Study on Application of Optimum Index Factor in the Electric Power Benchmarking

Wang Yao-ping¹, Tan ye xin² and Song Bin¹,*

¹ School of Electrical Engineering and Automation, Wuhan University, Wuhan, China
² School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan, China

Email: *592276269@qq.com

Abstract. Benchmarking is a widely used method of enterprise management, which can improve the management efficiency as well as the competitiveness of enterprises. In the process of electric benchmarking, a great deal of data that generated by multiple links such as power generation, transmission, substation, distribution, consumption and dispatching are not fully utilized. Accordingly, much effective information are not efficiently obtained. Aiming at this and to give a comprehensive consideration to the various factors of the power system, the method of data mining based on hierarchical optimum index factor is proposed to give priority to the development of a suitable electric benchmarking system to appraise the benchmarking professional index. This method takes various factors into account comprehensively and quantifies the development priority hierarchy of indexes. It is simple in form and also easy to calculate so that the key index can be figured out easily. Thus it can enable electric power companies to better determine the key emphasis in work, improve the efficiency of benchmarking management and fully play the role of benchmarking management.

1. Introduction

Benchmarking is a widely used management tool, which is not only beneficial for improving the product quality and managerial level of enterprises, but also can enhance the core competitiveness and competitive advantages of enterprises. The benchmarking work has been carried out in China in recent years [1], the system and methods of which have been gradually improved, which has effectively promoted the rationality and scientificity of enterprise management [2]. However, due to the complexity of indicators, it is difficult to conduct an in-depth analysis of each indicator. Thus, it is the primary task of benchmarking to make analysis of the key indexes with high efficiency and low difficulty to improve. Considering that the power grid company’s index dimension of standard management evaluation is large [3], the data is heterogeneous [4] and contains a large amount of hidden information [5], it is difficult to find the existing problems from data accurately and quickly with conventional benchmarking management methods [6].

Data mining is based on artificial intelligence, machine learning, pattern recognition, statistics, database and visualization technologies [7], and the main aim of the data mining process is to extract the useful information from the dossier of data and mold it into an understandable structure for future use [8]. The large amount of data generated in the power system contains extremely rich information, but the data obtained by people is only a small part of the information [9], which leads to the fact that
the majority of effective information has not been effectively obtained. In order to get more effective hidden information, power benchmarking data are evaluated from power generation, transmission, substation, distribution, consumption and dispatching. In the meantime, the method of data mining is applied to analyze data more efficiently to assist decision making [10]. Xue W.L. [11], Liu L.D. [12], Liu X.L. [13], Cheng Y.L. [14], and Yang X. [15] respectively expounds the applications of cluster analysis, correlation analysis, variance analysis and other methods in the electric benchmarking. Nevertheless, much work so far has not considered the development priority of index by considering various indexes as a whole. Drawing inspiration from the optimal index factor [16] to select part of the bands for analysis from multiple bands in remote sensing and combined with the idea of data mining, the optimum index method suitable for electric power benchmarking is proposed in this paper to select the key indexes from the complex indexes. This method takes various factors into account comprehensively and quantifies the development priority hierarchy of indexes, which can enable electric power companies to better determine the key emphasis in work and improve the efficiency of benchmarking management.

2. Optimum Index Factor
Considering the similarity between selecting the key index from various index and selecting the optimum band from multiple bands, the method of Optimum Index Factor is used to analyze the priority degree of the benchmarking index in this paper. The Optimum Index Factor (OIF) was developed by Chavez et al. (1982, 1984) as a method for determining the three-band combination that maximizes the variability in a particular multispectral scene. The OIF is based on the amount of total variance and correlation within and between all possible band combinations in the dataset. It is simple in form, easy to calculate, thus make it a very popular mathematical tool [16]. OIF is calculated as shown below:

$$OIF_{ij} = \frac{\sum_{k=1}^{n} S_k}{\sum_{k=1}^{n} |R_{ij}|}$$

(1)

Where: \( N \) is the number of bands of remote sensing images, \( S_i \) is the standard deviation of the i-band and \( R_{ij} \) is the correlation coefficient of the i-band and j-band. For the n-band image, calculate the standard deviation of the single-band image first, then calculate the correlation coefficient matrix between each band. Finally, the OIF index corresponding to all possible band combinations can be calculated, and the advantages and disadvantages of each band combination can be estimated according to the size of the index value.

The band combination corresponding to the maximum OIF index is the optimum band combination.

Deformed appropriately to make it suitable for the benchmarking index, the OIF which can be applied in getting the key index can be obtained. Conditions are supposed to be meet as follow to be the key index: ① The gap of this index between this company and other companies is small so that it is easy to improve the index. ② The improvement of this index can promote the progress of other indexes, that is, the benefit of the improvement of this index is greater. ③ The weight of this index is high, so that the index can be improved to get a higher score. The correlation factors of the combination features of remote sensing image bands mainly include standard deviation and correlation coefficient. Standard deviation represents the degree of variance in data distribution.

The larger the standard deviation, the greater the variance in data distribution. On the contrary, the smaller the standard deviation is, the more centralized the distribution of the data is. According to this, standard deviation can be used as the first criterion for selecting key index. Then correlation coefficient reflects the degree of linear correlation between two variables, thus makes it appropriate to describe the degree of correlation between indexes. So Correlation coefficient is applied as the second criterion for selecting key index. Considering that all index weights have been defined that can be used directly, the original OIF index was modified as shown below:

$$OIF_t = w_t \cdot \sum_{i=1}^{n} R_{ij} / S_i$$

(2)
Where: \( w_i \) is the ith index weight, and \( S_i \) is the standard deviation of the ith index, which can be calculated as: 
\[
S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}},
\]
where \( S \) stands for sample variance, \( x_i \) is the variable, \( \bar{x} \) is the sample mean and \( n \) is the number of samples. \( R_{ij} \) represents the Pearson correlation coefficient between index \( i \) and index \( j \), which can be calculated as shown below:
\[
R = \frac{1}{n-1} \sum_{i=1}^{n} \left( \frac{x_i - \bar{x}}{s_x} \right) \left( \frac{y_i - \bar{y}}{s_y} \right)
\]
(3)
Where: \( n \) is sample size, \( x_i, y_i \) are the individual sample points indexed with \( i \), \( \bar{x}, \bar{y} \) are the sample mean. \( OIF_i \) represents the influence degree of the ith index on the total score.

For all indexes with gaps, the variance of each index is calculated statistically first, and then the correlation coefficient matrix between each index is calculated. The index with smaller variance is easier to improve, and index with larger correlation coefficient can contribute to the improvement of more indexes. After calculating the OIF value of all indexes, sort the OIF value. The index with larger value are supposed to promote preferentially.

3. The example analysis

Based on the data of some company in 2016, the calculation and analysis were carried out to get the development priority hierarchy of indexes which enable the company to find out the key emphasis in work.

According to the above, the variance of each index is calculated first, and then the sum of the correlation coefficients of each index and other indexes is calculated. By summarizing and analyzing these data, the OIF value of all indexes can be calculated. Then sort the OIF value and the development priority hierarchy of indexes are obtained.

Firstly, the standard deviation of each index is calculated according to the score of all the companies in the province. Table 1 is part of the result:

| NO. | Benchmarking Index | Standard deviation |
|-----|--------------------|--------------------|
| 1   | N-1 pass rate of 110~35 kV power grid | 0.019 |
| 2   | N-1 pass rate of 10 kV public distribution line | 0.014 |
| 3   | 220 kV standardized line ratio | 0.173 |
| 4   | Insulation rate of 20 ~ 6 kV urban public distribution network overhead line | 0.177 |
| 5   | Absolute value of load ratio deviation from the optimal range of 110kV system | 1.999 |
| 6   | Maximum load rate of 110(66) ~ 10 kV circuit | 0.293 |
| 7   | Total asset turnover | 0.406 |
| 8   | Turnover of current assets | 0.831 |
| 9   | Payable temporary risk rate | 2.090 |
| 10  | Cash flow liability ratio | 0.260 |

Secondly, Pearson correlation coefficients of each index and other indexes are calculated. Table 2 is the result: (There are 15 indexes in ‘Excellent Performance’, among which ‘The proportion of high-end talents’ has no valid data. And ‘Personal casualty’, ‘Mis-operation’, ‘Transmission accident rate’, ‘Substation accident rate’ and ‘Information event’ are indexes that are easy to change abruptly, thus they are impossible to predict. Besides, as indexes related to accidents and human security, they are supposed to be the focus of the enterprise management. Therefore, these five indexes are outside the scope of the analysis.)
Table 2. The correlation coefficient of index in ‘Excellent Performance’ of some company in 2016.

| Correlation coefficient                                      | Index a | Index b | Index c | Index d | Index e | Index f | Index g |
|-------------------------------------------------------------|---------|---------|---------|---------|---------|---------|---------|
| Full-bore labor productivity (a)                            | 1.00    |         |         |         |         |         |         |
| Full-bore labor cost input-output efficiency index (b)      | 0.90    | 1.00    |         |         |         |         |         |
| Market competition index €                                   | -0.19   | 0.05    | 1.00    |         |         |         |         |
| Cost-to-income ratio (d)                                    | -0.91   | -0.89   | -0.19   | 1.00    |         |         |         |
| Operating maintenance cost growth rate per 10,000 yuan of grid assets € | 0.89    | 0.75    | 0.09    | -0.96   | 1.00    |         |         |
| Business contribution (f)                                   | 0.82    | 0.82    | 0.37    | -0.98   | 0.93    | 1.00    |         |
| EBITDA Profit margin (g)                                    | 0.57    | 0.44    | 0.44    | -0.80   | 0.87    | 0.86    | 1.00    |

A conclusion can be drawn according to Table 2 that the improvements of ‘Business contribution’ and ‘EBITDA Profit margin’ have bigger contribution to the promotion of other indexes, as they have higher correlation coefficient values.

Lastly, the OIF value can be calculated with the results above. Table 3 is the result.

Table 3. The result of OIF value of index in ‘Excellent Performance’ of one company in 2016.

| Benchmarking Index                                           | Standard deviation | Index weight | Sum of correlation coefficient | OIF    |
|--------------------------------------------------------------|--------------------|--------------|--------------------------------|--------|
| Full-bore labor productivity                                | 2.00               | 3.00         | 2.69                           | 4.04   |
| Full-bore labor cost input-output efficiency index          | 1.50               | 2.25         | 2.69                           | 4.04   |
| Cost-to-income ratio                                        | 2.00               | 10.80        | 2.85                           | 15.39  |
| Operating maintenance cost growth rate per 10,000 yuan of grid assets € | 1.50               | 7.20         | 0.27                           | 1.30   |
| Business contribution                                       | 2.00               | 8.10         | 3.23                           | 13.08  |
| EBITDA Profit margin                                        | 2.00               | 8.10         | 2.59                           | 10.49  |
| Electricity alternative market development index             | 0.63               | 2.25         | 4.90                           | 17.42  |

According to the result, it can be found that ‘Electricity alternative market development index’ has the highest OIF value, which indicate this index is of vital importance. However, the company has gotten full score in this index according to the data. Thus this index needs to be maintained attentively. Meanwhile, the OIF value of three indexes reaches higher than 10, and all of them have the room for improvement. The other three indexes with OIF value lower than 10 have the room for improvement as well. Accordingly, the priority development hierarchy diagram is shown in Figure 1.

![Figure 1. The hierarchical map of priority development for indexes of Excellent Performance.](image-url)
By sorting the OIF values, the development priority hierarchy is shown as follows: ‘Cost-to-income ratio’ > ‘EBITDA Profit margin’ > ‘Business contribution’ > ‘Operating maintenance cost growth rate per ten thousand yuan of grid assets’ > ‘Full-bore labor productivity’ > ‘Full-bore labor cost input-output efficiency index’. Considering there is no room for ‘Electricity alternative market development index’ to improve, this index should be maintained attentively.

Analogously, the development priority degree of other indexes can be calculated by using this same method, which is unnecessary to go into details here.

4. Conclusion
This paper uses the method of data mining to select the key index, which enables the electric power company to find out key emphasis in work and get higher grades in benchmarking, and also has good impact on improving the rationality and scientificity of enterprise management.

Meanwhile, this research contributes to improving the objectivity, accuracy, synergy and efficiency of benchmarking management. Nevertheless, more further researches are needed to figure out the inner driving factors of index that can help to fundamentally improve the effect of benchmarking.

References
[1] Cheng H.J, 2017 Research on benchmarking direction of provincial electric power companies. China Power Enterprise Management, 2017(31).
[2] Yang J, 2013 Discussion on the setting of benchmarking index system for county-level power supply enterprises. Morden Enterprise Culture, 2013(11):74-74
[3] Guo Z.Y, 2013 The research of benchmarking based on “Three Intensive and Five System” of Tianjin electric power corporation. Tianjin University, 2013
[4] Li B, Wei Z, Zhu Y et al. 2015 The utility model relates to a benchmarking evaluation system and evaluation method for the comprehensive management of line loss in a county-level power supply enterprise. CN104951866A. 2015.
[5] Yu J, 2015 The research and application of benchmarking in the electric power supply enterprise. North China Electric Power University, 2015.
[6] Liu S.S, Gan X.D, 2016 Research on the method of benchmarking management of electric power enterprises. Engineering Technology, 2016(6):00029-00029.
[7] Rhodes D R, Yu J, Shanker K, et al. 2004 Oncomine: a cancer microarray database and integrated data-mining platform. Neoplasia, 2004, 6(1):1-6.
[8] Shivam A, 2013 Data mining: data mining concepts and techniques. IEEE Conf. 2013.
[9] Wang X, 2016 Research on electric power company operation system based on big data technology. Business, 2016(4):214-214.
[10] Ye J.J, 2006 Research and application of intelligent decision support in power system information management. Zhejiang University, 2006.
[11] Xue W.L, Liu X.L, Li X.L et al. 2013 Application of K-means algorithm in indicators comparison of power industry. Water Resources and Power, 2013(6):240-242.
[12] Liu L.D, 2006 The application of data mining in electric benchmarking. North China Electric Power University (Beijing), 2006.
[13] Liu X.L, 2014 Research on benchmarking for grid enterprises based on data mining. North China Electric Power University, 2014.
[14] Cheng Y.L, Zhang Hong-qin. 2015 The application of improved K-means algorithm in the benchmarking of electric power enterprises. Modern Business Trade Industry, 2015, 36(16):69-71.
[15] Yang X. 2015 Discussion on application of information technology in benchmarking management of electric power enterprises. Modern SOE Research, 2015(18).
[16] Kienast-Brown S, Boettinger J L. 2010 Applying the optimum index factor to multiple data types in soil survey. Digital Soil Mapping. Springer Netherlands, 2010:385-398.