A New Approach to Construction Project Risk Assesment Based on Rough Set and Information Entropy

Huawang Shi Wanqing Li Wenqing Meng
School of Cicil Engineering, Hebei Institute of Engineering, Handan 056038,P.R.China
E-mail: stone21st@163.com

Abstract

With the need for improved performance in the construction industry and increasing contractual obligations, the requirement of an effective risk management approach has never been more necessary. In this paper, a new approach to construction project risk assesment based on rough set and information entropy is proposed. Firstly, the construction project risk evaluation index system is established, and then the indexes are reduced with no information loss through rough set approach. Secondly, risk factor weights are qualitatively described with information entropy. Thirdly, the qualitative results are transformed to quantities value and the result of evaluation is worked out by using unascertained number algorithm. Finally, a case study was carried out on the risk assesment of a sample project using the prototype. The results show that the rough set and information entropy can help understand the uncertainties in construction risk assesment, and the relationships between risk sources and the consequences on project performance measures can be identified and quantified consistently.

1. Introduction

Construction projects are one-of endeavours with many unique features such as long period, complicated processes, abominable environment, financial intensity and dynamic organization structures and such organizational and technological complexity generates enormous risks. With the need for improved performance in the construction industry and increasing contractual obligations , the requirement of an effective risk management approach has never been more necessary. Risk assessment is a complex subject shrouded in vagueness and uncertainty. As practiced today, construction risk assessment generally rely on experts' intuitive experience. Scientific methods should be developed and employed during project planning and design stages in order to raise risk assessment t accuracy.

In this paper, a new method of evaluation of construction risk based on rough set and information entropy is proposed. The attribute quantity was reduced by rough set. The main characteristic attributes were withdrawn, the complexity of evaluation system and the computing time was reduced, as well. The indexes after reduced by rough sets were given evaluation weights with information entropy. and the evaluation results of construction risk was got based on unascertained number. Compared with the fuzzy synthetic evaluation, this model based on Rough set and information entropy is more precise, finer and higher in resolving power, and hence is more suitable to pre-liminary evaluation of construction risk.

2. Reduction algorithm based on Rough Sets

Rough Sets theory proposed by Z.Plawlak (1982) is one of such techniques. It is a novel mathematic method to study uncertain data, deficiency of data, incomplete data, or even inconsistent data. The algorithm has 2 steps, such as attribute reduction and attributes value reduction as follows.

(1) For an information system $S = (U, A, V, f)$, $A = C \cup D$, $B \subseteq C$, if $\gamma_C(D) = \gamma_B(D)$ and B is individual in relation to *D*, then *B* is the simplification of attribute *D* in relation to *C*, as in $RED_D(C)$. The calculation is shown as follows.

Input: *C, D, and U*

Output: attribute reduction *C* in relation to *D*

Step1 $s \leftarrow 0, RED(s) \leftarrow \phi$;

Step2 $i \leftarrow 1$;

Step3 $j \leftarrow 1, m \leftarrow 0$;

Step4 For subset *C(i,j)* of *C*, covering *j* subset of element *i*

978-0-7695-3435-0/08 $25.00 © 2008 IEEE
DOI 10.1109/ICIII.2008.85

① $t \leftarrow 0$

②If $(RED(t) \neq \phi) \wedge (RED(t) \subseteq C(i,j))$,

$m \leftarrow m+1$,if $m = C_{|C|}^{i}$, turn to Step7, else turn to Step5

③If $t \geq s$ turn to (5)

④ $t \leftarrow t+1$,turn to (2)

⑤ If $\gamma_C(D) = \gamma_{C(i,j)}(D)$ turn to (6), else turn to Step5

⑥ $s \leftarrow s+1$, $RED(s) = C(i,j)$

Step5 If $j \geq C_{|C|}^{i}$ turns to Step6,

else $j \leftarrow j+1$,turn to Step4

Step6 If $i \geq |C|$ ends, else $i \leftarrow i+1$, turn to Step3

Step7 Output $RED(s)$

(2) For in information system $S = (U, RED_D(C) \cup D, V, f)$, the calculation is shown as follows.

Input: $S = (U, RED_D(C) \cup D, V, f)$, $RED(C) = \{C_1, C_2, \cdots, C_n\}$

Output: core value table S' of S

Step1 $S' = (U, C \cup D, V' \leftarrow Null, f')$

Step2 For each condition attribute C_k (repeat as follows)

For each $x_i \in U$ and $C_k'(x_i) = Null$ (repeat as follows)

If $\exists x_i((x_j \neq x_i) \wedge \forall C_l(C_l \neq C_k \wedge C_l(x_j) = C_l(x_i)) \wedge (D(x_j) \neq D(x_i))$,Then

$C_k'(x_j) = C_k(x_j)$, $C_k'(x_i) = C_k(x_j)$

Step3 Output S'

3. Evaluation index system of construction projects risk

3.1. Attribute reduction based on Rough Sets

In this paper, there are 28 indexes were selected to describe the construction risk, such as variations by the client, design variations, incomplete or inaccurate cost estimate, tight project schedule, materials on time, project funding problems, contractors'poor management ability, project funding problems, tight projects cheduleand, price ination of construction materials and so on. Then the indexes are described based on rough sets. The table of original indexes of construction projects risk is shown in Tab.1

In Tab.1, A is the sets of condition attribute i.e., original indexes. For example, we disperse the results of original indexes as Tab.1. Construction projects risk is divided into 3 grades {1,2,3} to represent {low,average,high}. Similarly, the customer satisfaction degree is also divided into 4 grades {1,2,3,4} to represent {lower,low,high,higher}. D is the desicion attribute, and is divided into 3 grades {1,2,3} to represent {bad,average,good}. U is construction projects number whose risk is evaluation.

Tab.1 The table of original indexes

U \ A	design variations	tight project schedule	project funding problems	materials on time	contractors' difficulty in reimbursemen	…	D
1	3	4	2	2	2	…	1
2	3	3	2	3	1	…	1
3	2	2	1	1	1	…	3
4	2	4	1	1	2		2
5	1	2	1	2	1	…	3
6	3	3	2	1	2	…	1
7	3	4	2	2	3	…	1
8	3	3	1	1	4	…	2
…	…	…	…	…		…	…

3.2 Establishing evaluation index system

From Tab.1, the table of original indexes is reduced as follow by using rough sets algorithm. Construction project risk is described by many factors, including project funding problems, contractors'difficulty in reimbursemen, inadequate site information(soiltest and survey report), price ination of construction materials, contractors'poor management abilit, project funding problems, materials on time, variations by the client, inadequate safety measures or unsafe operations, inadequate program scheduling, suppliers'incompetency to deliver and poor competency of labour.

3.3. Evaluation index weight by information entropy

Shannon proposed the conception of information entropy, which is employed to measure the uncertainty roughly. Following the method of determining the index's identification weight by using the information entropy will be introduced. For the discretes to chastic variables, their information entropy is

$$H(\chi) = -\sum_{i=1}^{m} p(x_i)\log p(x_i),$$

$$0 \le p(x_i) \le 1, \sum_{i=1}^{n} p(x_i) = 1 \qquad (1)$$

In this paper

$$H(j) = -\sum_{k=1}^{K} \mu_{ijk} \cdot \log \mu_{ijk} \qquad (2)$$

$$\gamma_j = 1 - \frac{1}{\log K} H(j) = 1 + \frac{1}{\log K}\sum_{k=1}^{K}\mu_{ijk}\cdot\log\mu_{ijk} \qquad (3)$$

Where $w_j = \gamma_j \Big/ \sum_{j=1}^{d} \gamma_j$, $w = (w_1, w_2, \cdots, w_d)$ (4)

Obversly, $0 \le w_j \le 1$ and $\sum_{j=1}^{d} w_j = 1$.

3.4. Synthesis appraisal system

As it is known that $\mu_{ijk} = \mu(x_{ij} \in c_k) \quad 1 \le i \le n, 1 \le j \le m$ is the unascertained measur and μ_{ijl} is unit x_i 's single factor's measure appraisal matrix, in which, μ_j^i means x_{ij} makes x_i has c_k grade in j row.

$$(\mu_{ijk})_{m\times K} = \begin{pmatrix} \mu_{i11}, \mu_{i12}, \cdots, \mu_{i1K} \\ \mu_{i21}, \mu_{i22}, \cdots, \mu_{i2K} \\ \vdots \quad \vdots \quad \cdots \quad \vdots \\ \mu_{im1}, \mu_{im2}, \cdots, \mu_{imK} \end{pmatrix} \qquad (5)$$

where,

$(i = 1,2,\cdots,n \quad j = 1,2,\cdots,m \quad k = 1,2,\cdots,K)$

If the single factor measure appraisal matrix above is known, the each factor's classific ation vector about x_i is (6), and then (7) got as fllow:

$$W^i = (w_1^i, w_2^i, \cdots, w_m^i) \qquad (6)$$

$$\mu^i = W^i \cdot (\mu_{ijk})_{m\times K} = (w_1^i, w_2^i, \cdots, w_m^i)\begin{pmatrix} \mu_{i11}, \mu_{i12}, \cdots, \mu_{i1K} \\ \mu_{i21}, \mu_{i22}, \cdots, \mu_{i2K} \\ \vdots \quad \vdots \quad \cdots \quad \vdots \\ \mu_{im1}, \mu_{im2}, \cdots, \mu_{imK} \end{pmatrix}$$

$$= \left(\sum_{j=1}^{m} w_j^i \cdot \mu_{ij1}, \sum_{j=1}^{m} w_j^i \cdot \mu_{ij2}, \cdots, \sum_{j=1}^{m} w_j^i \cdot \mu_{ijK}\right) \qquad (7)$$

So μ^i is x_i 's appraisal vector.

3.5. Principle of identification

Because the classification of the comment ranks is orderly. e.g., c_k is "better" than c_{k+1} , the identification principle of"maximum measure'' is not available. The credible identification principle is needed. Let the credible identification be $\lambda, (\lambda > 0.5)$, and 0.6 or 0.7 is always adopted.

$$k_0 = \min_k \left[(\sum_{l=1}^{k} \mu_{il}) \ge \lambda, k = 1,2,\cdots K \right] \qquad (8)$$

then x_i belongs to the rank c_{k_0} . It means that when x_i is not lower than c_k , the fiducial degree is λ ,or in other words lower than c_k is 1- λ .

4. Case study

According to Tab.1, description results of evaluation by one expert are shown in Tab.2. And description results of evaluation by other experts are ellipsised.

Tab.2 Experts description results of index evaluation

		I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9	I_{10}	I_{11}	I_{12}
Evaluation	higher	0.2	0.2	0.3	0.2	0.1	0.2	0.3	0.2	0.1	0.2	0.1	0.2
	high	0.2	0.2	0.3	0.2	0.1	0.2	0.3	0.2	0.1	0.2	0.1	0.2
	Average	0.3	0.1	0.1	0.2	0.3	0.2	0.1	0.2	0.3	0.2	0.3	0.1
	low	0.1	0.2	0.3	0.2	0.2	0.1	0.3	0.2	0.2	0.2	0.2	0.2
	lower	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2

As shown in Tab.2, the index set is{best, better, good, Average, Bad }and it is divided into five appraisal scales, and by all appearances, it is positive sequence. The each factor is total ten score. Then, each appraisal object is ten score and the distinguishment is that the degrees are different. The scoring principle is fairly and fit the measurement criterion. Based on the statistical data of the appraisal object, the single factor measurement matrix μ_{1jk} is as follows:

$$\mu_{1jk} = \begin{bmatrix} 0.139 & 0.235 & 0.400 & 0.210 & 0.016 \\ 0.096 & 0.365 & 0.384 & 0.106 & 0.049 \\ 0.044 & 0.129 & 0.351 & 0.403 & 0.073 \\ 0.063 & 0.182 & 0.379 & 0.231 & 0.145 \\ 0.221 & 0.327 & 0.319 & 0.117 & 0.016 \\ 0.018 & 0.174 & 0.456 & 0.335 & 0.017 \\ 0.145 & 0.214 & 0.369 & 0.201 & 0.071 \\ 0.139 & 0.235 & 0.400 & 0.210 & 0.016 \\ 0.063 & 0.231 & 0.432 & 0.124 & 0.150 \\ 0.355 & 0.241 & 0.210 & 0.143 & 0.051 \\ 0.145 & 0.214 & 0.369 & 0.201 & 0.071 \\ 0.223 & 0.249 & 0.317 & 0.120 & 0.091 \end{bmatrix}$$

According to the equations(2)-(4), the factor's weight vector is

$$\omega_j^i = (0.0761, 0.0620, 0.0928, 0.1482, 0.1116, 0.1005, 0.0514, 0.0457, 0.0947, 0.09150, 0.0645, 0.0610)$$

According to equation(7), the object's appraisal vector is

$$\mu^1 = W^1 \cdot (\mu_{1jk}) = (0.2904, 0.3026, 0.2050, 0.1960, 0.0048)$$

We adopt λ =0.6, according to the equation(8), when k_0=3, 0.2904+0.3026+0.2050= 0.798> 0.6.

So, the appraisal object is average. That is to say, this construction projects risk is average.

Based on the research data and the appraisal result, this company can know the construction projects' risk degree. In the mean time, the clients, designers, government bodies should take the responsibility to manage their relevant risks and work cooperatively from the feasibility phase onwards to address potential risks in time, contractors and subcontractors with robust construction and management knowledge should be employed to minimize construction risks and carry out safe, efficient and quality construction activities.

5. Conclusions

In this paper, the factor of evaluation index system of construction projects risk is sufficiently analyzed, and evaluation and weight of index are quantified. So, we can get some conclusions as follows.

(1) The algorithm of rough sets is feasibility in reducing indexes. And the evaluation index system is more proper than traditional.

(2) Employing the information entropy and credible identification principal, it overcomes the shortcoming of fuzzy comprehensive assessment and makes the result more objective.

(3) The result of evaluation is more sophisticated than traditional evaluation.

6. References

[1] Thompson PA, Perry JG. Engineering construction risks: a guide to project risk analysis and risk management[M]. London: Thomas Telford, 1992.

[2] Flanagan R, Norman G. Risk management construction [M]. UK: Blackwell, 1993.

[3] V.Carr, J.H.M.Tah. A fuzzy approach to construction project risk management and analysis:construction project risk management system[J]. Advances in Engineering software, 2001(32):847-857.

[4] PatrickX.W.Zou, Guomin Zhang, Jiayuan Wang. Understanding the key risks in construction projects in China[J]. International Journal of Project Management, 2007 (25):601-614.

[5] Z. Pawlak. Rough Sets-Theoretical Aspects of Reasoning about Data [M]. Klystron Academic Publisher, 1994.

[6] L Polkowski, A skowron. Rough Sets:A Perspective. Rough Sets in Knowledge Discovery(1,2), Physica-Verlag, Heidelberg,1998.

[7] Liu K D, Wu H Q, Pang Y J. Process and Application of Uncertain Information [M]. Beijing: Science Press, 1999(in Chinese)

[8] Institution of Civil engineerings and faculty and Institute of Actuaries. RAMP: risk analysis and management for projects. London: Thomas Telford, 1998.