Journal of Civil Engineering and Management
ISSN 1392-3730 / eISSN 1822-3605
2019 Volume 25 Issue 1: 1–18
https://doi.org/10.3846/jcem.2019.7462

EXPLORING THE PROPERTIES OF COST OVERRUN RISK PROPAGATION NETWORK (CORPN) FOR PROMOTING COST MANAGEMENT

Yun CHEN[1], Zhigen HU[2]*, Quan LIU[3]

[1, 2, 3]School of Water Resources and Hydropower Engineering, Wuhan University, Wuhan, China
[1, 2, 3]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, China

Received 28 June 2018; accepted 29 October 2018

Abstract. Construction cost overrun chronically plagues contractors. To address the issue, numerous studies have explored cost overrun risks (CORs). Nevertheless, their methods of identifying risk relationship are susceptible to experts' experience. In addition, they fail to unearth the relationship structure information and analyze the risk propagation effect. To fill these gaps, this paper intends to propose a methodology that integrates the engineering case analysis and complex network theory, so as to obtain a stable relationship structure and reveal its inherent property. First, 52 CORs and 158 risk paths are extracted from 156 engineering cases, followed by the establishment of a cost overrun risk propagation network (CORPN). Finally, the statistical properties of CORPN are explored. The results indicate that CORPN presents the topological property of heterogeneity. A large number of risk paths can be blocked through preventing the CORs with large total degree, like delay in construction period and engineering quantity increase. Meanwhile, CORPN shows the small-world property. The efficiency of risk propagation can be reduced through preventing the CORs with high betweenness centrality, such as lack of technical skill and experience. These findings contribute to the formulation of beforehand strategies that promote the cost management.

Keywords: cost management, construction cost overrun, cost overrun risks, cost overrun risk propagation network, complex network theory, case analysis.

Introduction

Cost overrun has always been a worldwide problem in the construction industry (Forcada, Rusinol, Macarulla, & Love, 2014; Gharaibeh, 2014; Derakhshanalavijeh & Teixeira, 2017). Flyvbjerg, Bruzelius, and Rothengatter (2003a) found that the cost overrun occurs in 90% of all megaprojects in 20 countries ranging from Europe to Asia. Sovacool, Gilbert, and Nugent (2014) investigated 401 electricity projects built between 1936 and 2014 in Asia, Europe, and North America, and the results showed that the cost overrun was $967 million, on average, per project. This problem generally appears in the owner and the contractor, while the contractor is more vulnerable to the cost escalation. In a world of growing competition between firms, the contractors have to reduce their markups to remain competitive (Akinci & Fischer, 1998). A marginal cost overburden could sweep away their project profit, and even lead to firm bankruptcy (Akinci & Fischer, 1998). In addition, the cost performance of the whole construction industry is heavily reliant on the contractors' performance (Doloi, 2013). Therefore, this paper concentrates on the cost overrun of the contractor. The construction cost overrun occurs as a result of cost overrun risk (COR) (Flyvbjerg, Holm, & Buhl, 2003b). COR is the factor that may cause actual cost outcome to negatively deviate from the original cost estimate (Dikmen, Birgonul, & Han, 2007). Without controlling cost overrun risks (CORs), the contractors will be unable to keep the construction cost within budget (Cheng, 2014). As a result, it has become crucial for the contractors to analyze the CORs and control them, which could effectively prevent cost overrun and promote cost management performance (Akinci & Fischer, 1998).

As limited capital reserve sets an unrealistic capital environment, the contractors cannot control all CORs in the construction (Touran & Suphot, 1997). Key CORs should be identified to help contractors focus on the crucial control points where the best management effect can be

*Corresponding author. E-mail: *zhigenhu@whu.edu.cn*

obtained with less expenditure (Creedy, Skitmore, & Wong, 2010). When analyzing key CORs, the relationships among CORs need to be fully considered, since any cost overrun risk (COR) may induce additional CORs (Fidan, Dikmen, Tanyer, & Birgonul, 2011). Many scholars have performed the exploration on complex relationships among CORs (Iyer & Sagheer, 2010; Eybpoosh, Dikmen, & Birgonul, 2011; Boateng, Chen, Ogunlana, & Ikediashi, 2012; Khanzadi, Eshtehardian, & Esfahani, 2017). However, current researches mainly depend on the expert interview and Delphi, which are susceptible to experts' experience. Moreover, when there are many CORs, experts not only need to specify a large number of relationships, but also to be consulted several times to ensure the consistency of experts' opinions, which consumes a large amount of time. On the other hand, previous studies fail to unearth inherent structure information of risk relationships and analyze the risk propagation effect. It is difficult to find key points that impede the risk propagation, thereby the control measures of risk propagation cannot be provided.

This paper aims to present a methodology of integrating engineering case analysis and complex network theory to address these gaps. Lots of engineering cases about cost overrun will be firstly collected to identify CORs and extract propagation relationships between them. Then a cost overrun risk propagation network (CORPN) is constructed. Finally, complex network theory is applied to analyze the properties of CORPN and find key CORs to prevent the risk propagation. The next section reviews the researches that identify the COR relationships and apply the complex network theory to risk propagation. Section 2 depicts the implementation details of the proposed methodology. Section 3 analyzes the properties of CORPN, and Section 4 discusses the calculation results. Section 5 presents the conclusions, recommendations, and future work.

1. Literature review

With deep comprehension of risk characteristic, risks do not singly exist, but affect each other. Many scholars have considered the relationships between CORs. Iyer and Sagheer (2010) applied the Interpretative Structural Modeling (ISM) to build a hierarchical structure of COR relationships. In order to obtain the risk relationships used in the ISM, several experts were consulted to compare each pair of CORs and judge whether one COR could affect the other. By Delphi method, the experts' opinions on the judgment of mutual influences between CORs reached an agreement after several round interviews. An adjacency matrix was generated to indicate the direct effects of one COR on all other CORs. As the conventional ISM cannot measure the degree of relationship, Tavakolan and Etemadinia (2017) proposed a fuzzy weighted ISM by utilizing fuzzy logic. The improved ISM can calculate how much project risks influence or be influenced by other risks. Eybpoosh et al. (2011) indicated that a network is a better reflection of diverse interactive risks than hierarchical lists in real construction projects. By literature survey, they first got an initial network model consisting of several interactive risk paths. Some industry experts were then requested to further check and identify critical risk paths. Finally, the probability of the relationships between CORs was quantitatively estimated through building the Structural Equation Modeling (SEM). Boateng et al. (2012) analyzed the complex and dynamic feature of the risks of delay and cost overrun, and constructed a System Dynamics (SD) model that describes this feature and risk network relationship. A cause and effect feedback structure of the risks in the SD model was obtained by interviewing the owners, operators, customers, and project managers. Khanzadi et al. (2017) used the Bayesian Network (BN) model to forecast the probabilistic cash flow. The BN model comprehensively considered the identification of risk factors, interactions between the factors, and simultaneous occurrences of the factors. The literature survey and expert survey were applied to specify the risk factors and their relationships with cash flows.

The aforementioned researches made efforts to mitigate the negligence that did not take into account relationships between CORs. However, there mainly exist three deficiencies. First, the methods of identifying the COR relationships are mainly relied on experts' expertise due to a lack of engineering case database of cost overrun. The expert interview and Delphi can make the best use of experts' experience. Nonetheless, the process of identifying risk relationships needs to consume large amounts of time when using these methods. For instance, this paper identifies 52 CORs, for which the experts should specify 2652 ($52 \times 52 - 52 = 2652$) pairwise relationships. In order to greatly reduce the number of risk relationship identifications, Tavakolan and Etemadinia (2017) classified the CORs into several categories and Nasir, McCabe, and Hartono (2003) precluded a part of independent CORs, but there are still hundreds of relationships left to specify. On the other hand, the data gathered by expert interview may be inaccurate and inconsistent (Choudhry, Aslam, Hinze, & Arain, 2014), since the expert's risk perception is susceptible to an individual belief, attitude, judgment, and feeling (Akintoye & MacLeod, 1997). Thus, after one round interview, the experts' opinions may not reach an agreement. The Delphi process has to be carried forward, which requires additional time. Second, previous methods cannot analyze the topology level of relationships between CORs. Except for ISM establishing the hierarchical structure of risks, SEM, SD, and BN involve in constructing the risk network. These methods only analyze relationships in the local view, focusing on point to point, and fail to further unearth the inherent network information that could help managers to understand the mechanism of a problem from the holistic view. Third, previous studies rarely analyze the risk propagation effect in the network. The network composed by risks is one of the most challenging issues in the complex systems because the impact of a risk can easily spread out the whole network (Hu et al., 2016). When one COR occurs, how much will relevant CORs be affected? How fast will it cause the cost overrun of a project? In addition, the important CORs that play important roles in the risk propagation process should be identified. To address these issues is helpful for implementing the most effective risk propagation control.

To fill the first gap, this paper uses the method of engineering case analysis to extract the relationships between CORs, because adequate engineering cases provide almost all information about risks (Eteifa & El-adaway, 2018). Based on a reliable case set, the experts no longer need to identify a lot of risk relationships, which could save much time. On the other hand, objective case analysis overcomes the defect of being susceptible to expert cognition and the inconsistency of group preference during the expert interview. The structure of the risk relationship built by engineering case analysis will not change with experts' perception, which provides a stable research foundation for the analysis of network structure. To fill the second and third gaps, this paper introduces the complex network theory which can deeply reveal the statistical relationships of interconnected units and completely obtain the dynamic characteristic of the whole system. Many scholars have applied the complex network theory to the risk network propagation. Simonsen (2005) began to combine the diffusion with the complex network. Simonsen, Buzna, Peters, Bornholdt, and Helbing (2008) demonstrated that failure risk propagation may weaken the network robustness. Duenas-Osorio and Vemuru (2009) took the power transmission grids as an example and used numerical simulation to obtain the effect of cascading failure on the infrastructure network under the attacks of natural and intentional hazards. Zhang and Yang (2013) built a dynamic risk propagation model and simulated the risk propagation in the R&D network. Chen, Hu, Liu, and Zhao (2018) proposed a Cellular Automata Susceptible Infected Susceptible (CA-SIS) model to simulate the risk propagation of delayed payment in a stakeholder network. Except for risk propagation in the infrastructure network and stakeholder network, Zhou, Irizarry, and Li (2014), Zhou, Xu, Guo, and Ding (2015), and Li et al. (2017) collected a great quantity of cases about safety accident, obtained many accident chains, and established the complex networks of safety accident. The properties of these safety accident networks were explored by using complex network theory. Above studies analyzed the properties of various networks and simulated the risk propagation in the networks. They verified the applicability of complex network theory and provided the research reference for this study. With the assistance of complex network theory, it would be easier to explore the CORPN properties and recognize the key CORs to prevent the risk propagation.

2. Methodology

This section presents the three phases of the proposed methodology, as shown in Figure 1. First, hundreds of historical engineering cases about cost overrun are collected and processed. Second, CORs and risk paths are extracted from the engineering cases to generate a CORPN. The nodes of CORPN represent the CORs and the edges of CORPN indicate the risk paths. At last, complex network theory is applied to analyze the statistical properties and to check the topological properties emerged from real network.

2.1. Data gathering

Historical case statistic provides the essential data used for risk analysis. The risk sources, risk events, risk consequences, and risk-reducing measures can be compre-

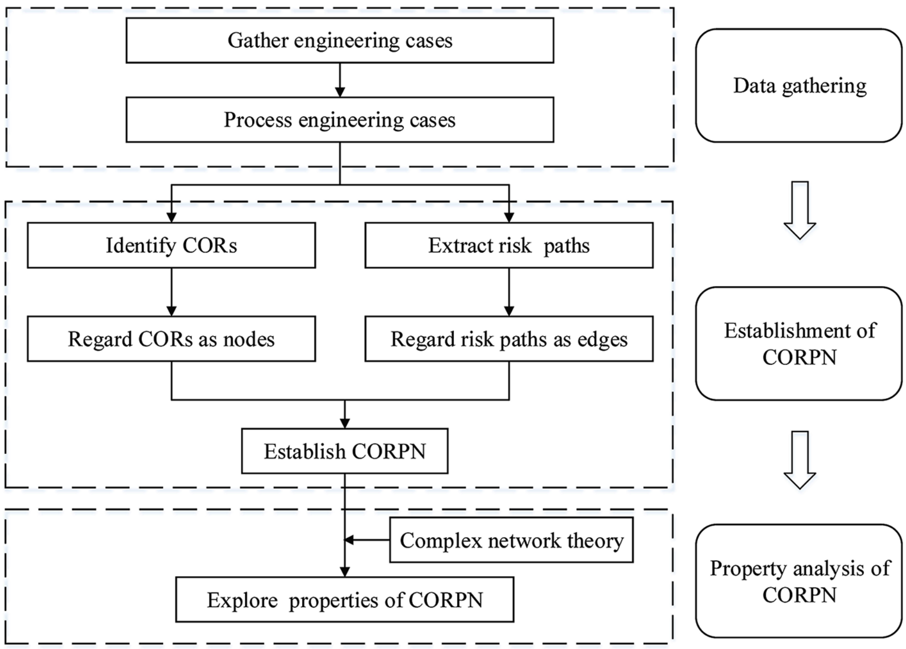

Figure 1. The flowchart of methodology

hensively obtained based on the method. On account of the absence of cost overrun cases collected by official authority departments, this research is based on a book: *Analysis of engineering claims and contract management in hydropower project construction*, which is compiled by Bai (2006). This book records 191 engineering cases regarding cost management of contractors. Some successful cost management cases are excluded by us because they do not belong to cost overrun. A total of 156 engineering cases about cost overrun are left. They happen in the hydropower project construction between 1991 and 2005, and are the most recent available complete cases. Whether these engineering cases can be used as a research basis needs further verification from three important aspects: (1) adequate case quantity, (2) detailed description of and background and occurrence process, and (3) deep reason analysis of cost overrun.

In the aspect of case quantity, until now, there is no standard statistical procedure that stipulates the sample size suitable for establishing a complex network (Eteifa & El-adaway, 2018). Accordingly, past researches using complex network theory in the similar field are reviewed to judge the suitability of sample size collected, as shown in Table 1. In these researches, minimum and maximum sample sizes are 100 cases and 203 cases respectively, and the mean number of sample size is 140 cases. The number of engineering cases collected falls between the minimum and maximum sample sizes and exceeds the mean sample size, meaning that the sample number is adequate. In the aspect of the description of occurrence process, Bai (2006) described the background and occurrence process of each case in detail, including project construction process, accurate time, stakeholder behaviors, detailed losses, etc. These detailed descriptions could help us to understand the reason analysis of cost overrun. In the aspect of reason analysis, as the book claims, to avoid the omission of event causes, Bai (2006) repeatedly discussed the problems and reasons of cost overrun in each case with several managers. The reasons for cost overrun are deeply unearthed from different levels, like underlying reasons, indirect reasons, and direct reasons. Through the above analysis, the 156 engineering cases about cost overrun can be used as this research basis.

Each engineering case includes the detailed description of occurrence process of cost overrun and in-depth reason analysis and assessment. However, these engineering cases are not concise enough, and they need to be further processed. The substantial background information is unnecessary after reason extraction. For the convenience of subsequent steps, key information of engineering cases should be extracted from the process description and reason analysis. The key information includes the case code, type of construction project, type of contract, simple background, and reasons of cost overrun. An engineering case of extracting key information is translated into English and used to present the final result of data collection, as shown in Table 2. Similarly, the same processing is conducted for 156 cases.

2.2. Establishment of CORPN

As can be seen from the reason statement of cost overrun, there is more than one COR in an engineering case. Moreover, the CORs influence each other rather than exist independently. One COR causes other CORs and finally leads to cost overrun, which may contain several risk paths (A→B→cost overrun, A→C→cost overrun, etc.). A CORPN could be a better solution to reflect all possible risk paths. Identification of CORs and extraction of their risk paths are two primary steps of CORPN establishment. After these two steps, CORPN can be visualized.

2.2.1. Identification of CORs and extraction of risk paths

The key information in each case intuitively presents the project type, time, stakeholders, detailed losses, etc. All these information can be directly captured, except for CORs hiding in the reason analysis. The CORs need to be further unearthed. To illustrate, two CORs are identified from the first reason of Case 3.6. The first COR can be inferred from the description: "in order to win the contract": eager to win the contract. The other description:

Table 1. Sample sizes in project management researches using complex network theory

Researchers	Network construction	Sample size
Eteifa and El-adaway (2018)	The causes of fatalities were extracted from 100 case files, and Social Network Analysis (SNA) model was used to construct a fatalities causation network.	100
Zhou et al. (2014)	102 available cases about subway construction accident were used to establish a subway construction accident network.	102
Deng, Song, Zhou, and Liu (2017)	126 typical coal mining accidents were collected to construct a coal mine risk network.	126
Li et al. (2017)	134 accidents happening in metro operation were collected to establish a metro operation hazard network.	134
Huang et al. (2016)	176 data points were analyzed to obtain fluctuation modes between China coal price index and coal mine accidental death, and a directed and weighted network was built, where nodes represent fluctuation modes and edges represent transformations between the fluctuation modes.	176
Zhou et al. (2015)	203 reports about railway accidents in UK were analyzed to establish directed weighted accident causation network.	203

Table 2. Key information of Case 3.6

Case code	Case 3.6
Type of construction project	Rock-fill dam
Type of contract	Low price contract
Simple background	The budget of the rock-fill dam was about RMB 88 million yuan, while the actual contract price was only RMB 35.93 million yuan.
Reasons of cost overrun	1. In order to win the contract, a contractor obtained a project at contract price below cost, and too low contract price resulted in that cost overrun had occurred before project construction. 2. What's worse, the construction site of the project was in bad weather, and incomplete hydrological data was provided by the designer. In September 2000, the project suffered a catastrophic flood, resulting in economic loss about RMB 3 million yuan. 3. Additionally, because of lack of construction experience, the contractor did not build the sand stone processing system for cushion material, and location of the temporary construction site was not reasonable. In the actual construction process, the temporary construction site had to be greatly increased, and the sand stone processing system was required to be built, leading to the too low original estimate.

Table 3. Example of identifying CORs and extracting risk paths from Case 3.6

No.	Reasons of cost overrun	CORs	Risk paths
Case 3.6	1. In order to win the contract, a contractor obtained a project at contract price below cost, and too low contract price resulted in that cost overrun had occurred before project construction.	Eager to win the contract; Low quoted price	Eager to win the contract → Low quoted price → Lost
	2. What's worse, the construction site of the project was in bad weather, and incomplete hydrological data was provided by the designer. In September 2000, the project suffered a catastrophic flood, resulting in economic loss about RMB 3 million yuan.	Weather effects; Incomplete design; Natural disaster	Weather effects & Incomplete design → Natural disaster → Lost
	3. Additionally, because of lack of construction experience, the contractor did not build the sand stone processing system for cushion material, and location of the temporary construction site was not reasonable. In the actual construction process, the temporary construction site had to be greatly increased, and the sand stone processing system was required to be built, leading to the too low original estimate.	Lack of technical skill and experience; Engineering quantity increase; Low quoted price	Lack of technical skill and experience → Engineering quantity increase → Low quoted price → Lost

Note: "&" means that next CORs occur only if several CORs simultaneously happen.

"contractor obtained a project at contract price below project cost" shows the second COR: low quoted price. In the aspect of extraction of risk paths, the first reason analysis of Case 3.6 is also taken as an example to explain the extraction process. According to the description, the "eager to win the contract" causes the "low quoted price", and the "low quoted price" leads to cost overrun eventually. There exist causal relationships between the eager to win the contract and the low quoted price, the low quoted price and cost overrun. These relationships form a risk path: eager to win the contract → low quoted price → cost overrun. The CORs and risk paths in other two reasons of Case 3.6 are extracted by the above method, as shown in Table 3.

To avoid isolated CORs in the risk paths, last COR in each path is set as "lost" (cost overrun), and it is regarded as the consequence of COR. Additionally, some CORs are induced only if their previous CORs simultaneously happen. In the second reason of Case 3.6, if the contractor had obtained accurate design information and safe construction plan had been designed, the economic loss might have been reduced even though the project encountered a catastrophic flood. Thus, only when the weather effect and incomplete design occur at the same time can natural disaster result in cost overrun.

Similarly, the reasons of cost overrun of 156 cases are analyzed to identify all CORs and corresponding risk paths. After that, the authors discussed the preliminary list of CORs with some experts in cost management. Some CORs carrying similar meanings are grouped and renamed, which removes the redundancy and reduces the complexity of CORPN. For example, variation in market, variation in material cost, high fluctuation in labor cost, and price fluctuation are unified as "price increase". The original CORs in corresponding risk paths are replaced by unified CORs at the same time. A total of 52 CORs are identified, as shown in Table 4, and a total of 158 risk paths without repetition are obtained and numbered in Appendix.

2.2.2. Visualization of CORPN

There are 158 risk paths that interweave each other. It is difficult to form a network by merging these paths one by one. An adjacency matrix is a mathematical means of showing which nodes of a network impact one another

Table 4. All CORs identified by engineering case analysis

Abb.	CORs	Abb.	CORs
R1	No permit/approval	R27	Insufficient survey of sources of funds
R2	Political ferment	R28	Change in contract
R3	Unstable condition of host country	R29	Contract failure
R4	Legislations and regulations change	R30	Lack of understanding of contract terms
R5	Price increase	R31	Ambiguity in contract clauses
R6	Project administration cost increase	R32	Construction accident
R7	Inadequate forecasting of market	R33	Error in construction
R8	Relocation problem	R34	Idle worker
R9	Resettlement problem	R35	Idle machine
R10	Delay in land acquisition	R36	Engineering quantity increase
R11	Lack of material or unqualified material	R37	Low quoted price
R12	Poor management ability of owner or supervision	R38	Insufficient site investigation
R13	Fail to provide construction site	R39	Lack of technical skill and experience
R14	Delay of owner	R40	Misunderstand drawing
R15	Delay in construction period	R41	Low management competency
R16	Project acceleration cost	R42	Lack of claims
R17	Design scope change	R43	Not buying insurance
R18	Delay in payment	R44	Poor quality
R19	Pay for project in advance	R45	Reworks
R20	Fraudulent practices	R46	Problem of cash flow
R21	Error in subcontractor	R47	Change in construction
R22	Error in detail design	R48	Bad construction environment
R23	Delay in design	R49	Weather effects
R24	Incomplete design	R50	Unknown geological conditions
R25	Poor communication and coordination	R51	Strike by workers
R26	Eager to win the contract	R52	Natural disaster

(Wambeke, Liu, & Hsiang, 2012), and thus it is used to establish the CORPN. The first column and first row of the adjacency matrix are occupied by all CORs. The off-diagonal cells of the adjacency matrix stand for causal relationships between two CORs. If one COR causes the other, the value of the corresponding cell is equal to 1, otherwise, the value is 0. All of the diagonal cells are equal to 0 because there is no relationship between COR itself. After building the adjacency matrix, a network diagram can be drawn, where nodes represent the CORs and directed edges represent causal relationships between CORs.

The formative process of a part of CORPN based on path 4, 5, and 14 in Appendix is used to illustrate the visualization of CORPN, as shown in Figure 2.

First, an adjacent matrix is generated from the three risk paths. For instance, unknown geological condition (R50) can directly lead to the delay in construction period (R15) and design scope change (R17), namely, (R50, R15) and (R50, R17) are equal to 1. Unknown geological condition (R50) cannot cause the project acceleration cost (R16) and lost (L) directly, meaning that (R50, R16) and (R50, L) are equal to 0. Then, the network is visualized based on the adjacent matrix. In the visualization of CORPN, tiny circles represent CORs. Directed straight-line segments represent the causal relationships between the CORs. Additionally, the length of the straight-line has no significance. Ultimately, imitating the formative process, a complete CORPN is generated by synthesizing all risk paths. The complete adjacent matrix is input into the software *UCINET* 6 for *Windows Version* 6.212, and *NetDraw* 2.084 is utilized to draw CORPN. The complete CORPN consists of 53 nodes and 238 edges, as visualized in Figure 3. The complete adjacent matrix is provided in Supplement Material.

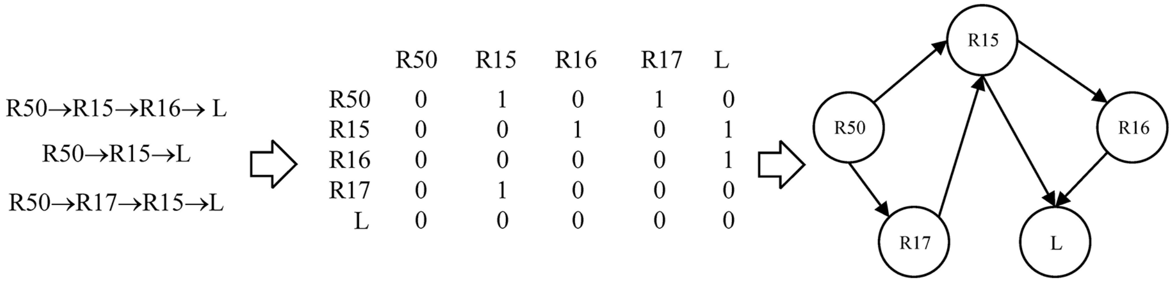

Figure 2. Formative process of a part of cost overrun risk propagation network (CORPN)

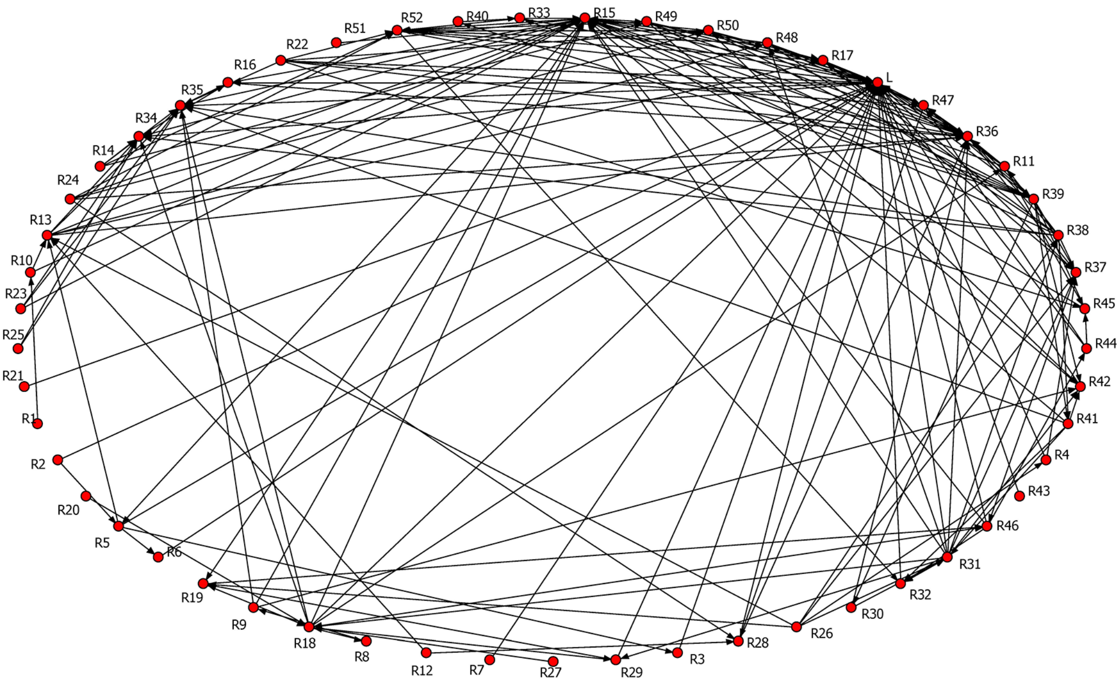

Figure 3. The visualization of complete cost overrun risk propagation network (CORPN)

2.3. Statistical properties of CORPN

From Figure 3, CORPN is too complex to intuitively obtain the network structure information. Complex network theory can capture the dynamic feature of components of network and contribute to point out critical nodes in the network. This theory is used to explore several representative statistical properties of CORPN, including node degree, degree distribution, average path and diameter, clustering coefficient, betweenness centrality and global efficiency.

Let $G = (V, E)$ be CORPN, where $V = (v_1, v_2, \ldots, v_n)$ is the set of nodes, standing for the CORs, and $E = (e_1, e_2, \ldots, e_n)$ is the set of edges, representing their relationships. The adjacent matrix of CORPN is shown as $\boldsymbol{A} = [a_{ij}]_{N\times N}$, where $a_{ij} = 1$ if ith node is connected with jth node, otherwise, $a_{ij} = 0$, and N is the number of the nodes.

1. Node degree and degree distribution

ith node's degree is the number of edges directly connected to the node. As CORPN belongs to directed network, the node's degree has three components: in-degree, out-degree and total degree (Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006). ith node's in-degree is the number of ith node's ingoing edges. ith node's out-degree is the number of ith node's outgoing edges. ith node's total degree is the sum of its in-degree and out-degree. They are calculated as follows:

$$\deg_i^{in} = \sum_i a_{ij},\ \deg_i^{out} = \sum_j a_{ij},\ \deg_i = \deg_i^{in} + \deg_i^{out}, \quad (1)$$

where: $\deg_i$ is total degree; $\deg_i^{in}$ represents in-degree; $\deg_i^{out}$ indicates out-degree.

Average degree is equal to total degree divided by total number of nodes. Degree distribution P(k) is defined as the probability that a node with the value of node's degree k is randomly selected.

2. Average path length and diameter

There may be more than one path between two nodes. The shortest path between ith node and jth node is the smallest number of edges that connect the two nodes in a network, which is denoted as $\min_{i,j} d_{ij}$. Average path length L is the average value of the shortest paths for all possible pairs of nodes (Boccaletti et al., 2006), which is shown as follows:

$$L = \frac{1}{N(N-1)} \sum_{i,j\in N, i\neq j} \min_{i,j} d_{ij}, \quad (2)$$

where: d_{ij} represents the path between ith node and jth node, which is equal to the number of edges that connect ith node and jth node; $\min_{i,j} d_{ij}$ is calculated by Dijkstra algorithm.

Diameter is defined as the longest of all calculated shortest paths in a network, which is denoted as $\max_{i,j} d_{ij}$.

3. Clustering coefficient

Clustering coefficient describes the property of a node regarding formation of cliques among its neighbors and measures which nodes tend to cluster together (Tabak, Takami, Rocha, Cajueiro, & Souza, 2014). Clustering coefficient of ith node is the proportion of edges between the nodes within its neighborhood divided by the number of edges that could possibly exist between them (Watts & Strogatz, 1998). In a directed network, a_{ij} is distinct from a_{ji}, and therefore there are at most $\deg_i(\deg_i - 1)$ edges that can exist between neighbors of ith node. The clustering coefficient of *i*th node is represented as C_i, which is calculated as follows:

$$C_i = \frac{l_i}{\deg_i(\deg_i - 1)}, \tag{3}$$

where l_i is the number of directly connected neighbors of *i*th node.

4. Betweenness centrality

Betweenness centrality includes node betweenness centrality and edge betweenness centrality. Because of similarity between them, only node betweenness centrality is analyzed in this study. Node betweenness centrality is the proportion of shortest paths through a node to all shortest paths between all possible pairs of nodes (Huang, Zhuang, Yao, & Uryasev, 2016). The betweenness centrality of *i*th node in the directed network is expressed as B_i and is calculated as follows (Gonzalez, Dalsgaard, & Olesen, 2010):

$$B_i = \frac{1}{(N-1)(N-2)} \sum_{s,t\in N, s\neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}, \tag{4}$$

where: *s*th node and *t*th node are two non-adjacent nodes; σ_{st} indicates total number of shortest paths from *s*th node to *t*th node; $\sigma_{st}(i)$ is the number of these paths that contains *i*th node.

5. Global efficiency

Efficiency has global efficiency and local efficiency, respectively describing the information propagation efficiency on a global and on a local network scale. Only global efficiency $E_{glob}(G)$ is considered to measure the risk propagation efficiency in CORPN, which is calculated as follows (Latora & Marchiori, 2001):

$$E_{glob}(G) = \frac{1}{N(N-1)} \sum_{i,j\in N, i\neq j} \frac{1}{\min_{i,j} d_{ij}}. \tag{5}$$

3. Results

The network property includes the statistical property and topological property (Jin, Zhang, & Li, 2016). For the sake of further unearthing the properties of CORPN, based on complex network theory in the previous section, this section explores typical statistical properties of CORPN and checks significant topological properties emerged from real network (Albert & Barabasi, 2002; Newman, 2003). The typical statistical properties are as follows: node degree (including in-degree, out-degree, and total degree), degree distribution, average path length, diameter, clustering coefficient, and betweenness centrality.

3.1. Node degree

The node degree representing the importance of a node is the most simple and important property of the node. The node degree in the directed network includes the in-degree, out-degree, and total degree. The larger total degree a node has, the more important the node is. A node with large in-degree is susceptible to other nodes. A node with large out-degree easily affects other nodes. The in-degree, out-degree and total degree of CORPN are calculated by Eqn (1) and shown as radar chart in Figure 4. R36 (engineering quantity increase) and R15 (delay in construction period) have higher in-degree with the value of 15 and 22 respectively (excluding L with the value of 32). That is to say, R36 and R15 are directly affected by 15 CORs and 22 CORs respectively. Noting that total degrees of the two nodes also rank in the top two. R39 (lack of technical skill and experience) has the highest out-degree with the value of 11, followed by R31 (ambiguity in contract clauses) with the value of 10, indicating that R39 and R31 could directly cause 11 and 10 risks respectively. The value of the average degree of CORPN is 6.9057, that is to say, each COR is averagely connected to 7 CORs.

3.2. Degree distribution

According to the total degree of all nodes, the number of each node's degree is counted, as shown in Table 5. The values of nodes' degrees below 10 (including 10) account for 73.58% and the values of nodes' degrees between 20 (including 20) and 30 (including 30) only account for 3.78%. Most of nodes' degrees are relatively small, but a few nodes' degrees are large, which presents the heterogeneity of degree distribution of CORPN.

The cumulative degree distribution of CORPN P(k), the fraction of nodes with degree greater than or equal to k, is measured to quantify the heterogeneity property of CORPN, because it can be used to reduce the noise in tail part of degree distribution (Chen, Yang, & Xu, 2012). After processed by MATLAB curve fitting toolbox, the cumulative degree distribution of CORPN is expressed in the single logarithmic coordinate system and general coordinate system, as shown in Figure 5. It is observed clearly that the cumulative degree distribution function of CORPN obeys an exponential distribution function $y = 1.2e^{-0.1582x}$ ($R^2 = 0.9927$), and imitative effect is good.

According to previous researches in the literature review (Zhou et al., 2014; Zhou et al., 2015; Li et al., 2017), the network composed by risk factors generally belongs

Table 5. The occurrence frequency of each node's degree

Node's degree	1	2	3	4	5	6	7	8	9
Number	7	7	4	4	6	3	5	1	2
Frequency /%	13.20	13.20	7.55	7.55	11.32	5.66	9.43	1.89	3.77
Node's degree	10	11	12	13	14	20	30	32	
Number	5	1	1	3	1	1	1	1	
Frequency /%	9.43	1.89	1.89	5.66	1.89	1.89	1.89	1.89	

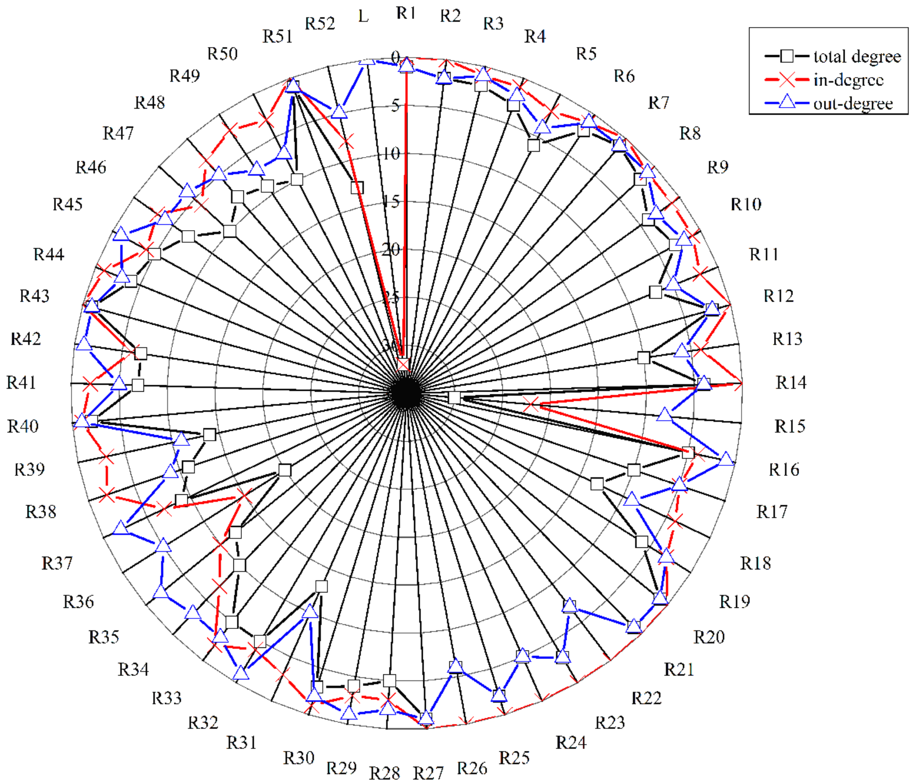

Figure 4. In-degree, out-degree and total degree of COR

to the scale-free network obeying power-law distribution. When the cumulative degree distribution of CORPN is fitted into the power-law distribution, the fitting function is $y = 1.169x^{-0.6181}$ ($R^2 = 0.8551$), as shown in Figure 6. Its power-law exponent is 1.6181 = 0.6181+1. However, the power-law distribution of scale-free network satisfies power-law exponent varying from 2 to 3 (Barabasi & Albert, 1999), and therefore CORPN deviates from the nature of the scale-free network. Compared with the fitting function of power-law distribution, the cumulative degree distribution of CORPN is more suitable for exponential distribution function $y = 1.2e^{-0.1582x}$ ($R^2 = 0.9927$). The properties of some real networks are consistent with the exponential networks, such as power grid (Liu & Tang, 2005). Additionally, Li and Chen (2003) found that a network's heterogeneity increases when the cumulative degree distribution of the network is developed from exponential distribution to power-law distribution. This indicates that CORPN displays weaker heterogeneity than the scale-free network. By comparing two fitting curves of CORPN, the slope of the

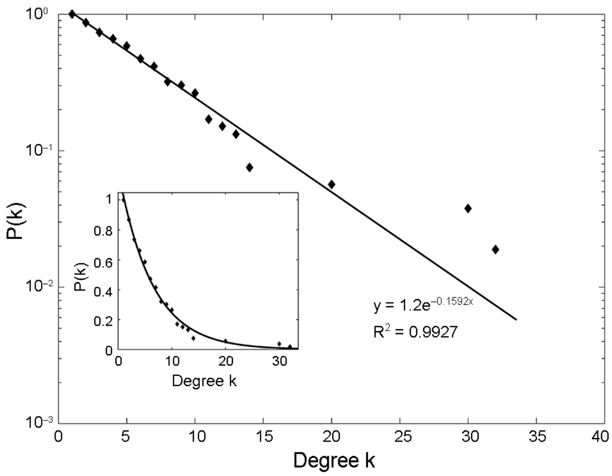

Figure 5. Cumulative degree distribution fitting exponential distribution is expressed in the single logarithmic coordinate system. The inset figure is the cumulative degree distribution fitting exponential distribution in the general coordinate system

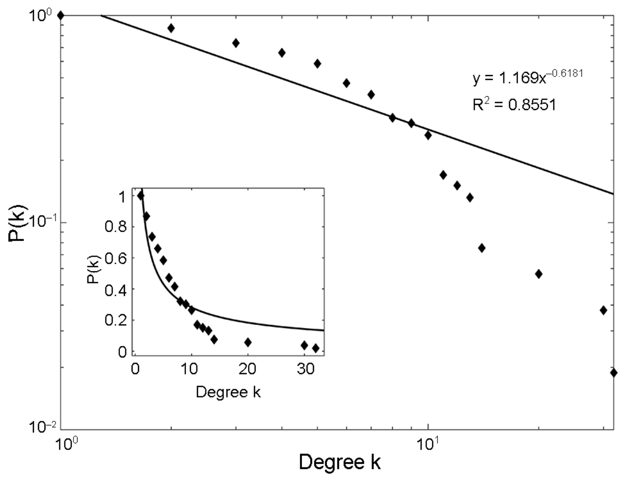

Figure 6. Cumulative degree distribution fitting power-law distribution is expressed in double logarithmic coordinate system. The inset figure is the cumulative degree distribution fitting power-law distribution in general coordinate system

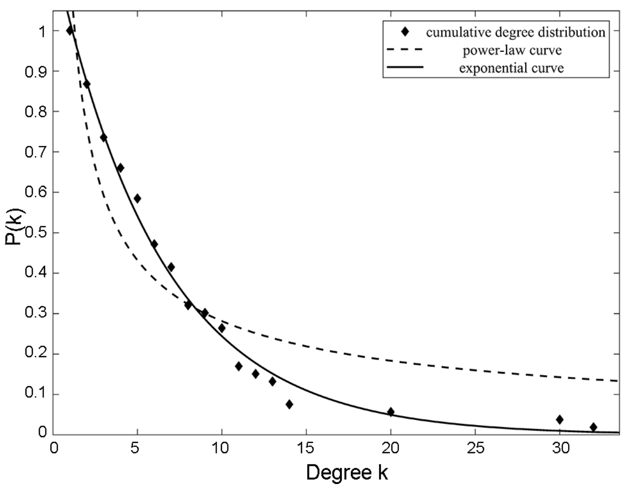

Figure 7. Comparison between power-law curve and exponential curve

power-law curve is steeper than the exponential curve in Figure 7, which further proves this conclusion.

3.3. Average path length and diameter

According to Eqn (2), the value of the average path length of CORPN is 3.116. That is to say, each COR propagates its negative effect to other CORs only through three steps on average. For instance, in the No. 20 path, R50 (unknown geological condition) gives rise to R52 (natural disaster). R52 leads to R15 (delay in construction period). R15 causes R5 (price increase). R50 has no direct relation with R5, while unknown geological condition can trigger the price increase by only three steps. The diameter of CORPN is 8. CORPN has more than one diameter. R2 (political ferment) to R38 (insufficient site investigation) is a pair of nodes with the longest distance in CORPN, namely, R2 (political ferment) →R5 (price increase) →R13 (fail to provide construction site) →R36 (engineering quantity increase) →R47 (change in construction) →R39 (lack of technical skill and experience) →R30 (lack of understanding of contract terms) →R31 (ambiguity in contract clauses) →R38 (insufficient site investigation). The political ferment may lead to the insufficient site investigation, which is likely to be ignored. The hidden risk relationships, like R2 → … →R38, reduce the effectiveness of risk control. In view of the two kinds of characteristic path length, although the number of CORPN's nodes is more than 50, the longest distance in CORPN is 8 steps and the average path length of CORPN is 3 steps, indicating that the size of CORPN is small.

3.4. Clustering coefficient

The values of the clustering coefficient of all nodes are calculated by Eqn (3) and shown in Figure 8. R1 (no permit/approval), R7 (inadequate forecasting of market), R20 (fraudulent practices), R21 (error in subcontractor), R27 (insufficient survey of sources of funds), R43 (not buying insurance), and R51 (strike by workers) get missing the clustering property because these nodes' degrees are equal to 1, and there is only one neighbor around them. The clustering coefficients of rest nodes fluctuate between 0 and 0.5, wherein R2 (political ferment), R3 (unstable condition of host country), R6 (project administration cost increase), R23 (delay in design), and R40 (misunderstand drawing) have the highest value 0.5. Because of no connection between adjacent nodes, the clustering coefficients of R8 (relocation costs), R12 (poor management ability of owner or supervision), and R25 (poor communication and coordination) are equal to 0. Additionally, it is found that the nodes with high total degree do not necessarily have high clustering coefficients. This is because they connect with many other nodes, many of which are linked with the same neighbor. This weakens the clustering property

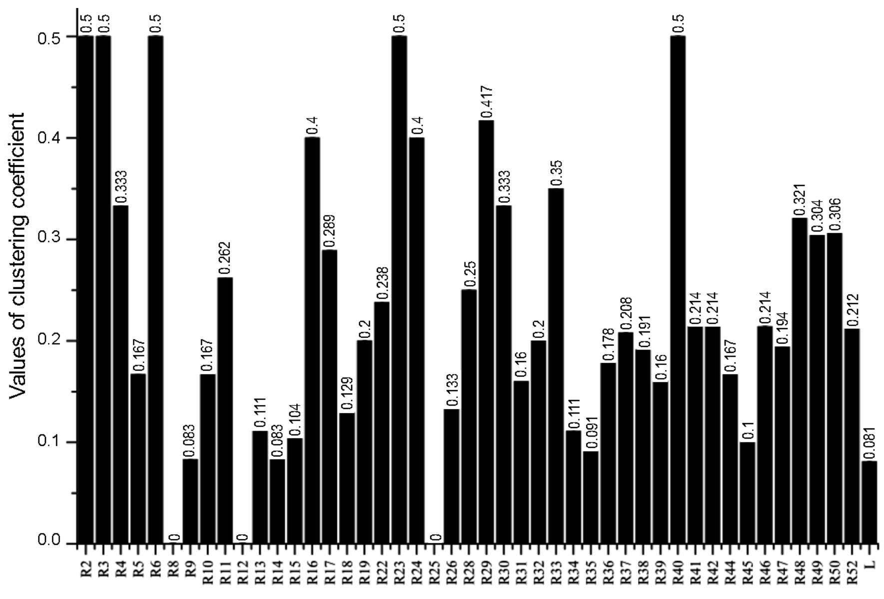

Figure 8. Clustering coefficient of each COR

of these nodes among their neighbors. Although the nodes with low degree are only connected to a few nodes, they are tightly interrelated, resulting in that their clustering coefficients are higher than the nodes with high total degree.

Small-world property indicates that most of the nodes can indirectly connect other nodes by a few steps. Small-world network has relatively small average path length and high clustering coefficient (Watts & Strogatz, 1998). The clustering coefficient of CORPN is 0.23, which is equal to the average value of clustering coefficient of all nodes. It is much greater than a random network having the same node set. The average path length of CORPN is 3.116. Therefore, the small-world property of CORPN can be deduced, meaning that most of the CORs can be affected by every other COR through a small number of intermediate CORs. The speed of risk propagation in CORPN will be extremely quick. It is hard to avoid cascading effects and also explains why the cost overrun becomes a common problem.

3.5. Betweenness centrality

According to the definition of betweenness centrality, the higher betweenness centrality of a node has the more risk paths will pass through it. The values of betweenness centrality of all nodes are calculated by Eqn (4). Because of the computational complexity of the algorithm, the calculation of betweenness centrality relies on *UCINET* software. The results are presented as a radar chart in Figure 9. There are 20 nodes whose betweenness centrality is equal to 0 because they do not serve as the party of intermediary in the interaction between other nodes. R15 (delay in construction period) has the highest betweenness centrality with the value of 0.2858, indicating that it has the largest number of shortest paths through it. That is followed by R39 (lack of technical skill and experience) with the value of 0.1197. The sum of the betweenness centrality of R15 and R39 is more than 0.4, namely, almost 40% of shortest paths pass through these two nodes. The values of betweenness centrality of other nodes are less than 0.1.

4. Discussion

Prevention of CORs could help contractor to achieve precise cost control. However, dozens of CORs are identified from the engineering cases. It is hard to control all CORs in the risk management. The identification of key CORs not only maintains the management expense, but also achieves the goal of boosting the cost management performance. In the identification of key CORs, the CORs' relationships and interactions should be considered simultaneously. Therefore, this paper establishes CORPN through engineering case analysis, measures the statistical properties of CORPN, and analyzes the COR propagation feature in CORPN. The results indicate two typical topological properties: heterogeneity and small-world network. Combined with the statistical properties, some key CORs are specified, which help managers to set up important monitoring points for cost overrun prevention in advance.

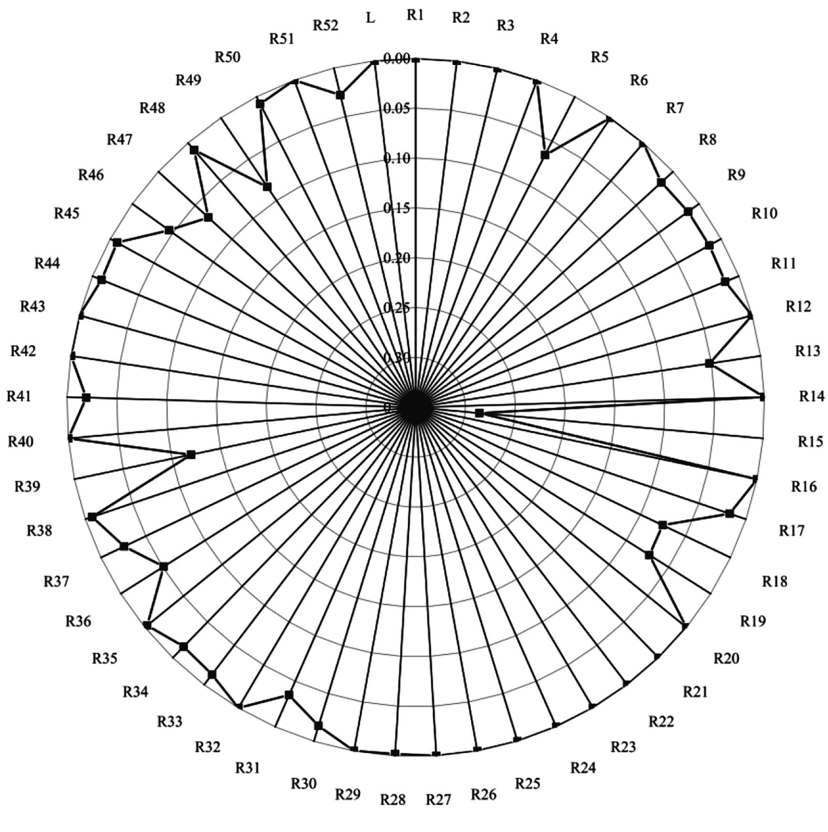

Figure 9. Betweenness centrality of each COR

4.1. Identification of key CORs from CORPN properties

The relationships between CORs and risk propagation feature bring the great challenge to the identification of CORs. If contractors consider CORs independently, the cost overrun may still occur, since CORs controlled may be not on the path of risk propagation. Not only consumes the management cost but also poor control effect will be obtained. Prevention of CORs means breaking the edges of CORPN. The CORs maintaining structural stability of CORPN should be highly valued. Once these CORs are prevented, the important risk paths in the risk propagation can be blocked. Minimum risk management expenditure could achieve the best control result. The heterogeneity of CORPN indicates that a few CORs are related to many other CORs, and a large number of CORs are poorly related. When R36 (engineering quantity increase) and R15 (delay in construction period) with large total degree are controlled, a large number of risk paths could be obstructed. It is beneficial to influence a lot of relationships between CORs and avoid triggering other CORs, which effectively reduces the range of risk propagation. For instance, if R15 having total degree with the value of 30 is controlled, 30 relationships between CORs can be affected, which accounts for 12.6% of all relationships. This is the result when only one COR is controlled.

Because of high clustering coefficient and short average path length, CORPN is a small-world network, which is characterized by the property that every node is related to almost every other node through a short relationship (Small & Tse, 2005). According to the research of Zhou et al. (2014), in general, moving nodes with high betweenness centrality could increase the average path length and diameter of the network to slow down the efficiency of risk propagation. The values of betweenness centrality of R15 (delay in construction period) and R39 (lack of technical skill and experience) rank in the top two. Nevertheless, assuming that R15 and R39 are moved out of CORPN, the average path length of CORPN falls to 2.442 from 3.116 and diameter of CORPN reduces to 7 from 8. This is, of course, wrong because R51 is isolated. In the unconnected network, average path length and diameter cannot be used to measure the propagation efficiency in CORPN (Crucitti, Latora, Marchiori, & Rapisarda, 2003). The global efficiency takes into account the isolated nodes and is a better measure to describe the propagation efficiency of a network. After removing R15 and R39, the global efficiency of CORPN falls to 0.1047 from 0.2240, based on Eqn (5). It indicates that prevention of these two CORs indeed slows down the efficiency of risk propagation. The No. 5 and No. 15 risk paths in Appendix are taken as examples to explain the efficiency change of risk propagation. In No. 5 risk paths (unknown geological condition (R50) → delay in construction period (R15) → lost (L)), unknown geological condition (R50) leads to the cost overrun by inducing the delay in construction period (R15). If the delay in construction period (R15) is avoided through accelerating construction, unknown geological condition (R50) leads to the cost overrun only by No. 15 risk path (unknown geological condition (R50) → design scope change (R17) → engineering quantity increase (R36) → lost (L)). The risk propagation path gets longer. The cost overrun may occur only when R50 causes R17, and R17 triggers R36 at the same time, which needs more demanding condition and dramatically declines the occurrence possibility of cost overrun.

In summary, compared with other CORs, prevention of the delay in construction period (R15), engineering quantity increase (R36), and lack of technical skill and experience (R39) could control risk propagation more effective, particularly R15 (delay in construction period) and R39 (lack of technical skill and experience). R15 simultaneously has large in-degree and total degree, and high betweenness centrality. R39 has large out-degree and high betweenness centrality.

4.2. Analysis of key CORs

In the aspect of the identification of key CORs, some researchers concluded similar results. For the delay in construction period (R15), as Shane, Molenaar, Anderson, and Schexnayder (2009) claimed, inflation rate and project schedule extension are two main components that could cause unanticipated cost increase in terms of time value of money. This is consistent with R15 (delay in construction period) having large total degree. Furthermore, Iyer & Sagheer (2010) found that time overrun has the highest dependence on other project risks, indicating that the construction schedule is susceptible to other factors. This is consistent with R15 (delay in construction period) having large in-degree. On the one hand, there are many contractors' own negative causes. Most commonly, unreasonable or lack of construction planning leads to disorganized site management, such as unbalanced allocation of labor and inappropriate use of mechanical equipment (Larsen, Shen, Lindhard, & Brunoe, 2016). These contractor factors seriously affect the construction efficiency; thereby the delay in construction period occurs. On the other hand, construction schedule also suffers from many negative factors caused by non-contractor, like bad natural condition, delay in transportation of materials, delayed payment, design change, etc. These uncontrollable risks bring the severe challenges for the contractor to complete the project within the specified construction period. Besides being susceptible, the delay in construction period could induce many other negative consequences related to cost overrun, which is ignored in prior studies. Once the construction schedule is disrupted, construction means, worker arrangement, and machinery deployment need to be readjusted to accelerate the construction progress.

The other key COR is R39 (lack of technical skill and experience) which has large out-degree and easily triggers other CORs. Similarly, Cheng (2014), Wright, Cho, and Hastak (2014) indicated that the lack of technical skill and experience has strong cost-influencing feature. In other words, this COR may produce many negative effects on cost management. This is because the construction of the entire project depends on various staffs. The lack of technical skill and experience means the lack of staff that has vast experience in construction management and technique. Before signing a contract, staffs without experience in contract negotiation may sign some unfavorable contract clauses, which will bring many difficulties in the construction cost control. Meanwhile, new and inexperienced staffs conducting cost estimate will result in poor estimating, particularly, underestimation (Shane et al., 2009). During construction, staffs with less construction experience are incapable of dealing with various uncertain technical challenges and problems, and they may make some construction errors, which could necessitate the rework (Kartam, 1996). Thus, it is difficult to economize the construction cost without professional staffs. Conversely, appropriate allocation of resource, their optimized utilization, and effective cash flow management by the experienced on-site contractor have significant influence on construction cost (Doloi, 2013).

The results of this paper show that R36 (engineering quantity increase) has large in-degree and easily occurs. Similarly, Eybpoosh et al. (2011) concluded that engineering quantity increase is at the high hierarchy of risk-path model of cost overrun, meaning that engineering quantity is a sensitive factor. This is because all engineering cases come from hydropower project construction which is characterized by various types of construction and complex construction environment. Hydropower project includes traffic engineering, dam, spillway, releasing tunnel, powerhouse, etc., and its construction period is generally 5 to 10 years, even 10 to 20 years. Over that period, it is hard to ensure that the project scope and design are unmodified (Shane et al., 2009). When new problems arise, various design changes will follow. After design changes, the contractor has to adjust original construction scheme, which may result in the increase of labor, material and management expenditure. Although the contractor could reduce economic loss by claim, the possibility of cost overrun increases. In addition, hydropower projects in China are generally located in the southwest region where geological conditions are extremely complex. For example, a fault zone easily causes the over break or overfilling in the construction. Because of the interference of the above factors, it is hard to accurately estimate engineering quantity.

Conclusions

The cost overrun has been a common problem in the construction industry. The analysis of CORs could allow contractors to pay attention to the crucial points of preventing cost overrun, so as to promote the cost management. The contributions of this study mainly include three aspects. First, the method of engineering case analysis effectively avoids the defects of expert interview and Delphi when identifying risk relationship. The engineering case analysis is used to identify 52 CORs and extract 158 risk paths. Second, a CORPN is established to present the complex relationships between CORs. Based on complex network theory, several typical statistical properties of CORPN are explored. The results show that CORPN has two significant topological properties: heterogeneity and small-world network. Third, how to prevent the risk propagation in CORPN is considered. The analysis results of the network properties help to identify three key CORs: delay in construction period (R15), engineering quantity increase (R36), and lack of technical skill and experience (R39). Controlling them can effectively prevent risk propagation to avoid the cost overrun.

CORPN presents the heterogeneity because the cumulative degree distribution of CORPN obeys the exponential distribution $y = 1.2e^{-0.1582x}$ ($R^2 = 0.9927$). If R15 (delay in construction period) and R36 (engineering quantity increase) with large total degree are prevented, a large number of relationships between CORs can be disrupted. It is thus necessary to arrange a reasonable construction scheme or plan to improve construction efficiency. For construction schedule delay caused by non-contractor, the contractor should claim for the owner and seek to extend the construction period. In order to prevent the engineering quantity increase, the contractor should make sufficient preparations including detailed site investigation before construction, repeated demonstration of design and construction scheme, careful study of contract documents, and reasonable claim, which could control the engineering quantity variation into an expected scope.

CORPN is a small-world network for high clustering coefficient and short average path length, meaning that CORs would propagate very quickly in CORPN. Dampening R39 (lack of technical skill and experience) with high betweenness centrality could slow down the efficiency of risk propagation. Therefore, the managers of the contractor are required to have rich experience in cost estimation and contract management, which could help contractors to accurately estimate budget and sign a favorable contract. The engineers with rich construction experience should be arranged to manage the project, so as to deal with various key technical challenges and problems.

There are some limitations in this study. First, how to integrate the dozens of CORs to reduce the complexity of CORPN is the next phase of research. Second, the engineering cases of this study originate from the hydropower project construction. Despite a hydropower project covers a wide range of project types, like housing, bridge, road, etc., the number of engineering cases and project types should be further enriched.

Acknowledgements

Thanks to Associate Professor Shu Chen and PhD Candidate Bo Shao for review and comment during the writing process. Thanks to the support received from the National Natural Science Foundation of China (Grant no. 51779195).

We would like to express our sincere gratitude to the reviewers of this paper for their constructive comments.

Funding

This work was supported by the "National Natural Science Foundation of China" under Grant [number 51779195].

Author contributions

Yun CHEN, Zhigen HU and Quan LIU conceived the study. Yun CHEN and Quan LIU were responsible for data collection, processing, and analysis. Yun CHEN and Zhigen HU wrote the paper.

Disclosure statement

The authors declare that they do not have any competing financial, professional, or personal interests from other parties.

References

Akinci, B., & Fischer, M. (1998). Factors affecting contractors' risk of cost overburden. *Journal of Management in Engineering*, *14*(1), 67-76. https://doi.org/10.1061/(asce)0742-597x(1998)14:1(67)

Akintoye, A. S., & MacLeod, M. J. (1997). Risk analysis and management in construction. *International Journal of Project Management*, *15*(1), 31-38. https://doi.org/10.1016/S0263-7863(96)00035-X

Albert, R., & Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47-97. https://doi.org/10.1103/RevModPhys.74.47

Bai, J. S. (2006). *Analysis of engineering claims and contract management in hydropower project construction*. Beijing: China Water & Power Press (in Chinese).

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509-512. https://doi.org/10.1126/science.286.5439.509

Boateng, P., Chen, Z., Ogunlana, S., & Ikediashi, D. (2012). A system dynamics approach to risks description in megaprojects development. *Organization, Technology and Management in Construction*, *4*(3), 593-603. https://doi.org/10.5592/otmcj.2012.3.3

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports – Review Section of Physics Letters*, *424*(4-5), 175-308. https://doi.org/10.1016/j.physrep.2005.10.009

Chen, G., Yang, X. H., & Xu, X. L. (2012). Weighted scaling in non-growth random networks. *Communications in Theoretical Physics*, *58*(3), 456-462. https://doi.org/10.1088/0253-6102/58/3/24

Chen, Y., Hu, Z. G., Liu, Q., & Zhao, M. Y. (2018). Risk propagation of delayed payment in stakeholder network of large hydropower project construction considering risk resistance and mitigation. *Mathematical Problems in Engineering*, 8013207. https://doi.org/10.1155/2018/8013207

Cheng, Y. M. (2014). An exploration into cost-influencing factors on construction projects. *International Journal of Project Management*, *32*(5), 850-860. https://doi.org/10.1016/j.ijproman.2013.10.003

Choudhry, R. M., Aslam, M. A., Hinze, J. W., & Arain, F. M. (2014). Cost and schedule risk analysis of bridge construction in Pakistan: Establishing risk guidelines. *Journal of Construction Engineering and Management*, *140*(7), 04014020. https://doi.org/10.1061/(asce)co.1943-7862.0000857

Creedy, G. D., Skitmore, M., & Wong, J. K. W. (2010). Evaluation of risk factors leading to cost overrun in delivery of highway construction projects. *Journal of Construction Engineering and Management*, *136*(5), 528-537. https://doi.org/10.1061/(asce)co.1943-7862.0000160

Crucitti, P., Latora, V., Marchiori, M., & Rapisarda, A. (2003). Efficiency of scale-free networks: Error and attack tolerance. *Physica A: Statistical Mechanics and Its Applications*, *320*, 622-642. https://doi.org/10.1016/s0378-4371(02)01545-5

Deng, Y., Song, L., Zhou, Z., & Liu, P. (2017). An approach for understanding and promoting coal mine safety by exploring coal mine risk network. *Complexity*, 7628569. https://doi.org/10.1155/2017/7628569

Derakhshanalavijeh, R., & Teixeira, J. M. C. (2017). Cost overrun in construction projects in developing countries, gas-oil industry of Iran as a case study. *Journal of Civil Engineering and Management*, *23*(1), 125-136. https://doi.org/10.3846/13923730.2014.992467

Dikmen, I., Birgonul, M. T., & Han, S. (2007). Using fuzzy risk assessment to rate cost overrun risk in international construction projects. *International Journal of Project Management*, *25*(5), 494-505. https://doi.org/10.1016/j.ijproman.2006.12.002

Doloi, H. (2013). Cost overruns and failure in project management: Understanding the roles of key stakeholders in construction projects. *Journal of Construction Engineering and Management*, *139*(3), 267-279. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000621

Duenas-Osorio, L., & Vemuru, S. M. (2009). Cascading failures in complex infrastructure systems. *Structural Safety*, *31*(2), 157-167. https://doi.org/10.1016/j.strusafe.2008.06.007

Eteifa, S. O., & El-adaway, I. H. (2018). Using social network analysis to model the interaction between root causes of fatalities in the construction industry. *Journal of Management in Engineering*, *34*(1), 04017045. https://doi.org/10.1061/(asce)me.1943-5479.0000567

Eybpoosh, M., Dikmen, I., & Birgonul, M. T. (2011). Identification of risk paths in international construction projects using structural equation modeling. *Journal of Construction Engineering and Management*, *137*(12), 1164-1175. https://doi.org/10.1061/(asce)co.1943-7862.0000382

Fidan, G., Dikmen, I., Tanyer, A. M., & Birgonul, M. T. (2011). Ontology for relating risk and vulnerability to cost overrun in international projects. *Journal of Computing in Civil Engineering*, *25*(4), 302-315. https://doi.org/10.1061/(asce)cp.1943-5487.0000090

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003a). *Megaprojects and risk: An anatomy of ambition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107050891

Flyvbjerg, B., Holm, M. K. S., & Buhl, S. L. (2003b). How common and how large are cost overruns in transport infrastructure projects?. *Transport Reviews*, *23*(1), 71-88. https://doi.org/10.1080/01441640309904

Forcada, N., Rusinol, G., Macarulla, M., & Love, P. E. D. (2014). Rework in highway projects. *Journal of Civil Engineering and Management*, *20*(4), 455-465. https://doi.org/10.3846/13923730.2014.893917

Gharaibeh, H. M. (2014). Cost control in mega projects using the Delphi method. *Journal of Management in Engineering*, *30*(5), 04014024.
https://doi.org/10.1061/(asce)me.1943-5479.0000218

Gonzalez, A. M. M., Dalsgaard, B., & Olesen, J. M. (2010). Centrality measures and the importance of generalist species in pollination networks. *Ecological Complexity*, *7*(1), 36-43.
https://doi.org/10.1016/j.ecocom.2009.03.008

Hu, X. B., Gheorghe, A. V., Leeson, M. S., Leng, S. P., Bourgeois, J., & Qu, X. B. (2016). Risk and safety of complex network systems. *Mathematical Problems in Engineering*, 8983915. https://doi.org/10.1155/2016/8983915

Huang, W. Q., Zhuang, X. T., Yao, S., & Uryasev, S. (2016). A financial network perspective of financial institutions' systemic risk contributions. *Physica A: Statistical Mechanics and Its Applications*, *456*, 183-196.
https://doi.org/10.1016/j.physa.2016.03.034

Huang, Y. C., Cheng, W. Y., Luo, S. D., Luo, Y., Ma, C. C., & He, T. L. (2016). Features of the asynchronous correlation between the China coal price index and coal mining accidental deaths. *Plos One*, *11*(11), 0167198.
https://doi.org/10.1371/journal.pone.0167198

Iyer, K. C., & Sagheer, M. (2010). Hierarchical structuring of PPP risks using interpretative structural modeling. *Journal of Construction Engineering and Management*, *136*(2), 151-159.
https://doi.org/10.1061/(asce)co.1943-7862.0000127

Jin, Y., Zhang, Q., & Li, S. P. (2016). Topological properties and community detection of venture capital network: Evidence from China. *Physica A: Statistical Mechanics and Its Applications*, *442*, 300-311.
https://doi.org/10.1016/j.physa.2015.09.029

Kartam, N. A. (1996). Making effective use of construction lessons learned in project life cycle. *Journal of Construction Engineering and Management*, *122*(1), 14-21.
https://doi.org/10.1061/(asce)0733-9364(1996)122:1(14)

Khanzadi, M., Eshtehardian, E., & Esfahani, M. M. (2017). Cash flow forecasting with risk consideration using Bayesian belief networks (BBNS). *Journal of Civil Engineering and Management*, *23*(8), 1045-1059.
https://doi.org/10.3846/13923730.2017.1374303

Larsen, J. K., Shen, G. Q., Lindhard, S. M., & Brunoe, T. D. (2016). Factors affecting schedule delay, cost overrun, and quality level in public construction projects. *Journal of Management in Engineering*, *32*(1), 04015032.
https://doi.org/10.1061/(asce)me.1943-5479.0000391

Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters*, *87*(19), 198701.
https://doi.org/10.1103/PhysRevLett.87.198701

Li, Q. M., Song, L. L., List, G. F., Deng, Y. L., Zhou, Z. P., & Liu, P. (2017). A new approach to understand metro operation safety by exploring metro operation hazard network (MOHN). *Safety Science*, *93*, 50-61. https://doi.org/10.1016/j.ssci.2016.10.010

Li, X., & Chen, G. R. (2003). A local-world evolving network model. *Physica A: Statistical Mechanics and Its Applications*, *328*(1-2), 274-286. https://doi.org/10.1016/s0378-4371(03)00604-6

Liu, J. Z., & Tang, Y. F. (2005). An exponential distribution network. *Chinese Physics*, *14*(4), 643-645.
https://doi.org/10.1088/1009-1963/14/4/001

Nasir, D., McCabe, B., & Hartono, L. (2003). Evaluating risk in construction-schedule model (ERIC-S): Construction schedule risk model. *Journal of Construction Engineering and Management*, *129*(5), 518-527.
https://doi.org/10.1061/(asce)0733-9364(2003)129:5(518)

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, *45*(2), 167-256.
https://doi.org/10.1137/s003614450342480

Shane, J. S., Molenaar, K. R., Anderson, S., & Schexnayder, C. (2009). Construction project cost escalation factors. *Journal of Management in Engineering*, *25*(4), 221-229.
https://doi.org/10.1061/(asce)0742-597x(2009)25:4(221)

Simonsen, I. (2005). Diffusion and networks: A powerful combination!. *Physica A: Statistical Mechanics and Its Applications*, *357*(2), 317-330. https://doi.org/10.1016/j.physa.2005.06.032

Simonsen, I., Buzna, L., Peters, K., Bornholdt, S., & Helbing, D. (2008). Transient dynamics increasing network vulnerability to cascading failures. *Physical Review Letters*, *100*(21), 218701.
https://doi.org/10.1103/PhysRevLett.100.218701

Small, M., & Tse, C. K. (2005). Clustering model for transmission of the SARS virus: Application to epidemic control and risk assessment. *Physica A: Statistical Mechanics and Its Applications*, *351*(2-4), 499-511.
https://doi.org/10.1016/j.physa.2005.01.009

Sovacool, B. K., Gilbert, A., & Nugent, D. (2014). Risk, innovation, electricity infrastructure and construction cost overruns: Testing six hypotheses. *Energy*, *74*, 906-917.
https://doi.org/10.1016/j.energy.2014.07.070

Tabak, B. M., Takami, M., Rocha, J. M. C., Cajueiro, D. O., & Souza, S. R. S. (2014). Directed clustering coefficient as a measure of systemic risk in complex banking networks. *Physica A: Statistical Mechanics and Its Applications*, *394*, 211-216.
https://doi.org/10.1016/j.physa.2013.09.010

Tavakolan, M., & Etemadinia, H. (2017). Fuzzy weighted interpretive structural modeling: Improved method for identification of risk interactions in construction projects. *Journal of Construction Engineering and Management*, *143*(11), 04017084.
https://doi.org/10.1061/(asce)co.1943-7862.0001395

Touran, A., & Suphot, L. (1997). Rank correlations in simulating construction costs. *Journal of Construction Engineering and Management*, *123*(3), 297-301.
https://doi.org/10.1061/(asce)0733-9364(1997)123:3(297)

Wambeke, B. W., Liu, M., & Hsiang, S. M. (2012). Using Pajek and centrality analysis to identify a social network of construction trades. *Journal of Construction Engineering and Management*, *138*(10), 1192-1201.
https://doi.org/10.1061/(asce)co.1943-7862.0000524

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics 'small-world' networks. *Nature*, *393*, 440-442.
https://doi.org/10.1038/30918

Wright, E. R., Cho, K., & Hastak, M. (2014). Assessment of critical construction engineering and management aspects of nuclear power projects. *Journal of Management in Engineering*, *30*(4), 04014016.
https://doi.org/10.1061/(asce)me.1943-5479.0000286

Zhang, Y. L., & Yang, N. D. (2013). Research on robustness of R&D network under cascading propagation of risk with gray attack information. *Reliability Engineering & System Safety*, *117*, 1-8. https://doi.org/10.1016/j.ress.2013.03.009

Zhou, Z. P., Irizarry, J., & Li, Q. M. (2014). Using network theory to explore the complexity of subway construction accident network (SCAN) for promoting safety management. *Safety Science*, *64*, 127-136. https://doi.org/10.1016/j.ssci.2013.11.029

Zhou, J., Xu, W. X., Guo, X., & Ding, J. (2015). A method for modeling and analysis of directed weighted accident causation network (DWACN). *Physica A: Statistical Mechanics and Its Applications*, *437*, 263-277.
https://doi.org/10.1016/j.physa.2015.05.112

Appendix. Risk paths

No.	Risk paths
1.	Low quoted price →Lost
2.	Low quoted price → Problem of cash flow→Lost
3.	Ambiguity in contract clauses & Unknown geological conditions→ Lack of claims→Lost
4.	Unknown geological conditions → Delay in construction period → Project acceleration cost →Lost
5.	Unknown geological conditions → Delay in construction period →Lost
6.	Weather effects & Unknown geological conditions → Delay in construction period →Lost
7.	Unknown geological conditions → Delay in construction period →Lost
8.	Unknown geological conditions → Delay in construction period → Weather effects → Design scope change → Change in construction → Quantity increase →Lost
9.	Ambiguity in contract clauses & Unknown geological conditions → Quantity increase →Lost
10.	Ambiguity in contract clauses & Unknown geological conditions → Quantity increase → Project acceleration cost →Lost
11.	Unknown geological conditions → Quantity increase →Lost
12.	Weather effects & Unknown geological conditions →Lost
13.	Unknown geological conditions →Lost
14.	Unknown geological conditions → Design scope change → Delay in construction period →Lost
15.	Unknown geological conditions → Design scope change → Quantity increase →Lost
16.	Unknown geological conditions → Lack of technical skill and experience →Lost
17.	Unknown geological conditions → Natural disaster → Construction accident →Lost
18.	Unknown geological conditions → Natural disaster → Quantity increase →Lost
19.	Unknown geological conditions → Natural disaster →Lost
20.	Unknown geological conditions → Natural disaster → Delay in construction period → Price increase →Lost
21.	Weather effects & Incomplete design → Natural disaster →Lost
22.	Weather effects → Natural disaster → Construction accident →Lost
23.	Weather effects → Natural disaster →Lost
24.	Low management competency & Resettlement problem & Weather effects → Lack of claims →Lost
25.	Low management competency & Resettlement problem & Weather effects → Delay in construction period→Lost
26.	Weather effects → Delay in construction period →Lost
27.	Weather effects & Not buying insurance →Lost
28.	Lack of material or unqualified material →Lost
29.	Lack of material or unqualified material → Delay in construction period →Lost
30.	Lack of material or unqualified material → Delay in construction period → Idle worker →Lost
31.	Lack of material or unqualified material → Delay in construction period → Idle machine →Lost
32.	Lack of material or unqualified material → Quantity increase →Lost
33.	Lack of material or unqualified material → Change in construction →Lost
34.	Lack of material or unqualified material → Reworks → Delay in construction period → Project acceleration cost →Lost
35.	Pay for project in advance → Delay in payment →Lost
36.	Legislations and regulations change →Lost
37.	Legislations and regulations change → Low quoted price →Lost
38.	Quantity increase → Low quoted price →Lost
39.	Strike by workers → Delay in construction period →Lost
40.	Error in subcontractor →Lost
41.	Low management competency → Idle machine →Lost
42.	Low management competency → Ambiguity in contract clauses →Lost
43.	Ambiguity in contract clauses & Price increase →Lost
44.	Ambiguity in contract clauses & Natural disaster → Quantity increase →Lost
45.	Ambiguity in contract clauses → Quantity increase →Lost
46.	Lack of understanding of contract terms → Quantity increase → Delay in construction period →Lost
47.	Ambiguity in contract clauses → Lack of claims →Lost
48.	Ambiguity in contract clauses →Lost
49.	Natural disaster & Ambiguity in contract clauses → Delay in construction period →Lost
50.	Political ferment & Ambiguity in contract clauses →Lost
51.	Natural disaster & Ambiguity in contract clauses →Lost
52.	Ambiguity in contract clauses → Contract failure →Lost
53.	Ambiguity in contract clauses → Low quoted price →Lost

No.	Risk paths
54.	Ambiguity in contract clauses → Delay in payment →Lost
55.	Ambiguity in contract clauses → Insufficient site investigation → Delay in construction period →Lost
56.	Ambiguity in contract clauses → Construction accident →Lost
57.	Eager to win the contract → Low quoted price →Lost
58.	Eager to win the contract → Low quoted price || Legislations and regulations change →Lost
59.	Eager to win the contract → Ambiguity in contract clauses → Low quoted price →Lost
60.	Eager to win the contract → Pay for project in advance → Problem of cash flow →Lost
61.	Eager to win the contract → Fail to provide construction site → Delay in construction period → Idle worker || Idle machine || Weather effects →Lost
62.	Eager to win the contract → Insufficient site investigation → Bad construction environment → Natural disaster → Quantity increase || Delay in construction period →Lost
63.	Eager to win the contract → Ambiguity in contract clauses → Bad construction environment || Lack of claims →Lost
64.	Inadequate forecasting of market & Price increase & Political ferment →Lost
65.	Price increase → Low quoted price →Lost
66.	Price increase → Unstable condition of host country →Lost
67.	Price increase → Project administration cost increase →Lost
68.	Natural disaster & Not buying insurance →Lost
69.	No permit & approval → Delay in land acquisition → Delay in construction period → Pay for project in advance → Contract failure →Lost
70.	Fraudulent practices → Delay in payment → Idle machine || Problem of cash flow || Idle worker →Lost
71.	Lack of claims →Lost
72.	Poor communication and coordination → Idle worker →Lost
73.	Poor communication and coordination → Idle machine →Lost
74.	Lack of technical skill and experience → Lack of claims →Lost
75.	Lack of technical skill and experience → Lack of understanding of contract terms → Ambiguity in contract clauses →Lost
76.	Lack of technical skill and experience → Quantity increase →Lost
77.	Lack of technical skill and experience → Quantity increase → Low quoted price →Lost
78.	Lack of technical skill and experience → Misunderstand drawing → Error in construction →Lost
79.	Lack of technical skill and experience → Delay in construction period →Lost
80.	Lack of technical skill and experience → Reworks → Delay in construction period →Lost
81.	Lack of technical skill and experience → Low management competency → Construction accident →Lost
82.	Lack of technical skill and experience → Unknown geological conditions → Delay in construction period →Lost
83.	Lack of technical skill and experience → Low quoted price →Lost
84.	Lack of technical skill and experience → Error in construction →Lost
85.	Lack of technical skill and experience → Error in construction → Delay in construction period →Lost
86.	Design scope change → Quantity increase→Lost
87.	Design scope change → Quantity increase→ Change in construction→Lost
88.	Design scope change →Lost
89.	Design scope change → Change in construction→Lost
90.	Design scope change → Change in construction→ Lack of technical skill and experience→Lost
91.	Design scope change → Low quoted price→Lost
92.	Error in detail design → Quantity increase→ Project acceleration cost→Lost
93.	Error in detail design → Quantity increase→Lost
94.	Error in detail design → Reworks→ Delay in construction period→ Project acceleration cost→Lost
95.	Error in detail design →Lost
96.	Error in detail design → Change in construction→Lost
97.	Error in detail design → Unknown geological conditions→ Design scope change→ Change in construction→Lost
98.	Error in detail design → Natural disaster→Lost
99.	Error in detail design → Idle machine→Lost
100.	Design scope change → Delay in construction period→Lost
101.	Delay in design → Delay in construction period→ Project acceleration cost→Lost
102.	Delay in design → Delay in construction period→Lost
103.	Delay in design → Delay in construction period→ Natural disaster→Lost
104.	Delay in design → Idle worker || Idle machine→Lost
105.	Incomplete design → Quantity increase→Lost
106.	Incomplete design → Unknown geological conditions→ Quantity increase→Lost
107.	Incomplete design → Delay in construction period→Lost

No.	Risk paths
108.	Incomplete design → Change in contract→ Quantity increase→Lost
109.	Error in construction → Natural disaster→ Delay in construction period→Lost
110.	Error in construction → Natural disaster→Lost
111.	Insufficient site investigation & Bad construction environment→ Delay in construction period→Lost
112.	Bad construction environment → Quantity increase→Lost
113.	Delay in land acquisition → Fail to provide construction site→ Idle worker→ Delay in construction period→Lost
114.	Delay in land acquisition → Fail to provide construction site→ Idle machine→Lost
115.	Insufficient site investigation → Low quoted price→Lost
116.	Insufficient site investigation → Idle worker || Idle machine→Lost
117.	Insufficient site investigation → Design scope change→ Delay in construction period→ Weather effects→ Lack of technical skill and experience→Lost
118.	Insufficient site investigation → Low management competency→ Lack of claims→Lost
119.	Insufficient site investigation → Low management competency→ Delay in construction period→Lost
120.	Insufficient site investigation → Lack of material or unqualified material→ Delay in construction period→Lost
121.	Insufficient site investigation → Low management competency→ Quantity increase→Lost
122.	Insufficient site investigation → Quantity increase→Lost
123.	Delay in payment → Delay in construction period→Lost
124.	Fail to provide construction site & Delay in payment → Delay in construction period→Lost
125.	Delay in payment →Lost
126.	Delay in payment → Lack of material or unqualified material→ Delay in construction period→ Project acceleration cost→Lost
127.	Delay in payment → Problem of cash flow→Lost
128.	Delay in payment → Delay in construction period→ Project acceleration cost→Lost
129.	Delay in payment → Relocation problem → Resettlement problem→ Delay in construction period→Lost
130.	Delay in payment → Contract failure →Lost
131.	Delay in payment → Delay in construction period→ Pay for project in advance → Contract failure→Lost
132.	Delay in payment → Pay for project in advance→ Problem of cash flow → Delay in construction period || Construction accident || Poor quality→Lost
133.	Poor management ability of owner or supervision → Fail to provide construction site→ Delay in construction period→Lost
134.	Poor management ability of owner or supervision → Change in contract→ Quantity increase→Lost
135.	Fail to provide construction site → Quantity increase→Lost
136.	Fail to provide construction site → Delay in construction period→Lost
137.	Fail to provide construction site → Design scope change→ Quantity increase→Lost
138.	Fail to provide construction site → Idle worker || Idle machine→Lost
139.	Fail to provide construction site → Bad construction environment→ Change in construction→ Change in contract→Lost
140.	Delay of owner → Natural disaster→ Design scope change→ Quantity increase→Lost
141.	Delay of owner → Natural disaster→Lost
142.	Delay of owner → Natural disaster→ Delay in construction period→Lost
143.	Delay of owner → Idle worker→Lost
144.	Delay of owner → Idle worker→ Project acceleration cost→Lost
145.	Delay of owner → Idle machine→Lost
146.	Delay of owner → Idle machine→ Project acceleration cost→Lost
147.	Delay of owner → Weather effects→ Quantity increase→Lost
148.	Resettlement problem → Delay in construction period→Lost
149.	Resettlement problem → Idle machine→Lost
150.	Political ferment → Price increase→Lost
151.	Poor quality → Reworks→ Delay in construction period→Lost
152.	Poor quality → Change in construction→Lost
153.	Insufficient survey of sources of funds → Delay in payment→Lost
154.	Insufficient survey of sources of funds → Delay in payment→ Delay in construction period→Lost
155.	Natural disaster → Delay in construction period→Lost
156.	Natural disaster →Lost
157.	Natural disaster → Delay in construction period→ Project acceleration cost→Lost
158.	Natural disaster → Delay in construction period→ Idle worker || Idle machine→Lost

Note: "&" means that cost overrun or a COR occurs only if several previous CORs simultaneously happen, and "||" represents that any of CORs may lead to cost overrun.