Research on Prediction Model of Line Loss Rate in Transformer District Based on LM Numerical Optimization and BP Neural Network

Yingmei Zhang*, Genghuang Yang*, Xiayi Hao and Shun Yao

School of Automation and Electrical Engineering, Tianjin University of Technology and Education, Tianjin 300222, China

*Corresponding author e-mail: ygenghuang@126.com, the_colourful_sky@163.com, tutezdh1203@163.com, admire95@163.com

Abstract. Due to the complex network structure of the low-voltage distribution network, the backwardness of electrical equipment and the large number of nodes, it is difficult to calculate the line loss rate accurately. The author established a simple linear loss rate prediction model with higher accuracy. Firstly, according to the actual power operation data, the typical characteristic parameters with great influence are obtained. Secondly, the K-means clustering algorithm is applied to classify each transformer district according to the different power consumption characteristics. Finally, the BP neural network model based on LM numerical optimization algorithm is designed and applied to predict line loss rate for data class of each transformer district. 1026 transformer district with complete power operation data in a region was selected as an example to verify the accuracy of the proposed prediction method, which provides data and theoretical basis for the loss reduction measures.

1. Introduction

Line loss rate is the percentage of the line loss load in the power network that accounts for the power supply load of the power network [1]. As an important technical and economic indicator of the power system, it is also a comprehensive technical and economic indicator to measure the level of operating and management of power companies. At present, the power grid has larger motivation to reduce line loss, especially in the low-voltage distribution network. But due to the complex network structure of the low-voltage distribution network, the backwardness of electrical equipment and the large number of nodes, it is difficult to calculate line loss rate accurately to meet the requirement of power grid running. Based on the above description, it is of great practical significance to propose a fast and accurate method for calculating the line loss rate in the transformer district.

The author established a simple linear loss rate prediction model with higher accuracy. Proceed as follows:

1. Select typical characteristic parameters with great influence.

2. Apply multidimensional K-means clustering algorithm. Because the algorithm is simple and efficient, it can process large data categories quickly and has good flexibility.

3. For each data category, use the neural network model of LM numerical optimization algorithm to predict the line loss rate.
2. Determine typical characteristic parameters
The parameters reflecting the operation of the line include the supply power $X_1$ (kW·h), the lost power $X_2$ (kW·h), and the line loss $X_3$ (kW). The parameters reflecting the average load level include the load rate $X_4$ (%) and the number of low-voltage users $X_5$. The range and unit of the above five typical characteristic parameters are different. To ensure the accuracy of the prediction results, the original data needs to be normalized by standard, so that it is not affected by the dimension.

3. K-means algorithm clustering
The central idea of the algorithm is the minimization of the clustering criterion function, which is the sum of the squares of the distances from each sample point in the cluster category to the center of the cluster. This paper selects the following clustering criteria function:

$$J_j = \sum_{x \in S_j(k)} \|x - z_j(k + 1)\|^2, \quad j = 1, 2, \ldots, K \quad (1)$$

The algorithm needs to set the number of clusters $K$ in advance. Due to the uncertainty of the clustering results, the paper selects the CH index to determine the optimal number of clusters. The index describes the tightness by the intra-class dispersion matrix, and the inter-class dispersion matrix describes Separation, the indicator is defined as follows:

$$CH(k) = \frac{trB(k)/(k-1)}{trW(k)/(n-k)} \quad (2)$$

Where $n$ is the number of clusters, $k$ is the current class, $trB(k)$ is the trace of the distance between the classes, and $trW(k)$ is the trace of the intra-class dispersion matrix. The higher the value of CH is, the closer the class itself is, and the more scattered the class is, the better the clustering result is.

4. BP neural network model based on LM numerical optimization algorithm
Traditional BP neural network has slow convergence and training, but is is likely to fall into local minimum point [2]. In this paper, the model based on LM numerical optimization algorithm is used to predict the line loss rate. The structure of the traditional BP neural network is shown in Fig. 1.

The learning process of BP neural network includes forward propagation of signals and back propagation of errors [3]. The forward propagation process is that the input vector $X = (x_1, x_2, \ldots, x_m)$ is processed from the input layer by the hidden layer, the processed vector is processed as the input of the output layer by the output layer, and the output vector $O = (o_1, o_2, \ldots, o_l)$ is finally obtained. If there is an error between the output vector and the expected vector $D = (d_1, d_2, \ldots, d_l)$, the error back propagation is performed, that is, the error after the output is used to estimate the error of the hidden layer, and thus gradually reversed. The weight $(wij, wjk)$ and the threshold $(bij, bjk)$ of each neuron are adjusted. This is the learning and training process of the neural network. When the error reaches the category range or reaches the number of learning, the training stops.
Traditional BP neural network uses the gradient descent method to adjust the weights and thresholds between layers, the error is continuously reduced. The optimized algorithm corrects the weights and thresholds of each layer by continuously solving the minimum value of the error \[4\]. Numerical optimization is to avoid calculating the Hessian matrix, express it with the following formula.

\[
H = J^T(k)J(k) \tag{3}
\]

The expression of the gradient vector is

\[
g = J^T(k)e(k) \tag{4}
\]

In (4), \(J(k)\) is the Jacobi matrix containing the first derivative of the network error versus threshold and weight, and \(e(k)\) is the error vector. The weights and thresholds are corrected by the following formula

\[
x(k+1) = x(k) - [J^T(k)J(k) + \mu J(k)]^{-1}J^T(k)e(k) \tag{5}
\]

When correcting the weight, \(x(k+1) = w(k+1), x(k) = w(k)\).
When correcting the threshold, \(x(k+1) = b(k+1), x(k) = b(k)\).

The optimized method can greatly improve the convergence speed, which is superior to the traditional method from the accuracy of approximation and the accuracy of training results \[5\]. But the method needs to calculate the second derivative of the error function, which increases the computational complexity. The algorithm is suitable for small and medium-sized networks with less complicated structures.

5. Application of prediction models
There is an example of 1026 transformer district with complete power running data in a region which is used to verify the correctness and feasibility of the prediction model, which proves the accuracy of the prediction model.

In order to facilitate the comparison, the incomplete and the non-conformity data are excluded from the samples of 1026 transformer district. The left 1000 complete station samples are selected, and each sample contains 5 independent variables and 1 dependent variable. The independent variable is the supply power \(X_1\) (kW·h), the lost power \(X_2\) (kW·h), and the line loss \(X_3\) (kW), load rate \(X_4\) (%) and the number of low-voltage users \(X_5\), and the dependent variable is \(R\) (%) line loss rate.
The number of clusters is tried to set as 2 to 6, and calculate the CH value. The change line graph is shown in Fig.2. When the number of clusters is 5, the CH value is the largest and the clustering effect is the best. Therefore, the number of K-means clusters is set to 5.

K-means multidimensional clustering analysis was performed on the samples. The clustering results are shown in Table 1. There are 315 samples in the 1\textsuperscript{st} category, and 51 samples in the 2\textsuperscript{nd} category, and 91 samples in the 3\textsuperscript{rd} category, and 128 samples in the 4\textsuperscript{th} category, and 415 samples in the 5\textsuperscript{th} category, with a total of 1000 samples.

According to the above classification results, the optimized model needs to be trained by 5 categories of samples separately. In order to ensure the accuracy of training results in each training process, it is necessary to standardize the samples of each category, and the processed data obey the normal distribution. It is divided into three parts for different purposes, one half for training, one quarter for testing, and the remaining quarter for validating. The error curve of each category is plotted in the training process and the linear regression result curve after training.

The first category of sample error curve is shown in Fig.3, and the linear regression result curve is shown in Fig.4.

From Fig.3, the verification error and the test error gradually become stable and tend to zero after the 5\textsuperscript{th} iteration, and the two error trends are basically the same, which indicates that the sample category is more reasonable. Fig.4 shows that the correlation coefficient of linear regression reaches 0.9784, and the training results of the network are better.

The second category of sample error curve is shown in Fig.5, and the linear regression result curve is shown in Fig.6.
Fig. 5 shows that the test error eventually stabilizes, but the verification error is not stable enough. Although the general trend is similar, the classification of this sample category is not reasonable enough. It has a certain potential for improvement. From Fig. 6, the linear regression correlation coefficient finally reaches 0.9642, which is larger than the 1st category of sample, but the training result is still good. This fully reflects that the superiority of the optimized neural network model is reasonable, when sample data has large individual errors in the input samples, input and output laws of the network are also difficult to be affected.

The 3rd category of sample error curve is shown in Fig. 7, and the linear regression result curve is shown in Fig. 8.

Fig. 7 shows that the test error tends to be stable after 6 iterations, but there is a certain error. The error of the verification error is large in the initial iteration but the final error becomes stable and the error is small. From Fig. 8, the correlation coefficient of linear regression is 0.9365, which is larger than the 2nd and 2nd samples.
The fourth category of sample error curve is shown in Fig.9, and the linear regression result curve is shown in Fig.10.

Fig.9 shows that the test error and verification error tend to zero very steadily after 10 iterations, the trend of change is consistent with each other. Sample category division is very reasonable. From Fig.10, linear regression correlation coefficient finally reaches 0.9998, and the training result is very good. This shows that the classification of this sample category is very reasonable, whereas this model may have a larger error when changing the input sample.

The 5th category of sample error curve is shown in Fig.11, and the linear regression result curve is shown in Fig.12.

Fig.11 shows that the trend of the test error and the verification error is very consistent, but the test error is stable less than 0.2 after 10 iterations, whereas confirmation error tends to zero. It is more stable than the 2nd and 3rd categories of error changes. Although the number of samples is large, from Fig.12, the linear regression correlation coefficient can reach 0.9564, the training error is small, and the result can be predicted well.

Combining the error and linear regression results curves of the above 5 categories, the overall error of the prediction model is small and relatively stable. The linear regression correlation coefficient can reach above 0.93. The model has good training results and can effectively predict the line loss rate.
6. Conclusion
In this paper, a linear loss rate prediction model is proposed and applied to the actual power running data of the transformer district. The following conclusions can be obtained through the verification process:

(1) Selecting typical feature parameters can effectively reduce the input dimensions of clustering and neural networks, which simplifies the calculation process.

(2) K-means clustering is performed on the samples with different electricity consumption behaviors before applying the optimized neural network prediction model. This effectively improves the accuracy and convergence speed of the prediction model, which is a good line loss rate prediction model. The model has large practical significance.

References
[1] Wang Haiyan, Zeng Jiang, Liu Gang. Enlightenment of foreign distribution network automation construction mode to China's distribution network construction [J]. Power System Protection and Control, 2009, 37 (11): 125-129+134.

[2] Feng Tianrui. Research on medium voltage distribution network line loss core index system and line loss model with characteristics of fuzzy clustering [D]. South China University of Technology, 2016.

[3] Wen Fushuan, Han Zhenxiang. The calculation of energy losses in distribution systems based upon a clustering algorithm and an artificial neural network model [J]. Proceeding of the CSEE, 1993, 13 (3): 41-50 (in Chinese).

[4] Xin Kaiyuan, Yang Yuhua, Chen Fu. An advanced algorithm based on combination of GA with BP to energy loss of distribution system [J]. Proceeding of the CSEE, 2002, 22 (2): 79-82 (in Chinese).

[5] Ding Shuo, Chang Xiaoheng, Wu Qinghui, Yang Youlin. Comparative Study of approximation performance of numerical optimization improved BP neural network [J]. Shandong Science, 2014, 27 (01): 68-72+91.