

# Fuzzy comprehensive evaluation of the factors affecting the data quality of power system

Yu Song & Yafei Jiao

*School of Control and Computer Engineering, North China Electric Power University, Baoding, Hebei Province, 071000, China*

**ABSTRACT:** With the development of information technology, large scale intelligent information system has been widely used in the field of electric power enterprise management. This paper combines data quality characteristics and the significant problems that power system faces with, Constructing the index system of quality evaluation of power system data, using fuzzy analytical hierarchy process (FAHP) to calculation index weight .Finally, in the end, an example is given to illustrate the level of index system.

**KEYWORD:** data quality; power system; fuzzy analytical hierarchy process; index weight

## 1 INTRODUCTION

### 1.1 *Data quality in Power system*

With the development of information technology, large scale intelligent information system has been widely used in the field of electric power enterprise management. System will cleaning and pre-treatment, integrated, annotated, summarized and re-organized the data come from heterologous into semantically consistent data storage, become subject oriented, stable and different time of the data set used to support management decision-making process and complex multidimensional queries, the quality of the integrated data become one of the key elements of the effective operation of the intelligent information system.

### 1.2 *Data quality general evaluation indicators*

The evaluation of data quality is generally classified as the following 5 aspects: accuracy, completeness, validity, correlation.

The Accuracy of data: the accuracy of the data is that the data must be truly and accurate to reflect the actual business, including the naming accuracy.

The integrity of the data: data integrity refers to the data is sufficient, there is no missing or omission, the specific said that the data can be fully described the various processes and operations in the business.

The Validity of data: One is to meet the requirements of the time, the other one is to meet the standards or rules, focusing on the availability of data.

The Contextuality of Data: Data does not exist in isolation, there are often a variety of relationships among the data, the data must be able to meet the association between the data, and not be able to contradict each other.

The security of Data there have two aspects about data security , the first is that data do not have the permissions of the user and data access is limited, the concrete manifestation is that data security is security and tamper resistance, the second is the data storage and transmission is not lost, i.e., data with good recovery.

### 1.3 *Data quality problem in Power system*

According to the requirements of data quality, the factors that affect the quality of data could be summarized as follows:

The quality factor of electric power operation data

There have a number of factors affects the quality of data on Power operation link, power transmission and distribution equipment failures, parameter setting, channel condition, weather or an unexpected disaster factors, RTU collection and so on all will bring problems of data quality.

2) The quality factors of market operation data in the process of Market operation, user power meter acquisition errors, power infrastructure failures and power supply company business information is not timely update will have an impact on the quality of the data.

3) The quality factors of production management data

The data produce from the process of the production and management stage can be as the basis as the decision-making of production equipment operation, maintenance and production accident treatment, the power plant use all kinds of relational database, real-time database to storage data efficiently, disk array, load balance, virtualization technology and even private cloud technology, all of them are used to provide services of store and analysis for these data. With the coming of the era of big data, the limits of storage space will bring hidden trouble of data security, and then bring influence to the quality of the data.

## 2 FUZZY EVALUATION METHOD

### 2.1 Set up evaluation index

Fuzzy comprehensive evaluation is a comprehensive evaluation method based on fuzzy mathematics. Based on the fuzzy mathematics, the comprehensive evaluation method is used to transform the qualitative evaluation into quantitative evaluation, Fuzzy comprehensive evaluation first determine the set of evaluation factors  $U$ , And set Reviews, record as  $V = \{v_1, v_2, v_3, v_4, v_5\}$  On behalf of good, preferably, general, bad, poor 5 grades, Determine the weight of indicators, this paper uses FAHP to determine the weight of indicators.

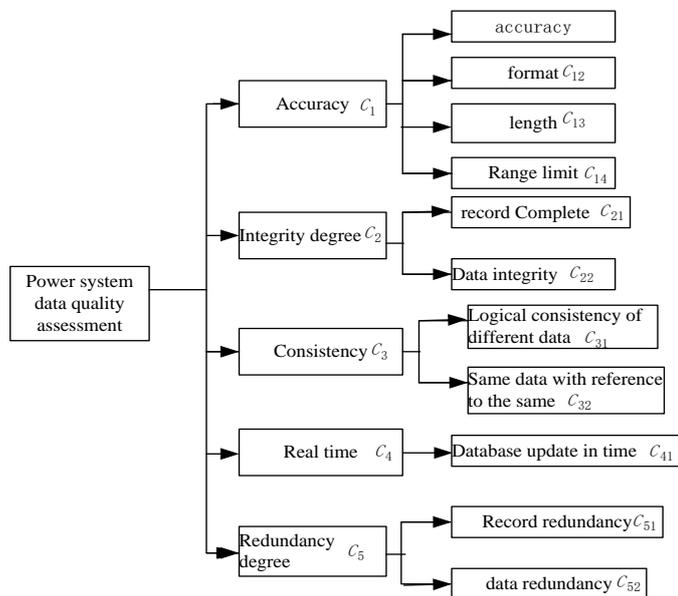


Figure 1. Two level structure of the index.

Based on the established evaluation index system, the evaluation index set is obtained.

This evaluation index is composed by target level  $A$  (data quality of power system) and criterion layer  $B$  (Including power operation data  $B_1$ , Market operation data  $B_2$ , Production management data  $B_3$ ) Index

layer  $C$  (including power transmission and distribution equipment maintenance  $C_{11}$ , Accurate parameter setting  $C_{12}$ , channel is in good condition  $C_{13}$ , prediction of weather or unexpected disaster factors  $C_{14}$ , complete RTU collections  $C_{15}$ , complete collection of user electric meter  $C_{21}$ , power infrastructure troubleshooting  $C_{22}$ , power supply company business information in a timely manner to improve  $C_{23}$ , data structure  $C_{31}$ , data storage security of various departments of production and management  $C_{32}$ ). as shown in figure 1.

### 2.2 Fuzzy evaluation method

Definition 1 Set matrix =  $R = (r_{ij})_{n \times n}$ , if  $0 \leq r_{ij} \leq 1, (i = 1, 2, \dots, n; j = 1, 2, \dots, n)$ , then  $R$  is a fuzzy judgment matrix.

Definition 2 if the fuzzy judgment matrix  $R = (r_{ij})_{n \times n}$ , when  $\forall i, j, k, r_{ij} = r_{ik} - r_{jk} + 0.5$ , The fuzzy judgment matrix  $R$  is a fuzzy consistent matrix.

Fuzzy judgment matrix  $R = (r_{ij})_{n \times n}$  Fuzzy consistent matrix, so:

(1) when  $\forall i (i = 1, 2, \dots, n), r_{ii} = 0.5$ ;

(2) when  $\forall i, j (i, j = 1, 2, \dots, n), r_{ij} + r_{ji} = 1$ ;

(3) The sum of the elements of the  $i$  and  $i$  columns of the  $R$  is  $n$ .

structure fuzzy judgment matrix

Determine two level rules sets, First level rule set  $U = \{U_1, U_2, \dots, U_m\}$ ,  $m$  is a number of first order rules,

Two level rule set  $U_i = \{a_1^{(i)}, a_2^{(i)}, \dots, a_{n_i}^{(i)}\}$   $n_i$  is the number of two level rules under the first level rule  $i$ , the important degree between elements  $a_i^{(i)}, a_j^{(i)}$  on rule  $i$  Relative to rule  $i$  obtain Fuzzy judgment matrix:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{n4} \end{bmatrix} = (r_{ij})_{n \times n}$$

$r_{ij}$  is the membership degree of elements  $a_i^{(i)}, a_j^{(i)}$ .

In the two level structure model, each rule is judged according to the above description. Obtaining the fuzzy judgment matrix of the upper level  $R, R_1, R_2, \dots, R_n$ , In order to make a quantitative description of the comparison between the two factors in accordance with the unified standard, it will use 0.1-0.9 scale method.

Set up fuzzy consistent matrix and calculation rule weight

Whether a fuzzy judgment matrix is a fuzzy consistent matrix, otherwise, the fuzzy judgment matrix is adjusted to the fuzzy consistent matrix, Line transformation:

$$r_{ij}' = (r_i' - r_j') / 2n + 0.5 \quad (1)$$

$$r_i' = \sum_{j=1}^n r_{ij}, i = 1, 2, \dots, n. \quad (2)$$

At this time, the fuzzy judgment moment is consistent.

The weight  $w_i^{(i)}$  of each factor  $a_i^{(i)}$  is calculated. Weight by sorting method:

$$w_i^{(i)} = \frac{1}{n} - \frac{1}{2a} + \frac{1}{na} \sum_{k=1}^n r'_{ik}; i = 1, 2, \dots, n \quad (3)$$

Where  $a$  is the parameter,  $a \geq \frac{n-1}{2}$ .

Here take  $a = \frac{n-1}{2}$ , Get the weight set of the first level rule  $W = \{w_1, w_2, \dots, w_m\}$ ,  $m$  is a number of first order rules, two level rule weight set  $w_i = \{w_1^{(i)}, w_2^{(i)}, \dots, w_{n_i}^{(i)}\}$ ,  $n_i$  is the number of the two level rules  $i$  under the first order rule.

Set up fuzzy consistent matrix and calculation rule weight

According to the evaluation index of power system, the rule set in this paper  $E = \{E_1, E_2, E_3, E_4, E_5\}$ , One of the sub rules,  $n$  is the number two level rules of sub rule set  $E_1$ , Grade one rule:

$$S_i = \frac{\sum_{j=1}^n w_j^{(i)} \times s_{jc}^{(i)}}{\sum_{j=1}^n w_j^{(i)}} \quad (4)$$

$i$  is the first level rules,  $j$  is the the two level rules under the first level rule  $i$ . Final result:

$$S = \frac{\sum_{i=1}^k W_i \times S_i}{\sum_{i=1}^k W_i} \quad (5)$$

### 3 EXAMPLE ANALYSIS

#### 3.1 Evaluation index system

Based on the analysis of the data quality factors of the power system, the hierarchical structure of data quality evaluation of power system is constructed according to the principle of hierarchy and scientific objectivity.

The target layer  $A$  includes three first level evaluation indexes,  $A = \{B_1, B_2, B_3\} = \{\text{Electric power operation, market operation, production management}\}$ . Criteria layer includes ten second levels evaluation indexes. Power operation stage' evaluating indicator include power transmission、distribution equipment failure, accurate parameter setting, channel is in good condition, prediction of weather or unexpected disaster factors, complete RTU collections,

$$B_1 = \{C_{11}, C_{12}, C_{13}, C_{14}, C_{15}\},$$

Evaluation index of market operation stage include the complete collection of user electric meter, power infrastructure troubleshooting, power supply company business information in a timely manner to improve  $B_2 = \{C_{21}, C_{22}, C_{23}\}$ . Evaluation index of production management includes data structure and data storage security of various departments of production and management,  $B_3 = \{C_{31}, C_{32}\}$ .

#### 3.2 Set of reviews

According to the above to establish the index system that impact data quality of power system, will be randomly selecting power enterprise executives, data analysts, database administrators 10 persons to make a data quality survey. According to the professional knowledge and their own experience to choose the indicators should belong to the category, the number of indicators to select the number / total number as the degree of membership, it determines the quality grade of the power operation data, as shown in table 2.

Table 2 the quality grade of the power operation data

First or- der index	second or- der index	Comment grade				
		1	2	3	4	5
$B_1$	$C_{11}$	0.6	0.3	0.1	0	0
	$C_{12}$	0	0	0.3	0.3	0.4
	$C_{13}$	0	0.5	0.4	0.1	0
	$C_{14}$	0	0	0.4	0.3	0.3
	$C_{15}$	0	0.2	0.5	0.3	0

And so on, to determine the grade of member of ship matrix the market operation data and production management data of, respectively  $R_1, R_2$ :

$$B_2 = \begin{bmatrix} 0 & 0.3 & 0.6 & 0.1 & 0 \\ 0.4 & 0.6 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \end{bmatrix} B_3 = \begin{bmatrix} 0 & 0 & 0.6 & 0.3 & 0.1 \\ 0.7 & 0.3 & 0 & 0 & 0 \end{bmatrix}$$

#### 3.3 Fuzzy hierarchy process to determine the weight

According to the steps given in the last section, the weight of the index is obtained. Firstly, 4 fuzzy judgments matrices are constructed

$$R = \begin{bmatrix} 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix} R_1 = \begin{bmatrix} 0.5 & 0.9 & 0.8 & 0.9 & 0.7 \\ 0.1 & 0.5 & 0.2 & 0.5 & 0.3 \\ 0.2 & 0.8 & 0.5 & 0.8 & 0.6 \\ 0.1 & 0.5 & 0.2 & 0.5 & 0.4 \\ 0.3 & 0.7 & 0.4 & 0.6 & 0.5 \end{bmatrix}$$

$$R_2 = \begin{bmatrix} 0.5 & 0.1 & 0.4 \\ 0.9 & 0.5 & 0.8 \\ 0.6 & 0.2 & 0.5 \end{bmatrix} R_3 = \begin{bmatrix} 0.5 & 0.3 \\ 0.7 & 0.5 \end{bmatrix}$$

The weights are obtained by the fuzzy analytic hierarchy process:

$$w = (0.3333, 0.3333, 0.3333)$$

$$w_1 = (0.3215, 0.1193, 0.2263, 0.1286, 0.2043)$$

$$w_2 = (0.2051, 0.5128, 0.2827)$$

$$w_3 = (0.3956, 0.6044)$$

### 3.4 Fuzzy hierarchy comprehensive evaluation

$$A_1 = w_1 \cdot B_1 = (0.1929, 0.2505, 0.3121, 0.1583, 0.0863)$$

$$A_2 = w_2 \cdot B_2 = (0.2051, 0.5106, 0.2644, 0.0205, 0)$$

$$A_3 = w_3 \cdot B_3 = (0.4231, 0.1813, 0.2379, 0.119, 0.0397)$$

$$A = [A_1, A_2, A_3]$$

$$S = w \cdot A = (0.2737, 0.3141, 0.2714, 0.0993, 0.042)$$

$$s = 0.2737 \times 1 + 0.3141 \times 2 + 0.2714 \times 3 + 0.0993 \times 4 + 0.042 \times 5 = 2.3233$$

The influence of index system on the data quality of power system is between 2~3, and the influence grade is "general", so the index system needs to be further summarized and improved.

## 4 CONCLUSION

There are many factors that influence the quality of power system data, which is determined by the complexity of the power system, only controlling the data quality effectively, it should have further attention and control on the elements influence on the data quality, evaluate the pros and cons of the quality of the data, and evaluate the level of the factors influencing the quality of the data value can be more efficient use of data and data mining.

## REFERENCES

- Ahmad, et al. 2014. Evaluating supplier quality performance using fuzzy analytical hierarchy process. International Conference on Quantitative Sciences and Its Applications. AIP Publishing (ICOQSIA); Langkawi 15-17 August 2014. Kedah Malaysia.
- Chen, H., et al. 2014. Analysis of Related Factors Influencing Reliability of Railway Signaling Systems Based on Fuzzy Analytical Hierarchy Process. In Proceedings of the 2013 International Conference on Electrical and Information Technologies for Rail Transportation (EITRT). Springer Berlin Heidelberg.
- Hezhan, Y. 2014. Application of Fuzzy Analytical Hierarchy Process in innovation education quality evaluation of higher education institution. In Information Management, Innovation Management and Industrial Engineering (ICIII), 25-26 October 2014. Xi'an China.
- JANJIC, A., et al. 2015. Renewable energy integration in smart grids-multicriteria assessment using the fuzzy analytical hierarchy process. Turkish Journal of Electrical Engineering & Computer Sciences, 23(6), 1896-1912.
- Kaur, P. 2014. Selection of Vendor Based on Intuitionistic Fuzzy Analytical Hierarchy Process. Advances in Operations Research, 3(23)14-20.

- Li, J., et al. 2009. Fuzzy Analytical Hierarchy Process for Evaluating Online Bookstores. In 2009 International Conference on Management and Service Science (MASS). 2009
- Mousavi, et al. 2014. A new hesitant fuzzy analytical hierarchy process method for decision-making problems under uncertainty. In Industrial Engineering and Engineering Management (IEEM), 09-12 December 2014. Sunway Resort Hotel & Spa Malaysia.
- Samvedi, A., et al. 2012. An integrated approach for machine tool selection using fuzzy analytical hierarchy process and grey relational analysis. International Journal of Production Research, 50(12), 3211-3221.
- Tavana, M., et al. 2010. A multiple criteria decision support system for autonomous underwater vehicle mission planning and control. International Journal of Operational Research, 7(2), 216-239.
- Yuyang, W., et al. 2015. Fuzzy Analytical Hierarchy Process Methods for Evaluating the Comfort of textiles. In Measuring Technology and Mechatronics Automation (ICMTMA), 13-14 June 2015. Nab Chang China.