Analysis of Power Service Bandwidth based on AHP and Clustering

Chen Yiming¹, Zhu Hechun¹, Yao Shiying¹, Huang Yunyang², Liao Jingru² and Xu Jie²

¹ Planning and Reviewing Center State Grid Sichuan Economic Research Institute, ² School of Information and Communication Engineering, University of Electronic Science and Technology of China

E-mail: xuj@uestc.edu.cn

Abstract: The current method of analyzing the relationship between communication service and bandwidth requirement in the power communication network is simple. It uses a single formula with predetermined constant parameters to calculate the bandwidth of each communication service for a given site which can not show the relative bandwidth demand of different services. This paper proposes a combined AHP and hierarchical clustering method to analyze the relative bandwidth requirement of each service. This method is based on AHP to analyze the impact of different factors on bandwidth and use hierarchical clustering method to aggregate services with similar communication needs into the same class, and it shows different services’ relative bandwidth requirement clearly. Experiments on data of various site shows that all the services with similar bandwidth requirements are grouped into the same class.

1. Introduction

With the development of the electric power industry, its communication network become much more important, which ensures the security of operation and production in the power grid [1].

As the power grid business constructing, the demand for interaction between different professions and different departments has risen sharply. However, the current method for the State Grid to calculate the bandwidth of the service is very simple. Since it just uses a single formula to calculate the bandwidth of a particular service at a site, it is difficult to meet the needs of power communication networks development. This paper proposes a combined analytic hierarchy process and hierarchical clustering method to analyse the relationship between service and bandwidth in a certain site.

At present, the research on bandwidth and business relationship mainly adopts the method based on the importance of each service. Hierarchical Structure Analysis (AHP) is a typical method to analyse the importance of power communication services. The method constructs the judgment matrix and do the consistency check to get the weights of each layer, so as to determine the final sort [2].

Wang Xiaoqing et al. [3] analyse the requirements of different services’ communication. They construct judgment matrices and calculate the importance of each service and the demand of the bandwidth is obtained finally. Li Chang Chao et al. [4] propose an AHP algorithm which decompose the asymmetric hierarchical model into several structural models and layered the criterion layers.

However, when it comes to the analysing the requirement of the services’ bandwidth in a certain site, those studies perform not very well. It is because the requirements of the same service in different
sites are different and the amount of data transmitted by the same service at different sites is also
different.

Secondly, the calculation of traditional AHP is cumbersome. It is because AHP algorithm needs to
repeat the consistency check when we calculating the index weights. In addition, some studies show
that the weights of some services are similar.

Therefore, we can find that services with similar bandwidth requirements can be classified into one
class. This paper proposes a combined model algorithm of AHP and clustering algorithm. AHP is used
to calculate the index weights, and the clustering algorithm is used to classify the services with similar
communication requirements into one class. This method provides a reference for future dynamic
bandwidth allocation or communication bandwidth allocation in the event of an emergency.

2. Theoretical background

2.1. Overview

According to the experts’ opinion, we get the judgment matrix of the communication indicator and the
communication requirements of each service. We do consistency check to ensure the reliability of the
judgment matrix. Then, we calculate the relative weight of the index. After that we obtain the
weighted data of service communication demand by calculating the index’s weight. Finally, the data is
clustered by the agglomerative hierarchical clustering. The norms of the center point are calculated
and sorted to obtain the results. Algorithm flow is shown in Figure 1.

![Algorithm flow.](image)

2.2. Calculate indicator weights based on AHP

2.2.1. Construct the judgment matrix. Considering the expert opinion and construct the AHP judgment
matrix, we adopt the ratio of the importance degree of the two indicators to indicate the relative
importance of the two indicators[5].

Since the weights are subjectively set, the results may be random. For example, we can assume that
A is more important than B, B is more important than C, and C is more important than A. It is
obviously contrary to common sense. Thus the consistency check is introduced internally to ensure
that the error is within the acceptable range [6].

2.2.2. Consistency check. According to AHP theory we define the consistency index (CI):

\[ CI = \frac{\lambda_{max} - n}{n-1} \]  

Where \( \lambda_{max} \) is the maximum eigenvalue of the judgment matrix, and \( n \) is the dimension of the
judgment matrix. And we define the consistency ratio (CR):

\[ CR = \frac{CI}{RI} \]  

Where RI is the average value of CI for random matrices and we only accept the matrix as a
consistent one if CR < 0.1[7].

2.2.3. Calculating indicator weights. We calculate the maximum eigenvalue of the judgment matrix
and the corresponding eigenvector. Each element in \( \omega_{max} \) is the weights of the indicators.
2.3. Agglomerative Hierarchical Clustering
Agglomerative Hierarchical Clustering is a method that takes each sample as a class at the beginning, iteratively iterates, clusters according to a rule, and finally aggregates all the samples into one class [8].

2.3.1. AHC flow. Step 1: Input the number of clusters and data set. Step 2: Take each sample as one class. Step 3: Calculate the distance of each class to other classes, and group the two closest classes into one class. Step 4: Repeat step 3 when the number of current clusters is less than the number of clusters entered, otherwise enter step 5. Step 5: Output

2.3.2. Weight-based center point norm for each class. We obtain the norm by using the following formula:

\[ S = \left( \sum_{i=1}^{n} \left( \frac{1}{m} \sum_{j=1}^{m} \beta_{ij} \right)^2 \right)^{1/2} \]  

(3)

Where \( S \) is the norm of the center point of the class, \( \beta_{ij} \) is the weighted demand for the jth service on ith indicator in the class, \( n \) is the number of indicators, \( m \) is the number of the services in the class.

3. Experiment and analysis
This experiment uses data of the communication requirements of the services in some substation which provided by Power Grid Corporation.

3.1. Indicator selection
According to the state grid regulations, indicators for measuring the quality of communication services include reliability, security, and real-time. However, the basic bandwidth of the service transmission is also a factor that affects the bandwidth allocation of the communication service. Therefore, we need to take the data volume into consideration. The indicators is shown in Figure 2.

![Figure 2. Power service communication indicator.](image)

3.2. Indicator weights
According to the opinions of experts, we set judgment matrix \( F \) as follow:

\[
F = \begin{bmatrix}
1.00 & 1.20 & 1.50 & 1.80 \\
0.83 & 1.00 & 1.33 & 1.50 \\
0.66 & 0.80 & 1.00 & 1.20 \\
0.55 & 0.66 & 0.83 & 1.00
\end{bmatrix}
\]

After calculation, we obtain the results vector \( (0.32, 0.27, 0.21, 0.18) \), the element of the vector are Data-Volume, Real-time, Reliability, Security.

3.3. Consistency Check
We use equation (1) and equation (2) to obtain CR and CI, the results are: \( CI = 4.16 \times 10^{-5} \), \( CR = 4.6 \times 10^{-5} \). Since \( CR < 0.1 \), the judgment matrix is a consistent matrix.
3.4. Agglomerative Hierarchical Clustering

3.4.1. Data. According to the data and business analysis report provided by the power company, we use five scales, namely “very low”, “low”, “medium”, “high” and “very high” to describe the corresponding indicators and quantify them as 1, 2, 3, 4, 5. Table 1 and Table 2 show the bandwidth demand of services scales for some services in the 110kv substations and 35kv substations respectively.

| Service             | Real-time | Reliability | Security | Data-Volume |
|---------------------|-----------|-------------|----------|-------------|
| Relay protection    | 5         | 5           | 5        | 4           |
| OMS                 | 1         | 1           | 1        | 1           |
| Dispatching telephone | 3       | 3           | 4        | 3           |

Table 1. 110kV substation

| Service             | Real-time | Reliability | Security | Data-Volume |
|---------------------|-----------|-------------|----------|-------------|
| Relay protection    | 4         | 5           | 3        | 2           |
| OMS                 | 1         | 1           | 1        | 1           |
| Dispatching telephone | 4       | 5           | 3        | 5           |

Table 2. 35kV substation

3.4.2. Simulation and results. The result shows that the method performs well when we set the number of clusters to 3. If the number of clusters is set to 4 or more, it may cause one class contains only one service, or the difference between the norms of the center points of the two clusters is small. That means the requirement of services in two clusters is similar, and it means the result may not be very suitable.

After that, we obtain the weighted data based on the weight of the index. Then the weighted data is clustered by hierarchical clustering method, and the norm of the center point of each class is obtained according to equation (3), and sorted from large to small. We define the Cluster with the largest norm of the center point as Cluster 1, the second largest one as Cluster 2, the lowest one as Cluster 3. Table 3 and table 4 show the results of the services in 35kV substation and 110kV substation respectively.

| Cluster number | Service                      | Norm of center point |
|----------------|------------------------------|----------------------|
| Cluster 1      | Dispatching telephone        | 3.91                 |
|                | Distribution automation      |                      |
|                | Power information collection |                      |
| Cluster 2      | Dispatching data network     | 1.83                 |
|                | Administration telephone     |                      |
|                | Relay protection             |                      |
| Cluster 3      | PMS, OMS, OA, GIS           | 0.87                 |

Table 3. Results of 35kV substation clustering

| Cluster number | Service                      | Norm of center point |
|----------------|------------------------------|----------------------|
| Cluster 1      | Relay protection,            | 4.12                 |
|                | Dispatching automation      |                      |
| Cluster 2      | Dispatching telephone,       | 2.04                 |
|                | Dispatching data network     |                      |
| Cluster 3      | Substation video surveillance,| 0.91                 |
|                | PMS, OMS, OA,               |                      |

Table 4. Results of 110kV substation clustering
As Table 3 and 4 show, the method aggregates services with similar bandwidth requirements into one class. Services in Cluster 1 is the highest, Services in Cluster 2 is medium, Services in Cluster 3 is the lowest. Comparing Table 1 with Table 2, we found that the bandwidth requirement of one service in different sites is different. For example, dispatching telephone has the most bandwidth requirement in 35kv substation relatively, but in 110kV substation it does not the most.

As Table 3 and 4 show, we can find that Relay protection bandwidth’s bandwidth requirement is greater than Dispatching telephone’s bandwidth requirement in 110kV substation. But in 35kV substation, we can find that Relay protection bandwidth’s bandwidth requirement is lower than Dispatching telephone’ bandwidth requirement. Thus one service has different bandwidth requirements at different sites.

3.4.3. Results Verifying. Table 5 and Table 6 show the bandwidth and clustering results for some services in the 110 kV substation and 35kV substation respectively.

| Service             | Bandwidth | Cluster number |
|---------------------|-----------|----------------|
| Relay protection    | 8.1M      | Cluster 1      |
| Distribution automation | 7.9M | Cluster 1      |
| Dispatching telephone | 4.1M  | Cluster 2      |
| Dispatching data network | 4.05M | Cluster 2      |
| OMS                 | 0.125M    | Cluster 3      |
| OA                  | 0.25M     | Cluster 3      |
| GIS                 | 0.2M      | Cluster 3      |

Table 5. 110 kV substation

| Service              | Bandwidth | Cluster number |
|----------------------|-----------|----------------|
| Power information collection | 4.0M | Cluster 1 |
| Dispatching telephone   | 3.95M    | Cluster 1 |
| Relay protection       | 2.05M    | Cluster 2 |
| Dispatching data network | 2.0M  | Cluster 2 |
| OMS                  | 0.125M   | Cluster 3 |
| OA                   | 0.125M   | Cluster 3 |

Table 6. 35 kV substation

As Table 5 and Table 6 show. We can find that services in the same class have similar actual bandwidth. For example, the difference of different services’ actual bandwidth in one cluster in 110 kV substation does not exceed 0.2M, and in 35 kV substation the difference is not more than 0.05M. In addition, the difference of actual bandwidth of services in different classes is quite large. For example, in 35kV substation, the bandwidth of services in Cluster1 is about twice bandwidth of those in Cluster 2 which is thirty times that in Cluster 3. Apart from that, we can find that Relay protection’s actual bandwidth is 8.1M which is the largest one in 110 kV substation, but in 35kV substation actual bandwidth of Relay protection is 2.05M which is half of Power information collection. The result proves that the actual bandwidth usage of the same service at different sites is different.

4. Conclusion
Researches on the use of the combined AHP and hierarchical clustering method in analysing the bandwidth requirement of services are presented in the paper. AHP is used as the base to obtain the weights of different indicators. Then we use agglomerative clustering algorithm to aggregate services with similar communication needs into one class. We use the data provided by the grid company to verify the algorithm. The results show that the combination algorithm of AHP and Agglomerative hierarchical clustering is an algorithm that can describe the relative demand of service bandwidth in a certain site.
Acknowledgments
This work was supported by State Grid Science and Technique Research Program.

References
[1] Wang H Q, Liu Y J. 2016 Analysis of the construction of smart grid communication management system Science and Technology Innovation and Application, vol 30 pp 217.
[2] Amandeep, F. Mohammad and V. Yadav, 2015 Automatic decision making for multi-criteria load balancing in cloud environment using AHP International Conference on Computing, Communication & Automation, Noida, pp. 569-576.
[3] Wang X Q et al. 2014 Automatic decision making for multi-criteria load balancing in cloud environment using AHP Information Systems Engineering vol 2 pp 21-21
[4] Li C C, Kang Z J, Yu H G, Zheng S C, Ri K H. 2019 Key node identification of power communication network considering the importance of power business Journal of Electrical Engineering: pp 1-11.
[5] Wu Y F, Cui Y P, Hu J W. 2019 Alarm Processing Method Based on Hierarchical Clustering Computer Science vol 04 pp 203-209.
[6] Xu J F, Wang J J, Zhu K L, Zhang P H, Ma Y F 2018 Credit index measurement method for android application security based on AHP Journal of Tsinghua University. vol 58 No.2 pp 131-136
[7] Ali K, Ullviye P, Cansu D 2017 A new approximation for risk assessment using the ahp and fine kinney methodologies Safety Science vol 91 pp 24-32.
[8] Adhitya N, Mahendra A H P, Heru A S, Junta Z, Ardytha L, Ayu P 2018 Determining The Senior High School Major Using Agglomerative Hierarchial Clustering Algorithm 2018 International Seminar on Application for Technology of Information and Communication, Semarang, pp 225-228.