Shenzhen is a coastal city in south China that borders Hong Kong. Shenzhen covers a total area of 1,952 km². China’s first special economic zone, Shenzhen underwent rapid population growth. Metro Line 9 was planned to ease traffic pressure.

The Metro Line 9 project is located at the center of Shenzhen City. It adopts the build-transfer mode and was built by China Construction Company Ltd. The total length of Metro Line 9 is approximately 25.38 km (all underground), which includes 22 stations (see Figure 2). The average length between adjacent stations is 1.17 km with a maximum station spacing of 2.985 km and minimum station spacing of 0.376 km. Construction began on July 2012. The
planned construction period is 40 months, and the target completion date is on December 30, 2016. The total investment is around 16 billion RMB.

Owing to its location, Metro Line 9 is surrounded by high-density buildings and structures, and part of it is close to the coastline. Safety risk analysis and mitigation are crucial because of the complicated geological conditions, tight construction period, and complicated construction technology.

4.2 ISM building

First, questionnaires were answered by designers, contractors, subcontractors, and consultants who have extensive metro construction experience. The data of risk factor pairwise correlation with preferable convergence were obtained after distributing two sets of questionnaires. Consistency check was conducted before building the 43 × 43 adjacency matrix \(A_{43}\), which is shown in Figure 3. 

Second, the calculation of reachable matrix was conducted by using MATLAB software according to Eq. 1. When power \(k = 12\), a reachable matrix \(M_{12}\) was obtained. Reachable matrix \(M_{12}\) is presented in Figure 4.

Third, Eq. 2-4 were used to calculate reachable set \(R(S_1)\), antecedent set \(A(S_1)\), and common set \(C(S_1)\). According to the reachable matrix, this system has a total of 12 levels of sets. By taking the first level as an example, Table 2 lists the reachable, antecedent, and common sets of the first level.

Then, according to Eq. 5, the initial set of the project is \(B(S_1) = \{S_1, S_{16}, S_{18}, S_{19}, S_{21}, S_{22}, S_{24}, S_{29}, S_{30}, S_{31}\}\). Because \(R(S_1) \cap R(S_{19}) \cap R(S_{12}) \cap R(S_{29}) \cap R(S_{10}) \cap R(S_{13}) \neq \emptyset\) and \(R(S_{19}) \cap R(S_{21}) \cap R(S_{22}) \neq \emptyset\), thus, all risk factors were divided into regions \(P_1\) and \(P_2\). Correlativity does not exist between factors of different regions, but each region exhibits internal correlation integrity, namely,

\[\Pi_1 = P_1, P_2 = \{S_1, S_{16}, S_{19}, S_{21}, S_{24}, S_{29}, S_{31}\} \cap \{S_{17}, S_{18}, S_{21}, S_{22}\}\]

According to Eq. 4 and Eq. 6,

\[C(S_1) = R(S_1) \cap A(S_1) = R(S_1)\]

\[C(S_1) = R(S_1) \cap A(S_1) = R(S_1)\]

\[C(S_1) = R(S_1) \cap A(S_1) = R(S_1)\]

\[C(S_1) = R(S_1) \cap A(S_1) = R(S_1)\]

Therefore, \(S_{14}, S_{15}, S_{17}, S_{19}\), and \(S_{21}\) can be deemed as the first level node, namely, \(L_1 = \{S_{14}, S_{15}, S_{17}, S_{19}, S_{21}\}\). The corresponding nodes of Table 2 were ruled out at the same time to obtain the reachable set and antecedent set of the second level.

By employing this analogy, the final level division of factors is shown in Table 3, namely, \(\Pi_0 = \{L_1, L_2, \ldots, L_{11}\}\). According to the descriptions of each risk factor, the levels were roughly classified into three layers. The risk factors from level \(L_1\) to \(L_4\) were classified as direct risks, which directly influence the safety of metro construction, and risk factors from level \(L_5\) to \(L_{11}\) were grouped as root risks of all risk factors that affect the safety level of metro construction by level. The remaining middle levels from level \(L_5\) to \(L_{10}\) were considered indirect risks.

Based on the preceding analysis, the hierarchical structure of the overall risk factors affecting the safety of metro construction was established and converted to the interpretative structural model shown in Figure 5.
the direction of arrows shows the direction of influence. The 

4.3 Results analysis

(1) Arrows from bottom to top indicate that the risk factors of the lower level influence that of the higher level because the direction of arrows shows the direction of influence. The analysis results verify that the risk factors in metro construction are not completely independent. However, the relationships of interaction and potential transmission are completely independent. The relationships of risk factors must be considered in risk evaluation and risk response.

(2) According to the results of region division, risk factors “Bad weather situation” (S11), “Unexpected situation of surrounding buildings” (S12), and “Unexpected situation of construction site” (S17) exhibit a common influence on risk factor “Complex environment of construction site” (S10). These risk factors are concerned with external environmental risks. Thus, these risk factors form a region separately and do not affect other risk factors. Risk factors “Unexpected underground pipeline” (S16), “Complex hydrologic condition” (S19), and “Complex geological condition” (S20) are risks from the external environment. However, these risk factors significantly influence risk factor “Deficiency in report of geological investigation” (S22). Thus, these risk factors are classified into another region.

(3) The structure is clustered into three layers, namely, root risks, indirect risks, and direct risks, as shown in Table 3 and Figure 5. Root risks do not directly result in accidents and are concerned with the status of geological investigation, design, safety systems, and personnel quality, such as “Inadequate geological investigation (S22),” “Incomplete design drawings (S23),” “Incomplete safety systems (S26),” and “Inadequate personnel quality (S27).” Root risks transmit the potential loss to indirect risks, which are concerned with the status of construction management, including “Variations of construction scope (S27)” and “Weak supervision in construction process (S28).” Direct risks are concerned with the status of materials and work operation, such as “Poor material quality (S20)” and “Inappropriate work operation (S21).” Direct risks are affected by indirect risks and lead to accidents once these risks occur.

(4) Risk factors S1, S2, S16, S19, and S20 are at the bottom level, which transmit uncertainties along the arrows
level by level. Geological investigation is contained by personnel quality, work time, and other aspects. Thus, the report on the investigation may not reflect the actual geological condition and may influence the design aspect, which leads to “Incomplete design drawings (SI3)” and “Irrational design parameters (SI6).” Therefore, conducting a complete and detailed investigation and providing design drawings in the preconstruction stage are crucial.

(5) Risk factor pairs “SI and SI4”, “SI2 and SI8”, and “SI3 and SI5” exhibit strong connections of mutual effects and interpenetrations. When one risk factor exists, other risk factors also occur. All the risk factors are located at the bottom of the model, which shows that these risk factors exhibit a strong potential transmission. Therefore, these three pairs of risk factors warrant further examination.

(6) Numerous arrows are indicated from “Poor material quality (SI3)” and “Ineffective safety performance measures for construction (SI6),” which means that these risk factors influence other risk factors and should be deemed as key objects of supervision and control.

**5. Conclusions**

Metro construction involves significant and complex safety risks, which are correlated with each other. To identify and analyze the relationships of safety risk factors, an overall risk factor list is presented from published investigations of accidents using the TM method. A model is proposed to analyze the relationships of the risk factors in metro construction based on the ISM method to increase the accuracy of risk assessment and effectiveness of risk response for decision makers. The main conclusions are as follows:

(1) The TM method is adaptive to safety risk identification in metro construction because the required input data are objective and available. The obtained risk list indicates that metro construction accidents are mostly caused by geological risks. The risks caused in geological investigation and design stage are potential risks and need to be dealt with properly. Otherwise, these risks may conceal and cause...
uncertainties in the construction stage.

(2) Safety risk factors are related and form a three-layered structure, including root risks, indirect risks, and direct risks. The risks in geology, design, safety systems, and personnel quality are the root risks with high transmissibility and concealment ability, which constantly and significantly affect the safety of metro construction. The risks in construction safety management are indirect risks, which indirectly lead to accidents and act as carriers that transmit root risks to direct risks. The risks in materials and work operation are direct risks that directly result in accidents.

(3) A case study verified that the ISM model operates well and produces rational results on the relationships of risk factors.

The current study presents a framework to analyze the relationships of risk factors in metro construction and provide targeted and effective solutions from the perspective of risk correlation and transmission. However, the relationships of risk factors in the interpretative structural model are obtained based on the experience of experts who are highly subjective. The degree of interaction of the risk factors is undetermined, which will be addressed and optimized in a follow-up study.

Acknowledgements
This work was supported by the Ministry of Housing and Urban-Rural Development of China under the project No. R32016036.

References