Association Rule Analysis on Impact Factors of Workforce Quota in Power Enterprise

Yuan ZHANG¹, Si-yu CHEN²,³, Meng-yi SHA²,³ and Min-gang GAO³

¹Dept. of Management Consulting, State Grid Energy Research Institute, Beijing, China, 102209
²University of Chinese Academy of Sciences, Beijing, China, 100049
³Institutes of Science and Development, Chinese Academy of Sciences, Beijing, China, 100190

*Corresponding Author

Keywords: Human resource, Workforce quota, Power enterprise, Association rule, Data mining.

Abstract. For the study of impact factors and their relationship of workforce quota in power supply enterprises, from the perspective of optimizing the management of human resources, an association rule analysis model for influencing factors of labor placement in power supply enterprises was established based on the data of category of main transformer maintenance in a city-level power supply enterprise. The results show that it is feasible to apply this method to study the influencing factors of workforce quota in power enterprise, and the association rule analysis conclusion can optimize the management of workforce quota in power enterprise.

Introduction

Electricity is an important support for national economic development. As the largest power producer in the world, China's power industry is becoming more and more important both at home and abroad. The promulgation of "Some Opinions on Further Deepening the Reform of Electricity System" in March 2015 marks the beginning of a new round of electricity market reform. With the reform of transmission and distribution price, the salary of employees will be taken as part of the new allowable cost in the supervision cycle of transmission and distribution price, and the requirement of human resource management in power supply enterprises will be strengthened.

Facing the changes of the external environment and the adjustment of the service focus of power supply enterprises, power supply enterprises need to establish a scientific and efficient human resource management system. Due to the wide coverage and great differences between different jobs of power supply enterprises, there are a lot of historical problems left over in the human resources planning, such as overall overcrowding and structural shortage of staff, planning and formulation of the actual work of various professions are out of line. The scientific and quantitative management of labor quota is one of the most basic and important parts in the intensive management of enterprises. With the increase of the proportion of networking operations, cloud computing technology can be used to supervise the human resource management of power supply enterprises in real time, accurately grasp the demand side in the balance of human supply and demand.

Researchers at domestic and abroad are focusing on the factors affecting the workforce quota of power supply enterprises, which can be roughly divided into subjective and objective factors. The subjective factors on the impact of labor placement is generally recognized by the industry, but the classification of personnel factors is different. Peng Hong (2004) sorted out the influencing factors under different methods of staffing, pointing out that the historical analysis method can be divided into the number of historical staffs and the quality of personnel, the strategic decomposition analysis method pays attention to the personal preferences of managers, the ability of managers and subordinates, and the econometric analysis method emphasizes labor efficiency. Tang Wei et al. (2012) put forward the producers' professional knowledge, operational skills, managerial level, employment structure and total quantity. Che Xiuyan et al. (2006) used the national policy, the size of labor intensity and the degree of difficulty of labor and other factors. Qian Cheng et al. (2015) were
more concerned about the impact of internal factors on staff management, such as organizational strategy. Wang Hua et al. (2015) started from the electric power enterprises, emphasizing the topography, power supply service area, and number of business units. Jorne et al. (2013) also proposed factors such as personnel labor contract relationship, mode of operation, and job rotation. It can be found that the main influencing factors are the level of external economic development, as well as the production technology, equipment conditions, labor environment and other conditions within the enterprise. The identification of objective factors in each study is not the same, and mainly relies on literature research, experience summary and other qualitative methods, without quantitative verification.

With the improvement and accumulation of PMS and other systems and work order information, power supply enterprises have accumulated a large number of operational data, but these data currently mainly play a record and personal performance appraisal role, do not excavate the potential value of improving the level of human resources management. Therefore, this study attempts to use association rules method, with a municipal power supply company's main transformer maintenance professional work order data from 2014 to 2017 as a sample, establish a model of impact factors of workforce quota in power enterprise, mine the influencing factors of power supply enterprise labor appointment and analyze their relationships, then put forward suggestions for optimizing labor appointment management of power supply enterprise.

Association Rules and Apriori Algorithm

Association rules can discover the relevance of each item, and the mining process can be divided into two steps: first, find all frequent item sets satisfying the degree of support greater than or equal to the minimum support from the transaction set; second, generate association rules from these frequent item sets, calculate the confidence of these association rules, and then retain the rules of satisfying the degree of confidence greater than or equal to the minimum confidence.

Apriori algorithm is a classical algorithm, which uses downward closed attributes. Its principle is to generate (k+1) - itemsets from k-itemsets by using iterative method of layer-by-layer search. First, scan the transaction set to find the set of frequent 1-itemsets, which is called $L_1$, then use $L_1$ to find the set $L_2$ of frequent 2-itemsets, and then use $L_2$ to find $L_1$, so iterate until the location of frequent k-itemsets can not be found, in which each $L_k$ must traverse the transaction set. From the description above, we can see that there are association rules between any two transaction sets in a given transaction set. In order to find meaningful association rules, we should select the appropriate minimum support and minimum confidence in the mining process.

The Model of Impact Factors of Workforce Quota in Power Enterprise

Analysis of Influencing Factors

Firstly, combine with literature research and enterprise research to sort out the influencing factors of staffing, and organize expert discussions or interviews to form a preliminary index database. Secondly, considering the quantitative support of index data, and eliminate the factors that could not be quantified; considering the correlation between indicators, then eliminate the indicators that can be reused. Finally, classify and analyze the influencing factors of the nine types of internal staffing standards, and improve the index database, as shown in Table 1.
Table 1. Impact factor set of workforce quota in power enterprise.

| No. | Type              | Impact factors                                                  |
|-----|-------------------|----------------------------------------------------------------|
| 1   | Economic factor   | Electricity sales, operating income, fixed assets, corporate profits, and regional per capita GDP |
| 2   | Equipment condition| Equipment import substitution rate, equipment life and equipment load |
| 3   | Personnel quality | Length of service, level of vocational skills and education     |
| 4   | Service situation | Operation radius                                               |
| 5   | Others            | Impact of new jobs                                             |

Based on the index database of influencing factors of manpower quota in power supply enterprises, this paper divides the influencing factors into five categories: economic factor, equipment condition, personnel quality, service condition and other factors. Because of the great difference in the positions of power supply enterprises, through the sorting and analysis of the nine categories of fixed-size items in the current standard, the index set of influencing factors for each fixed-size profession formed after supplementing the factors not taken into account in the standard, as shown in Table 2.

Table 2. Impact factor set of workforce quota for each post in power enterprise.

| No. | Post                        | Impact factors                                                                 |
|-----|------------------------------|--------------------------------------------------------------------------------|
| 1   | Maintenance and overhaul     | - Line length and number of equipment with different natural conditions and voltage levels |
|     |                              | - Length of service, level of vocational skills and education, and equipment life |
| 2   | Operation and dispatch       | - Number of equipment and lines, and number of business households             |
|     |                              | - Length of service, level of vocational skills and education, and equipment life |
| 3   | Information and communication| - Number of dispatching objects (number of substations, etc.) and number of different types of devices |
|     |                              | - Length of service, level of vocational skills and education, and equipment life |
| 4   | Construction                 | Output value (design fee, supervision fee, construction cost); the number of items selected for evaluation, technical supervision |
| 5   | Plan                         | Incremental index of grid volume, such as five-year planning of increment of investment in grid infrastructure with different voltage levels, and increment of installed capacity of power generation, etc. |
| 6   | Goods and materials          | - Number of material requirement plan or material contract, and contract amount |
|     |                              | - Material procurement cycle                                                  |
| 7   | Marketing                    | - Business volume                                                             |
|     |                              | - Length of service, level of vocational skills and education, and regional differences (level of technological development, cultural quality, etc.) |
| 8   | Training and education       | The number of trainees in the year, the number of personnel evaluated in the previous year, the number of management objects such as equipment and books, etc. |
| 9   | Management                   | Head office approval, business quota and subordinate (if there are subordinate) number of management staff |

**Attribute Determination and Data Processing**

According to the actual data, this study only obtains the relevant data of the main transformer maintenance post of a municipal power supply company, so now without considering the economic factors, chooses five factors as the attribute of the set in the association rules: voltage level, equipment operating life, length of service, skill level and educational background.

Based on the data obtained from the record of work, staff information, basic equipment information and others of main transformer maintenance post, excluding unqualified data such as null value, invalid value, and other unrelated items, there are 721 effective records with higher work efficiency.

To do data mining, the data should be discretized and transformed into corresponding transactional data in association rule mining. Voltage level is divided into KV1 [110kV] and KV2 [220kV]; equipment operating life is divided into EP1 [0 years, 10 years], EP2 [11 years, 20 years] and EP3 [21 years, 31 years], according to the professional practical significance; length of service is divided into SE1, SE2, ..., SE9; skill level is classified into SL1 [intermediate worker], SL2 [senior worker], SL3
[technician], and SL4 [senior technician]; education background is classified into AD1 [secondary college graduate], AD2 [junior college graduate], and AD3 [undergraduate graduate]. The discrete data is shown in Table 3. All attributes of each work order record can be viewed as a transaction, and the attributes of all work orders form a transaction set.

Table 3. Preprocessed data.

| No. | Voltage level | Equipment operating life | Length of service | Skill level | Educational background |
|-----|---------------|--------------------------|-------------------|-------------|------------------------|
| 1   | KV2           | EP1                      | SE3               | SL3         | AD3                    |
| 2   | KV2           | EP1                      | SE2               | SL4         | AD2                    |
| 3   | KV2           | EP1                      | SE2               | SL1         | AD3                    |
| 4   | KV1           | EP1                      | SE3               | SL3         | AD3                    |
| 5   | KV1           | EP1                      | SE2               | SL2         | AD3                    |
| ... | ...           | ...                      | ...               | ...         | ...                    |

Support and Confidence Threshold Setting

The key to get accurate association rules is whether the support and confidence threshold settings are appropriate. The support degree and confidence level are not fixed, and they could be continuously revised according to the results of mining. The setting of two thresholds has a great influence on the size of the generated rules. If the thresholds are set too large, useful association rules are easily lost; if the thresholds are set too small, the number of rules generated will be very large. Due to the limited data accumulation foundation of the influencing factors of the company's workforce quota, the amount of sample data is small and it is difficult to collect, the threshold of support and confidence should be small. The threshold setting will draw lessons from the idea of trial and error method, and constantly adjust the threshold to get the ideal mining results.

Association Rules Mining and Analysis for Influencing Factors of Labor Quota

The non-commercial data mining software Weka is selected as the platform for model building and analysis. The pre-processed data is first converted into ARFF format, and then Apriori algorithm is selected for data mining.

Based on the work order records of the main transformer maintenance team of a municipal power supply company from January 2014 to August 2017, the data preprocessing is completed according to the mining steps of association rules of influencing factors. The data model is established by using data mining software Weka, and the support and confidence thresholds are set up. And generate frequent itemsets and strong association rules. As the experimental debugging, set the minimum support of 0.15, minimum confidence of 0.75, and find the best association rules as shown in Table 4.

Table 4. Association rules of influence factors in main transformer maintenance post.

| Rule ID | Antecedent                        | Consequent          | Case | Confidence coefficient |
|---------|-----------------------------------|---------------------|------|------------------------|
| 1       | 2 years of service                | Undergraduate       | 110  | 0.92                   |
| 2       | 110kV, secondary college          | Senior technician   | 125  | 0.89                   |
| 3       | Secondary college                 | Senior technician   | 175  | 0.88                   |
| 4       | 6 years of service, secondary college | Senior technician | 110  | 0.87                   |
| 5       | 10-20 years of equipment life, senior technician | 110kV | 111 | 0.82                   |
| 6       | 6 years of service                | Senior technician   | 141  | 0.79                   |
| 7       | 6 years of service, Senior technician | Secondary college | 110  | 0.78                   |
| 8       | 110kV, 1 year of service          | Undergraduate       | 111  | 0.77                   |
| 9       | 10-20 years of equipment life     | 110kV               | 199  | 0.77                   |
| 10      | 1 year of service, undergraduate  | 110kV               | 111  | 0.76                   |

a. The efficiency of two-year undergraduate staff is relatively higher. In recent years, the company has increased the requirement of qualifications when recruiting employees, and the number of undergraduate graduates has increased. This kind of staff may have a strong learning ability. After two
years of training and practical operation, they can reach the level of complete proficiency and their work efficiency is higher.

b. Senior technicians graduated from secondary specialized schools have higher efficiency in 110kV equipment with 6 years of service. Employees with relatively low academic qualifications are more efficient when they have high skill levels and rich work experience through training and practical work. The company can strengthen staff training and skill level assessment, emphasizing the actual operation degree. In the process of workforce quota management, it is necessary to consider whether there is a substitution relationship between the actual post's academic level and skill level, and focus on the factors needed to improve the actual work efficiency of the specific post.

c. Employees are higher efficient in operating equipment with a life span of 10-20 years. The main reason is that the new equipments are in debugging stage and the old equipments are in maintenance stage, their operating efficiency is lower, in line with the equipment failure rate’s U-shaped trend.

Conclusion
In this study, firstly, the influencing factors of workforce quota in power supply enterprises are analyzed, and the identification model of influencing factors of workforce quota in power supply enterprises is constructed by using association rules, and the association rules are analyzed by using the work order data of a municipal power supply enterprise to find out the influencing factors of high work efficiency and the relationship between them. The data sample size of the power supply enterprise is small, but if the model is applied to the provincial, municipal and rural power supply companies, and various types of appointment posts, it will help to improve the efficiency of power supply enterprises and optimize the management of labor appointment.

Acknowledgement
This research was financially supported by SGCC's Science and Technology Research Project - Quantitative Analysis of Workforce Quota in Power Supply Enterprises Based on Cloud Sharing Platform.

References
[1] Peng Hong. Redesign of manpower appointment in D company's production guarantee department [D]. Sichuan University, 2004.
[2] Tang Wei, Fang Shuzhen. Research and exploration of flexible human resource allocation mechanism based on labor employment [J]. China Electric Power Education, 2012 (15): 36-38.
[3] Che Xiuyan, Mao Xue Kai. Talking about the fixed quota of enterprises [J]. China Electric Power Education, 2006 (S2):104-105.
[4] Qian Cheng, Zeng Sheng Qin. How to do a good job in the labor employment of enterprise [J]. Personnel Management, 2015 (12): 98-100.
[5] Wang Hua, Lv Yanli, Zhang Jianjun, et al. Research and implementation of labor quota management system in State Grid Corporation of China [J]. Power Information and Communication Technology, 2015, 13 (9): 77-81.
[6] Pan Yanhua, Li Guang, Li Xiangyuan. Association rules analysis of factors affecting shipbuilding cost [J]. Journal of Wuhan University of Technology (Information and Management Engineering Edition), 2017, 39 (3): 353-358.
[7] Jorne V D B, Beliën J, De Bruecker P, et al. Personnel scheduling: a literature review.[C]//IEEE International Conference on Image Processing. IEEE, 2013:1329-1332.
[8] R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases[J]. ACM SIGMOD Rec., vol. 22, no. 2, pp. 207-216, 1993.