The relationship between management and control indexes of distribution network construction based on Apriori algorithm

YANG Zhuo1, YU Jinxiong2,*, LI Huitao3, ZHENG Zhicong4, WANG Lili1, WANG Kehui2, PAN Zhaolun2, SUN Huijun2

1 State Grid Henan Electric Power Company, Henan Zhengzhou, China
2 North China Electric Power University, Beijing Changping, China
3 State Grid Henan Electric Power Company Economic and Technology Research Institute, Henan Zhengzhou, China
4 State Grid Pingdingshan Eletic Power Company, Henan Pingdingshan, China

Abstract. There are a large number of distribution network construction projects with different types, and a large number of complex data relationships are generated in the construction process. This paper combs the key management and control indicators of distribution network construction projects, and constructs the management and control index system of each stage of the project. Based on Apriori algorithm, this paper studies the relationship between the key management and control indicators, and provides reference for the management and control of distribution network construction.

1 Introduction
Distribution network is an important public infrastructure for national economic and social development, which is the key link to connect the transmission network and users and ensure the reliability of power supply. However, the distribution network has the characteristics of large-scale construction, various types of equipment, wide distribution range and large regional differences. In addition, the distribution network has many engineering stages, complex control objects, large amount of data involved, strict management requirements, close connection between the project stage and the control object, and complex data relationship, which makes the management and control of the whole process of the distribution network project more difficult. Therefore, it is necessary to study the key factors of the whole process management and control of the distribution network engineering construction.

At present, the management research of distribution network construction projects mainly focuses on single distribution network projects, and few studies on key control indexes in the construction process. This paper starts with the construction of the key management and control index system of distribution network, finds out the relationship between the control indexes, and provides reference for the management and control of the construction of distribution network.

2 Key index system for whole process control of distribution network construction
This chapter searches for the key control indicators of the whole process of distribution network construction project, and divides them into decision-making, design, construction and completion phases in chronological order, and each phase is divided into cost, quality, and progress.

2.1 Decision stage indicators
In terms of cost, the research fee mainly includes two major costs: tool use fee and survey labor cost; in terms of quality, the site selection rationality of relevant projects is considered; in terms of progress, the survey progress is mainly composed of two progress indicators: Inspection progress and decision-making progress.

2.2 Design stage indicators
In terms of cost, the design input cost is considered. In terms of quality, the rationality of design and the adequacy rate of capital preparation are considered. In terms of progress, the design progress and project approval progress are considered.

* Corresponding author: jinxiong.yu@163.com
© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
2.3 Construction stage indicators

In terms of distribution network construction cost, four indicators are selected: labor cost, material cost, machinery cost and measure cost. The measure cost is subdivided into construction increase cost and temporary facility cost. In terms of quality, construction method, material qualification rate and construction qualification rate are considered. In terms of progress, the five indicators of equipment operation efficiency, construction personnel efficiency, material supply rate, social and economic factors and adverse weather delay rate are considered for control.

\[
\sigma_j = \sum_{i=1}^{n} x_{ij} \quad \text{(4)}
\]

\[
lift(j \Rightarrow k) = \frac{\text{support}(j \Rightarrow k)}{\text{support}(j) \cdot \text{support}(k)} \quad \text{(5)}
\]

\[
\text{support}(j) = \frac{\sigma_j}{n} \quad \text{(6)}
\]

3 Association relationship mining model of key management and control indicators for distribution network construction projects

3.1 Association rules for key management and control indicators of distribution network construction projects

Assuming that m key indicators are set during the construction process, these indicators need to be evaluated. Collected n index records related to the distribution network construction project, each record contains the completion of the set m key indicators. Each distribution network construction project index record has at least one key management and control index deviation.

\[
x_i = \begin{cases} 0, & \text{There is no deviation in the j-th index of the i-th project} \\ 1, & \text{There is deviation in the j-th index of the i-th project} \end{cases} \quad \text{(1)}
\]

Where: \( x_{ij} \) is the implementation of the j-th index of the i-th distribution network construction project; \( i = 1, 2, ..., n \); \( j = 1, 2, ..., m \).

Calculate the \( \text{support}(j \Rightarrow k) \), \( \text{confidence}(j \Rightarrow k) \), \( \text{lift}(j \Rightarrow k) \) between the j-th and k-th indicators.

\[
\text{support}(j \Rightarrow k) = \frac{\sigma_{jk}}{n} \quad \text{(2)}
\]

\[
\text{confidence}(j \Rightarrow k) = \frac{\sigma_{jk}}{\sigma_j} \quad \text{(3)}
\]

2.4 Completion stage indicators

In terms of cost at the completion stage, the test tool usage fee and acceptance labor fee are used as the key control indicators at this stage. In terms of quality, it is believed that the acceptance criteria will affect the quality of the project, so it is listed as an indicator. In terms of progress, the acceptance progress is introduced as an indicator.

In summary, the key index system for the whole process of distribution network construction is shown in Figure 1.
3.2 Association relationship mining based on Apriori algorithm

The mining of frequent itemsets in the Apriori algorithm is gradually realized through multiple iterations. In the first iteration, calculate the support, confidence and lift of each item in the indicator system and record all the frequent items found as \( k \). In the subsequent rounds of iterations, the \( k \)-item set is used as the seed itemset of this round of iteration to generate a candidate \((k+1)\)-item set. In this iteration, in order to mine all \((k+1)\)-frequent itemsets as the seed itemsets of the next iteration, it is necessary to calculate the actual support, confidence and lift of each candidate \((k+1)\)-itemsets. The iterative process continues until no new frequent itemsets can be generated.

According to the definition of frequent itemsets, in order to find out all frequent itemsets and determine whether various combinations are frequent items, it is necessary to exhaust all combinations of indicators in a distribution network construction project, and calculate the support, confidence and lift of each combination. Apriori algorithm uses the following two attributes to reduce the computation of item set combination.

Property 1: any nonempty subset of a frequent itemset is frequent.

Property 2: any superset of an infrequent itemset is non frequent.

4 Empirical analysis

The support, confidence, and lift threshold settings will have a greater impact on the results of the association relationship. If the threshold is too small, some less closely related associations will be generated; if the threshold is too large, the closely related associations will be deleted. According to the principle of not missing important associations, the thresholds of support, confidence, and lift are set to 0.3, 0.7, and 1, respectively.

In this study, 27 key indicators are determined temporarily. According to the data records of 178 distribution network construction projects collected, the support degree, confidence degree and lift degree between the indicators are calculated, and compared with the threshold value, and some correlation relations are selected to describe and interpret. Taking the rationality of site selection as an example, Table 1 shows the calculation results of the indicators related to the rationality of site selection.

| No. | Association | Support | Confidence | Lift |
|-----|-------------|---------|------------|------|
| 1   | Rationality of site selection -- design progress | 0.38    | 0.79       | 1.12 |
| 2   | Rationality of site selection -- approval progress | 0.35    | 0.75       | 1.05 |
| 3   | Rationality of site selection -- labor cost | 0.62    | 0.86       | 1.05 |
| 4   | Rationality of site selection -- machinery cost | 0.63    | 0.76       | 1.21 |
| 5   | Rationality of site selection -- equipment operation efficiency | 0.53    | 0.84       | 1.02 |
| 6   | Rationality of site selection -- efficiency of construction personnel | 0.50    | 0.78       | 1.17 |

It can be seen that in the first association, the support degree of rationality of site selection -- decision progress is 0.38, which indicates that 38% of the records in the project concentration are biased in the rationality of site selection and the design progress. The confidence level of 0.79 indicates that there is a 79% probability that the design progress will also deviate in the project records with deviation in site selection rationality. The increase degree is greater than 1, which indicates that the two are positively correlated, and the higher the value, the higher the correlation.

Approval schedule, labor cost, machinery cost, equipment operation efficiency and construction personnel efficiency are also related to the rationality of site selection. It shows that when there is deviation in the rationality of site selection, the indicators related to the rationality of site selection are more likely to be biased. In the process of management and control, when the rationality of site selection is found to be deviated, the related indicators can be controlled in advance.

The material qualification rate, construction qualified rate, construction personnel efficiency, material supply rate are related to the capital adequacy rate, which indicates that when the capital adequacy rate deviates, the indicators related to the capital adequacy rate tend to deviate. The shortage of funds may lead to the low qualified rate and supply rate due to insufficient material purchase fee, and the working enthusiasm and efficiency of construction personnel will be reduced, so the qualified rate will also decline.

The efficiency of construction personnel is related to labor cost, which indicates that the deviation of labor cost will lead to the decline of construction personnel efficiency.

The qualified rate of construction is related to the construction method. Improper construction method is easy to lead to unqualified construction.

The qualified rate of construction and the progress of acceptance are related to the qualified rate of materials, which indicates that when the qualified rate of materials has problems, the two indicators of construction qualified rate and acceptance progress are more likely to have deviation.

The acceptance progress is associated with the construction pass rate, which indicates that the construction acceptance rate is not up to standard, which is more likely to lead to the deviation of the acceptance progress.

Labor cost and construction increase cost are related to the delay rate of social and economic factors, which indicates that the frequent occurrence of delay caused by social and economic factors is more likely to lead to excessive labor cost and increased construction cost.

Labor cost, construction increase cost and temporary facilities cost are related to the delay rate of bad weather, which indicates that when the delay caused by bad weather is more, it is easier to lead to deviation of labor cost, construction increase cost and temporary facility cost.

By integrating the above relationships, the key management and control indicators association diagram of distribution network construction project is obtained as
shown in Figure 2.

![Diagram showing key management and control indicators of distribution network construction projects]

**5 Conclusion**

In this paper, the key management and control index system is established according to the three dimensions of cost, quality and progress in each stage of distribution network construction project. The association algorithm is first applied to the research on the association relationship of key management and control indicators, and the mining model of association relationship is constructed. According to the collected data of distribution network construction project, the support degree, confidence degree and promotion degree between the key control indicators are calculated, and compared with the set value. The correlation between the key control indicators is obtained by comparing the thresholds. In the future management of distribution network construction projects, if a certain index deviates, the related indicators can be controlled in advance to reduce the impact on the project implementation and achieve the effect of prior control.

**Acknowledgment**

This work is supported by Research on Big Data Correlation Analysis of Key Nodes in Power Distribution Network Engineering Project Management (SGTYHT/18-JS-209).

**References**

1. LI Weiguo, XU Wenwen, QIAO Zhenyu, WANG Xuguang. Distribution section fault location method based on transient zero sequence current bump feature[J]. Power System Protection and Control, 2020, 48(10):164-173.

2. YANG Feiran, YU Yongjin. Fault diagnosis of distribution network based on time-constrained hierarchical fuzzy Petri net[J]. Power System Protection and Control, 2020, 48(02):99-106.

3. Z. Q. Bo, X. N. Lin, Q. Wang, Y. H. Yi, F. Q. Zhou. Developments of power system protection and control[J]. Springer Singapore, 2016, 1(1):1-8.

4. Winters Tobey. Retail electricity markets require marginal cost real-time pricing[J]. The Electricity Journal, 2001, 14(9):74-81.