A research on line loss calculation based on BP neural network with genetic algorithm optimization

Yukun Jin¹, Zeng Li¹, Yipin Han¹, Xiaopeng Li¹, Pingting Li¹, Guangdi Li²³ and Hao Wang²

¹National Network Liaoning Electric Power Co., Ltd. Anshan Power supply Co., Ltd; ²College of Information Science and Engineering, Northeastern University, Shenyang, 110004, P.R. China
³Email: liguangdi@ise.neu.edu.cn

Abstract. In order to realize the calculation of the line loss of the distribution network with complex structure and low-voltage station area, this paper presents a line loss calculation method based on BP neural network with genetic algorithm optimization. The proposed method is based on the actual operation data of the distribution network. Firstly, build an error back propagation (BP) neural network model to compute the theoretical line loss of the distribution network, then use genetic algorithm (GA) to optimize the neural network and establish the GA-BP model. Based on the proposed model, the calculation demonstrates that the neural network line loss rate calculation model with genetic algorithm optimization shows better performance than the single BP neural network model, such as better nonlinear fitting ability and higher calculation accuracy. Therefore, the line loss calculation method proposed in this paper based on the BP neural network with the genetic algorithm optimization can improve the accuracy of the distribution network line loss rate calculation model.

1. Introduction

Distribution network line loss, which is the power loss on the distribution line, usually includes theoretical line loss and statistical line loss. The former refers to the losses caused by the various components in the power system, which can be calculated theoretically; the latter is the difference between the power supplied by the grid and the power sold. The internal operation and management mechanism of the distribution network can be reflected by analyzing and comparing theoretical line loss and statistical line loss.

The line loss rate is an important indicator in line loss analysis. It is the proportion of network loss in the total power supply, usually expressed as a percentage. Commonly used methods for calculating the theoretical line loss rate of distribution networks include equivalent model algorithms and power flow improvement algorithms such as PQ decomposition method, improved iteration method, and forward-backward method. With the development of the power system, the amount of data collection and network information in the distribution network is increasing. For traditional calculation methods, the expansion of the amount of information will make the calculation more difficult; while the machine learning method requires a large amount of data. For learning, the larger the amount of data, the more conducive to the establishment of the model, and the higher the calculation accuracy. Therefore, the use of machine learning to establish the calculation model of the power network line loss rate can solve the problems of complex power network structure and large data volume. Artificial
Neural Network (ANN) is one of many machine learning algorithms. ANN handles complex, parallel, and nonlinear problems by simulating the working methods of human brain neurons. Because of its strong robustness, good fault tolerance and high fitting performance, artificial neural network has been widely used in various fields.

Root mean square current method is a common theoretical line loss calculation method at present. [1]. Another calculation method is the power flow calculation method, but the calculation process of this method is very complex, and the data requirements are relatively high [2-3].

In recent years, the development and application of neural network theory [4-9] has provided a novel idea for line loss rate calculation. In [4], a calculation method of BP Neural Network improved by Particle Swarm Optimization algorithm has been proposed; In [5], after preprocessing the original data set, the BP neural network is optimized by genetic algorithm, and then the power network line loss is calculated by the optimal algorithm; In [6], the dynamic clustering algorithm is used to classify the sample data set, and then the radial basis function neural network is used to calculate the power network line loss. However, there are many shortcomings in those methods, for example, it is difficult to calculate, and the calculation error is not accurate enough. The clustering algorithm can efficiently obtain the characteristics of the original power data, therefore, it is widely used in calculating power network line loss [10-11].

In order to alleviate the aforementioned constraints and limitations, this paper uses the neural network to compute the line loss rate of the power network, and studies the effect of the BP neural network in the calculation of the line loss rate; and optimizes the BP neural network through the genetic algorithm to improve the accuracy of the calculation model of the line loss rate of the distribution network.

2. Line loss calculation model based on BP neural network

2.1. BP neural network

Figure 1 shows the process of calculating line loss using neural network. By collecting a large amount of power grid data, after normalizing the data, the data set is divided into training samples and test samples, and then the training sample set is trained to obtain the neural network, and then the independent variables in the test sample set are put into training. In the neural network, the line loss rate is obtained, and the calculated line loss rate is compared with the actual line loss rate to test the fitting performance of the network.

![Figure 1](image_url)

**Figure 1.** Flowchart of line loss calculation of distribution network based on neural network.

This article uses the most widely used BP (Back Propagation) neural network, which is a multilayer feedforward network based on the error back propagation algorithm. The learning algorithm of the BP neural network consists of forward propagation (calculation error) and back propagation (adjusting weights, threshold). The BP algorithm has a strong ability to fit non-linear functions, and can effectively deal with the non-linear relationship between characteristic parameters and line loss rate of the distribution network. The simulation model uses Matlab neural network toolbox to design the neural network structure, and build a line loss rate calculation model based on BP neural network.
2.2. Characteristic parameters
To compute the line loss rate of the distribution network through the neural network, the characteristic parameters that determine the line loss of the power network must be determined firstly. In order to ensure the power factor of the load on the user side, the system is generally equipped with sufficient reactive power compensation equipment, so we mainly calculate the active power loss $\Delta W$ in the distribution network station area. Assuming that a certain station area has $n$ branches and $m$ load points, the resistance of the $i$-th branch is $R_i$. In this line loss calculation period $T$, the active power loss has the following expression:

$$\Delta W = \sum_{i=1}^{n} \left(\frac{W^2_i + Q^2_i}{U_i^2 T}\right) R_i$$  \hspace{1cm} (1)

In (1), $n$ is the number of branches in the station area; $i$ represents the $i$-th branch; $U_i$ is the average voltage of the $i$-th branch in time $T$; $W_i$ is the active power consumed at the end of the $i$-th branch; $Q_i$ is the reactive power consumed at the end of the $i$-th branch; $R_i$ is the resistance of the $i$-th branch.

It can be concluded from the above formula that the active power loss $\Delta W$ is determined by the operating parameters $W_i$, $Q_i$, and $U_i$, and the structural parameters of the system $R_i$. $\Delta W$ is a nonlinear function composed of $W_i$, $Q_i$, $U_i$, and $R_i$, as shown in (2).

$$\Delta W = f(W_i, Q_i, U_i, R_i)$$  \hspace{1cm} (2)

For a selected station area, under normal operating conditions, its resistance $R_i$ remains basically unchanged. Under the action of the reactive power compensation device, its load side voltage varies little, so it can also be considered as maintaining Invariable, so the expression of the station area loss can be simplified to the relational expression with the active power $W_i$ and the reactive power $Q_i$.

Therefore, according to the acquired data, the monthly active power and the monthly reactive power of each station area of the distribution network are used as the characteristic parameters of the neural network model, and the line loss rate serves as the output of the model.

2.3. Data acquisition and preprocessing
This paper adopts the same monthly power data of a 110kV cooperative transformer 10kV feeder of a power supply company. We use the same monthly active and same monthly reactive power of the 8 stations on the line for 48 consecutive months as the input of the network. Collect electricity at zero daily to calculate monthly active power and monthly reactive power, and use monthly line loss rate as the output of the network. Since the annual electricity load has a similar annual load curve, in the case of less sample data, we reuse the sample data to make the trained network model more accurate. This article will reuse the sample data twice for a total of 144 sample data.

Since the active power and reactive power of the station area are relatively large, in order to remove the dimensional influence and speed up the network convergence, it is significant to normalize the original data to form a $16 \times 144$ training sample set.

The test sample set includes the daily active power, reactive power, and line loss rate of the line for 6 consecutive days. The independent variable sample data is also normalized to obtain a $16 \times 6$ test sample set. To normalize the above data, the formula is expressed as follows:

$$x' = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$$  \hspace{1cm} (3)

In (3), $x_{\text{max}}$ is the maximum value of a certain type of sample data; $x_{\text{min}}$ is the minimum value of a certain type of sample data.
The normalization of sample data can also be processed by directly calling the normalization function (*mapstd*) in the MATLAB toolbox, which directly normalizes the sample data of the independent variables to standard data, that is, the variance is 1, the mean is 0. The commonly used call format of this function is as follows:

\[
\begin{bmatrix} Z, Ps \end{bmatrix} = \text{mapstd}(X)
\]

(4)

Where \(Z\) is the standard normalized independent variable matrix, \(X\) is the sample original data matrix, and \(Ps\) is a structured data used to store data information.

2.4. BP neural network structure

The determination of the neural network structure is mainly to determine the network type, the number of layers of the BP neural network, the number of neurons in each layer, and the form of the activation function. The function from the intermediate layer (hidden layer) to the output layer selects the sigmoid function; the number of input layers is determined by the dimension of the sample data. The output layer is the line loss rate; there is currently no unified theoretical method for determining the number of intermediate layers, and it is determined by an empirical formula, which is expressed as follows:

\[
l = \sqrt{m + n + a}
\]

(5)

Where \(l\) is the number of neurons in the hidden layer; \(n\) is the number of neurons in the input layer, and \(a\) generally takes a constant between 1 and 10.

Since the input data includes the active power and reactive power of 8 stations, it is concluded that the number of neurons in the input layer of BP neural network is 16; the number of neurons in the middle layer is determined by an empirical formula, so the number of neurons in the hidden layer is 4; the output is the line loss rate of the distribution line, and the number of neurons is 1. It can be determined that the BP neural network structure is of type 16-4-1, as shown in Figure 2, where \(\omega\) is the weight between each layer, and \(b\) is the threshold value of each layer.

Figure 2. Structure diagram of power network line loss calculation model based on back propagation neural network.

3. Line loss calculation model of BP neural network based on genetic algorithm optimization

3.1. GA-BP neural network

Since the initial weight and threshold of the BP neural network are randomly generated between [0, 1], the results obtained after each retraining are different, and it is often necessary to select a better fitting network after multiple training; at the same time, back propagation neural network has some limitations: the error gradient changes very little, the adjustment time is long, the number of iterations are many, the convergence speed of neural network is slow, and the neural network output layer can easily fall into a local minimum.
Figure 3. GA-BP neural network model algorithm flowchart.

Genetic algorithm selects, crosses and mutates the original population by simulating the principle of "survival of the fittest" in nature, so as to produce a new generation of more adaptive population. According to the genetic algorithm, the weights and thresholds of the BP neural network are optimized to obtain the optimal individual. The optimal weight and threshold are used to predict and calculate the theoretical line loss of the distribution network to avoid the BP neural network from falling into the local optimum, thereby obtaining a more accurate line loss value.

Therefore, this paper presents a line loss rate calculation model based on BP neural network with genetic algorithm optimization, namely the GA-BP model. The specific algorithm flow of the model is shown in Figure 3.

3.2. Fitness function formula
Fitness indicates the pros and cons of the population in the genetic algorithm. In this paper, the reciprocal of the sum of squared errors between the calculated line loss rate and the actual value is used as the individual fitness function. The expression is expressed as follows:

$$E_i = \frac{1}{\sum_{k=0}^{s} (y(k) - s(k))^2}$$  \hspace{1cm} (6)

Where $y(k)$ is the predicted output result of the GA-BP model; $s(k)$ is the expected output result of the GA-BP model; $s(k)$ is the number of samples. The genetic algorithm is proceeding in the direction of increasing fitness.

3.3. Genetic manipulation
Genetic operations mainly refer to selection operations, crossover operations and mutation operations.
3.3.1. Selection operations. The method used in this article is the roulette method. Its advantage is that chromosomes with better fitness values have a greater chance of being selected to enter the next generation, and even chromosomes with relatively poor fitness values have the same chance. This can ensure the diversity of individuals in the group. At the same time, each generation of chromosomes can be infinitely approximated to the optimal solution, thereby avoid local extremes. The formula for choosing the roulette method is expressed as follows:

\[ f_i = \frac{k}{F_i} N; \quad p_i = \frac{f_i}{\sum_{i=1}^{N} f_i} \]  

(7)

Where \( F_i \) is the fitness value of the individual \( i \), \( k \) is the coefficient, \( N \) represents the total number of chromosomes.

3.3.2. Crossover operations. Crossover processing needs to find a pair of chromosomes in the individual, the \( m \) chromosome \( m \) and the \( l \) chromosome \( l \) cross at the \( j \) point with a certain probability to obtain a new individual. The position of the crossover operation is random. The formula is obtained as follows:

\[
\begin{align*}
  a_{mj} &= a_{mj}(1-b) + a_{lj}b \\
  a_{lj} &= a_{lj}(1-b) + a_{mj}b
\end{align*}
\]  

(8)

In (8), \( b \) is a random number between [0,1].

3.3.3. Mutation operations. The mutation operation needs to randomly find a certain chromosome \( x \), and the \( y \) gene \( a_{xy} \) mutates according to a certain mutation probability to produce a new individual. The mutation processing formula is obtained as follows:

\[
a_{xy} = \begin{cases} 
  a_{xy} + (a_{xy} - a_{\text{max}}) \times f(g), & r < 0.5 \\
  a_{xy} + (a_{\text{min}} - a_{xy}) \times f(g), & r \leq 0.5
\end{cases}
\]  

\[ f(g) = r_2 (1 - g / G_{\max})^2 \]  

(9) (10)

Where the subscripts max and min are the upper and lower bounds of the gene, \( r_2 \) is a random number, \( g \) is the number of iterations, \( G_{\max} \) is the maximum number of evolutions, and \( r \) is a random number from 0 to 1.

3.4. Realization of genetic algorithm

In the MATLAB, the GA-BP model is constructed using the Genetic Algorithm Toolbox (GAOT). The algorithm uses floating-point encoding: the input layer is 16 dimensions, the hidden layer is 4 dimensions, and the output layer is 1 dimension. A total of \( 16 \times 4 + 4 \times 1 = 68 \) weights, \( 4 + 1 = 5 \) thresholds, the length of chromosome is \( 68 + 5 = 73 \). In the genetic algorithm, the genetic algebra is 100, the population size is 50, and adopt the arithmetic crossover and roulette methods.

4. Comparison and analysis of simulation results

4.1. BP neural network simulation results

Open the Neural Net Fitting module in the status bar application in the Matlab software. Under the data selection page, we need to select the standard sample matrix file to be imported, and Targets select the target output matrix file to be imported; in the page of Validation and Test Data, select 70% for training data, 15% for validation data, and 15% for testing data; choose the number of hidden layers under the Network Architecture page is 4.
After training the neural network with the training sample set, as shown in Figure 4, the network has reached the minimum gradient requirement after 15 iterations. In order to prevent overfitting, we verify that the mean square error of the sample data has not dropped for 6 consecutive times. After reaching the default check value of the neural network toolbox, the training is stopped, and the best effect is achieved at the 9th iteration.

![Figure 4. BP neural network mean square error change curve.](image)

The test sample set (16×6) is imported into the trained neural network model, the line loss rate is calculated by the BP neural network model, and the comparison diagram and the fitting diagram of the calculated value and the actual value are drawn through Matlab, as shown in Figure 5 and Figure 6. The fitting value of the test sample is 0.94762. It can be seen that the fit of the BP model calculation needs to be further improved.

![Figure 5. Comparison chart of BP neural network line loss calculation results.](image)  ![Figure 6. BP neural network line loss calculation result fitting diagram.](image)

4.2. GA-BP neural network simulation results

After the genetic algorithm is used to optimize the BP model, the change curve of the optimal individual fitness value is shown in Figure 7. It can be seen that the optimal individual fitness value is $3.00138645377362 \times 10^{-4}$ when it evolves to the 92th generation, the inheritance is terminated when the conditions are met, and the neural network is trained after obtaining the initial values of the optimal weight and threshold.
It can be concluded from Figure 8 that the network has reached the minimum gradient requirement after 16 iterations and achieved the best results. At this time, the iterative gradient value and mean square error of the training sample are both small, which are $4.35 \times 10^{-20}$ and $1.35 \times 10^{-9}$, respectively.

As shown in Figure 9 and Figure 10, the test data set (16×6) is imported into the trained GA-BP neural network, and the comparison graph and the fitting graph of the calculated value and the actual value are drawn through Matlab. The degree of coincidence between the calculated value and the true value is very high. The fitted value $R$ of the test sample reaches 0.99928. Compared with the single BP model (Figure 5 and Figure 6), the degree of fit has increased by 5.166%.

**4.3. Comparison of simulation results**

The calculated values of BP neural network and GA-BP neural network are derived and compared with the true line loss rate, the results are seen in Table 1. The maximum absolute error of BP neural network is 0.751% and the maximum relative error is 0.1275%. However, the maximum absolute error of all calculated data of the GA-BP neural network is 0.14%, and the maximum relative error is 0.0204%. Compared with a single BP neural network, the average absolute error of the GA-BP neural network is reduced by 0.273%. The error is reduced by 0.048%. 
Table 1. Error comparison of test results of BP and GA-BP neural network.

| Sample | Actual value (%) | Calculated value (%) | Absolute error (%) | Relative error (%) |
|--------|------------------|----------------------|--------------------|-------------------|
|        | BP               | GA-BP                | BP                 | GA-BP             | BP                 | GA-BP              |
| 1      | 5.900            | 5.140                | 5.890              | 0.751             | 0.010              | 0.1275             | 0.0017             |
| 2      | 5.540            | 5.139                | 5.460              | 0.401             | 0.080              | 0.0724             | 0.0144             |
| 3      | 7.240            | 7.297                | 7.200              | 0.066             | 0.040              | 0.0091             | 0.0056             |
| 4      | 8.220            | 8.132                | 8.160              | 0.079             | 0.060              | 0.0096             | 0.0073             |
| 5      | 6.860            | 6.662                | 6.720              | 0.189             | 0.140              | 0.0276             | 0.0203             |
| 6      | 4.680            | 5.160                | 4.680              | 0.480             | 0                | 0.1028             | 0                 |
|        | Maximum error value |                     |                    | 0.751             | 0.140              | 0.1275             | 0.0204             |
|        | Mean error value |                      |                    | 0.328             | 0.055              | 0.0580             | 0.0100             |

Comparing the test results of the BP and GA-BP neural network, it can be found from the above table that the calculation accuracy of the neural network optimized by the genetic algorithm is greatly improved, which verifies that the method has better convergence, prediction accuracy and universality. Chemical ability has high theoretical and practical significance.

5. Conclusions

In this paper, the genetic algorithm and the BP neural network are combined to calculate the line loss, and the results are compared with single BP neural network model. The results show that the calculated value of the proposed BP neural network with genetic algorithm optimization is closer to the actual value, and the degree of fit is higher, and the line loss rate of the distribution network can be calculated more accurately. Compared with traditional line loss rate calculation methods, the GA-BP neural network calculation model proposed in this article has its unique advantages:

1. The neural network has strong fault tolerance and robustness, which can avoid the influence of data errors caused by human factors in the data collection process on the calculation accuracy, and can achieve accurate calculations under special circumstances such as large load fluctuations;

2. The strong generalization ability of neural network can make a model suitable for a variety of application scenarios. Because the annual power load has a similar annual load curve, for the same distribution line, the model obtained from the monthly line loss rate data of the same period can be used to calculate the daily line loss rate or the annual line loss rate of the same period.

3. The global optimization of the genetic algorithm prevents the BP neural network from falling into the local optimum. The optimal weight and initial threshold value can improve the fitting effect and calculation accuracy of the neural network model.

There are still many innovations in the direction of line loss calculation of distribution network. In the future, we can try to improve the neural network structure or study the real-time prediction of line loss.

Acknowledgements

This paper was funded by the Science and Technology Project of State Grid Liaoning Electric Power Company Ltd (2020YF-16), Research on Topology Analysis and Calculation Method of Daily Line Loss Based on Optimal Configuration.

References

[1] L Ni, L Yao, Z Wang, J Zhang, J Yuan and Y Zhou 2019 A Review of Line Loss Analysis of the Low-Voltage Distribution System 2019 IEEE 3rd International Conference on Circuits, Systems and Devices (ICCSD), Chengdu 111-114

[2] W Pan, X Chen and Y Li 2018 Calculation Method of Corona Loss of Transmission Line Based on AC/DC Power Flow 2018 IEEE International Conference on High Voltage Engineering and Application (ICHVE), ATHENS, Greece 1-4
[3] C S Chen, J C Hwang, M Y Cho and Y W Chen 1994 Development of simplified loss models for distribution system analysis in *IEEE Transactions on Power Delivery* 9(3) 1545-1551

[4] B Jianghong, L Liping, W Qi, J Mu, Y Jun and S Dunwen 2018 A method of evaluating 10kV Distribution Network Line Losses Based on Intelligent Algorithm *2018 International Conference on Power System Technology (POWERCON)*, Guangzhou 3324-3329

[5] K Y Xin, Y H Yang, F Chen 2002 An advanced algorithm based on combination of GA with BP to energy loss of distribution system *Proceedings of the CSEE* 22(2) 79-82

[6] H L Jiang, M An, X J Liu 2005 The calculation of energy losses in distribution systems based on RBF network with dynamic clustering algorithm *Proceedings of the CSEE* 25(10) 35-39

[7] F Ni and J Yu 2009 Line losses calculation in distribution network based on RBF neural network optimized by hierarchical GA 2009 *International Conference on Sustainable Power Generation and Supply*, Nanjing 1-5

[8] S Chao, L Zhensheng, H Jinlei, K Zhenxing and W Zheng 2018 Line Loss Calculation in Power Distribution Network Based on Power Measurement Data and BP Neural Network *2018 International Conference on Power System Technology (POWERCON)*, Guangzhou 4107-4112

[9] L Liu, J Bai, Y Zhang, M Jiang, Y Sun and Q Wang 2018 An Evaluation Method of 10k V Distribution Network Line Loss Based on Improved BP Neural Network *2018 China International Conference on Electricity Distribution (CICED)*, Tianjin 2401-2406

[10] Z Wang, Y Li and H Liu 2019 Analysis and Calculation of Line Loss Data Based on Hybrid Clustering *IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS)*, Beijing 310-314

[11] Q Zhou, K Yu, X Chen and S Liu 2018 Calculation Method of the Line Loss Rate in Low-voltage Transformer District Based on PCA and K-Means Clustering and Support Vector Machine *2018 International Conference on Power System Technology (POWERCON)*, Guangzhou 4264-4271