

Classification and Identification of Power Quality Disturbances Based on Stochastic Forest Model

Xin Feng^{1*}

¹College of Applied Engineering, Henan University of Science and Technology, Sanmenxia, Henan, China

*Corresponding Author.

Abstract:

With the emergence and use of a large number of new power electronic equipment, the power supply department has begun to pay extensive attention to the problem of power quality, and the majority of users have put forward higher requirements for the quality of power supply. This paper studies the classification and identification method of power quality disturbances based on random forest model. In this paper, according to IEEE power quality standard, the normal waveform and 16 common power quality disturbance waveforms are mathematically modeled, and the power quality disturbance signal is analyzed by S-transform. In this paper, the power quality disturbance identification algorithm based on random forest is optimized. Experimental data show that the optimized method has higher disturbance recognition accuracy and better anti noise ability. Therefore, using the power quality disturbance identification method proposed in this paper to monitor the power quality of power grid is of great significance to ensure the safe and stable operation of power grid and improve economic benefits.

Keywords: *Stochastic Forest Model, Power Quality Disturbance, Disturbance Identification, Power Grid Security.*

I. INTRODUCTION

With the expansion of industrial scale and the development of science and technology, the load structure in modern power system is complex [1-2]. New processes and technologies are widely used in all aspects of industrial production and people's life. More and more users adopt high-tech equipment with good performance and high efficiency, but sensitive to the change of power supply characteristics, which makes the problem of power quality more and more serious, and puts forward higher requirements for power quality [3].

The increasingly prominent problem of power quality has attracted the general attention of power supply departments and power users [4-6]. Power quality problems can be divided into steady-state quality problems and transient quality problems according to their generation and duration [7]. The harm and loss caused by the decline of voltage quality is very serious. Low quality electric energy leads to large-scale power outage of power grid, resulting in the decline of user productivity. In order to ensure the power supply quality and the safe, stable and economic operation of the system, it is very necessary and urgent to study the power quality problem scientifically and then control it effectively [8].

Classify and study all kinds of power quality phenomena, deeply analyze and explore the electromagnetic interference problems to be understood, find out the causes of power quality problems and take targeted solutions. It is of great significance to ensure high-quality power supply for social production and people's life, and minimize the impact on modern industrial enterprises and important power users.

II. BASIC THEORY OF WAVELET TRANSFORM AND ARTIFICIAL NEURAL NETWORK

1. Wavelet transform

The so-called wavelet essentially refers to a cluster of functions that are localized in both time domain and frequency domain. From the time domain, wavelet has oscillation and attenuation, and its mean value is zero. For the function $\psi(t) \in L^2(\mathbb{R})$, if and only if its Fourier transform $\psi(\omega)$ satisfies the following conditions:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (1)$$

$\psi(t)$ is a basic wavelet or Mother Wavelet. The generating function $\psi(t)$ is translated and expanded to obtain a family of wavelet sequences $\{\psi_{a,b}\}$

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a > 0, b \in \mathbb{R} \quad (2)$$

Where, the coefficient $1/a$ is the energy normalization coefficient; a is the scaling factor, also called the scaling factor, which determines the frequency domain center and bandwidth of wavelet $\psi_{a,b}(t)$; b is the translation factor, which together with a determines the time domain center of wavelet.

The continuous wavelet transform (CWT) of the square integrable function $f(t) \in L^2\mathbb{R}$ is defined as:

$$WT_f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^*\left(\frac{t-b}{a}\right) dt = f(t), \psi_{a,b} \quad a > 0 \quad (3)$$

The corresponding reconstruction formula is the inverse transform of continuous wavelet:

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} W_f(a, b) \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) db \quad (4)$$

In order to make the wavelet transform have changeable time and frequency resolution and adapt to the nonstationarity of the signal to be analyzed, we naturally need to change the size of a and B to make the wavelet transform have the function of "zoom". In practice, the most commonly used is the binary dynamic sampling grid, that is, $a_0 = 2$, $b_0 = 1$. The resulting wavelet:

$$\psi_{2^j k}(t) = 2^{-j/2} \psi\left(\frac{t-k}{2^j}\right) \quad j, k \in Z \quad (5)$$

It is called dyadic wavelet.

Dyadic wavelet has the function of zoom in signal analysis. Suppose there is a magnification of 2^j , which corresponds to a part of the observed signal. If you want to further view the smaller details of the signal, you need to increase the magnification, that is, reduce the j value; Conversely, if you want to understand the coarser content of the signal, you can reduce the magnification, that is, increase the j value.

Dyadic wavelet transform is:

$$Wf_{2^j}(b) = f(t) * \psi_{2^j, b}(t) = 2^{-j/2} \int_R f(t) \psi\left(\frac{t-b}{2^j}\right) dt \quad (6)$$

The corresponding reconstruction formula, i.e. inverse dyadic wavelet transform, is:

$$f(t) = \sum_{R \in Z} \int_R Wf_{2^j}(b) \psi_{2^j, b}^\perp(t) db \quad (7)$$

Dyadic wavelet is between continuous wavelet and discrete wavelet. It only discretizes the scale parameters, but the translation parameters in the time domain still change continuously. Therefore, dyadic wavelet transform still has the time-shift covariance of continuous wavelet transform, which is its unique advantage compared with discrete wavelet transform. Because of this, it is very useful in singularity detection, signal analysis and image processing.

2. Artificial neural network

Artificial neural networks (ANN) is an algorithmic mathematical model that imitates the behavior characteristics of animal neural networks for distributed parallel information processing. It is composed of a large number of neurons, which can be divided into different types according to different combination methods, such as back propagation neural network (BP neural network), self-organizing neural network, generalized regression neural network and probabilistic neural network.

BP (back propagation) neural network was proposed by a team of scientists led by Rumelhart and McClelland in 1986. It is a multilayer feedforward neural network trained according to the error back propagation algorithm (BP algorithm). BP network can learn and store a large number of input-output mode mapping relationships without revealing the

mathematical equations describing this mapping relationship in advance. The network adopts gradient descent method for learning, and continuously adjusts the weight and threshold of the network through back propagation to minimize the sum of squares of the network error. Because of its simple structure, many adjustable parameters and good controllability, BP neural network has been widely used in practice.

BP neural network is a neural network with three or more layers of neurons, including input layer, middle layer (hidden layer) and output layer. The upper and lower layers are fully connected, but there is no connection between neurons in the same layer, and there is no feedback connection between neurons in each layer, forming a feedforward neural network system with hierarchical structure.

III. CLASSIFICATION OF POWER QUALITY DISTURBANCES BASED ON WAVELET TRANSFORM AND PSO-BP NEURAL NETWORK

The power quality disturbance classification based on wavelet transform and artificial neural network is generally divided into the following two steps: extracting disturbance features and classifying. Using the characteristics of multi-resolution analysis of wavelet transform, the energy distribution of each disturbance signal in each frequency band can be obtained by decomposing the disturbance signal to different scales. The energy distribution does not change much when the disturbance occurs. The energy distribution of different kinds of disturbances in each frequency band is different, which can be used as the basis for distinguishing each disturbance.

Particle swarm optimization (PSO) is a biological evolution algorithm jointly proposed by American social psychologist Kennedy and electrical engineer Eberhart in 1995. PSO algorithm adopts the speed position search model, which compares the particles to a point without mass and volume but with speed and position. The particles fly at a certain speed in the search space, and the flight speed is dynamically adjusted according to the flight experience.

In order to improve the convergence performance of the algorithm, Shi and Eberhart introduced the inertia weight W into the speed term of PSO algorithm in 1998, and proposed to dynamically adjust the inertia weight in the evolution process to balance the global convergence and convergence speed. The evolution equation has been called the standard PSO algorithm by relevant scholars. The BP neural network adopts a three-layer structure, the number of neurons in the input layer is 10, the number of neurons in the hidden layer is 10, and the number of neurons in the output layer is 7. S-type tangent function (Tansig) and S-type logarithm function (logsig) are selected for the transfer functions of hidden layer and output layer respectively. The learning function is tranlm. The network structure is shown in Figure 1.

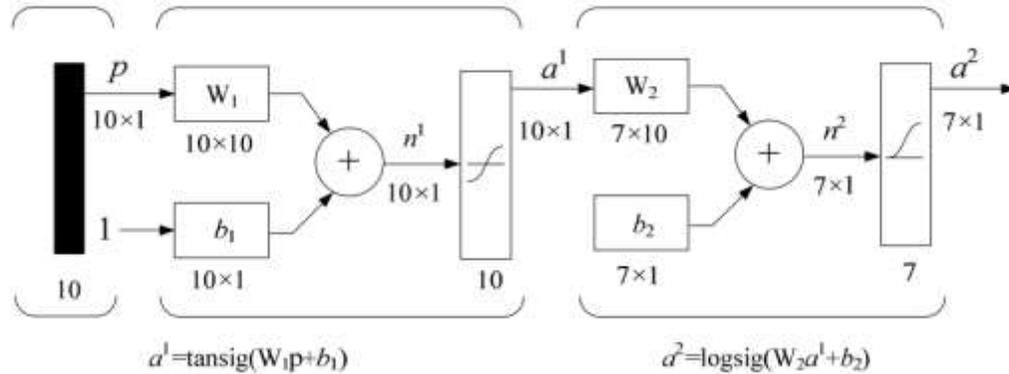


Fig 1: The network structure

The particle status update operation is as follows:

$$v_{id} = wv_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \quad (8)$$

$$x_{id} = x_{id} + v_{id} \quad (9)$$

$$w(t) = w_{max} - (w_{max} - w_{min})t / t_{max} \quad (10)$$

Where, w is the weighting coefficient, $0 < w < 1$. Generally speaking, w is taken between $[0.1, 0.9]$. w_{max} is the maximum weighting coefficient and w_{min} is the minimum weighting coefficient; r_1 and r_2 are a random number between $(0, 1)$ respectively; D is the dimension, $d = 1, 2, \dots, D$; T_{max} is the total number of iterations of the algorithm, and t is the current number of iterations; c_1 and c_2 are learning factors and non negative constants (usually $c_1 = c_2 = 2.0$). C_1 is the weight coefficient of the particle tracking its historical optimal value, which represents the particle's own understanding, and c_2 is the weight coefficient of the particle tracking the group optimal value, which represents the particle's understanding of the whole group knowledge.

The basic steps of PSO algorithm are as follows:

when PSO algorithm runs, first initialize the size of a group of particles and the initial position and initial velocity of each particle, and calculate the fitness value of each particle according to the fitness function; Then, the fitness value of each particle is compared with its own best position p_{best} . If it is better, the current position is taken as its own best position p_{best} ; The fitness value of each particle is compared with the best position g_{best} experienced by the population. If it is better, it is regarded as the best position g_{best} of the population; Finally, update the velocity and position of the particle according to formulas (8) and (9), and continue to calculate the next particle. The iteration termination condition is generally selected as the maximum number of iterations and the optimal position searched by particle swarm optimization to meet the adaptation threshold.

When optimizing BP algorithm with PSO algorithm, firstly, each connection weight and threshold of neural network are taken as the position vector of particles, that is, each network weight and threshold are taken as a particle swarm. Initialize the position vector x , and then according to the steps of particle swarm optimization algorithm, using the infinite

approximation ability of PSO algorithm, search and obtain a near optimal position vector as the initial weight and threshold of BP neural network, so as to minimize the following mean square error index (fitness value):

$$f_t = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C (y_{j,t}^d - y_{j,t})^2 \quad (11)$$

Where n is the number of samples in the training set; $y_{j,t}^d$ is the ideal output value of the j -th network output node of the i -th sample; $y_{j,t}$ is the actual output value of the j -th network output node of the i -th sample; C is the number of network output neurons.

After a set of near optimal network weights and thresholds are obtained by the above method, BP algorithm is used to further optimize according to these weights and thresholds, so as to obtain the optimal values of network weights and thresholds. The essence of BP algorithm is gradient descent method based on optimization. The learning process of the algorithm consists of two parts, namely "forward propagation of information" and "inverse propagation of error". The difference between the actual output and the expected output is the error. Through the continuous correction of the weights and thresholds of neurons in each layer of the network, the error function decreases along the gradient direction. With the alternating and repeated process of "forward propagation of information" and "inverse propagation of error", the total error can change in the direction of reduction, and finally the error can reach the expected goal ϵ until.

IV. POWER QUALITY DISTURBANCE CLASSIFICATION BASED ON WAVELET COEFFICIENT PCA + KPCA AND PNN

Principal component analysis (PCA), also known as principal component analysis, aims to transform multiple indexes into a few comprehensive indexes by using the idea of dimension reduction. In statistics, principal component analysis is a technique to simplify data sets. It is a linear transformation. This transformation transforms the data into a new coordinate system, so that the first variance of any data projection is on the first coordinate (called the first principal component), the second variance is on the second coordinate (the second principal component), and so on. PCA is often used to reduce the dimension of data sets while maintaining the characteristics of the data sets with the largest contribution to each other's difference. Therefore, PCA is often used for feature extraction in classification and recognition to reduce the dimension of the original sample data.

The basic idea of KPCA is through a nonlinear transformation Φ . The input data x ($x \in \mathbb{R}^n$) is mapped to the feature space F , $F = \{ \Phi(x) : x \in \mathbb{R}^n \}$, that is, the data that cannot be linearly classified in the input space can be transformed into the feature space, and then PCA can be

used for feature extraction to realize linear classification.

For the general PCA method, that is, by solving the characteristic equation:

$$\lambda v = Cv \quad (12)$$

There are three kinds of kernel functions widely used by KPCA:

(1) Polynomial kernel function:

$$K(x, x_i) = [(x \cdot x_i) + 1]^q \quad (13)$$

(2) Multilayer perceptron kernel function:

$$K(x, x_i) = \tanh[V(x \cdot x_i) + c] \quad (14)$$

(3) Gaussian radial basis function (RBF) kernel function:

$$K(x, x_i) = \exp\left(-\|x \cdot x_i\|^2 / \sigma^2\right) \quad (15)$$

In this paper, the third Gaussian radial basis function kernel function will be used.

The dimension reduced wavelet coefficients are used as the input of PNN to realize the correct classification of various power quality disturbances. PNN is a simple neural network which can be used for pattern classification. Its essence is a parallel algorithm based on Bayesian minimum risk criterion. Compared with BP network, it has the advantages of fast training, new samples do not need to be retrained, and the optimal solution under Bayesian classification criterion can be obtained as long as there are enough training data. Therefore, it is widely used in various pattern classification problems.

When training the network, the network directly stores the training sample vector as the pattern sample vector of the network without any modification. It only needs to estimate the smoothing factor of Gaussian function by empirical statistics. The process is very simple: when the network works, the sample x to be identified is directly sent from the input layer to each category unit of the pattern layer, and the dot product of vector x and W is carried out in the pattern unit. After the nonlinear processing is completed, it is sent to the summation layer. In the summation layer, each unit is only connected with the corresponding category of mode units, and the probability of each category is estimated according to the Parzen method; In the decision layer (output layer), according to the probability estimation of the input vector, the input vector is divided into the category with the maximum a posteriori probability according to the Bayesian classification rule. The first mock exam is to identify the output of the same mode in the pattern layer and multiplied by the cost factor. The decision-making layer selects the disturbance pattern corresponding to the largest output in the summation layer as the classification result.

PSCAD / EMTDC is used to simulate the normal voltage and 6 kinds of power quality disturbance waveforms as sample data. A total of 140 training samples (20 in each class) and 70 test samples (10 in each class) are generated. The amplitude of three-phase power supply is 110KV, the initial phase angle is 90 degrees, the frequency is 50Hz, the sampling frequency is

6.4kHz, the waveform data of 8 cycles are taken, a total of 1024 points are sampled, and the time length is 0.16s. The voltage fluctuation results are shown in Figure 2.

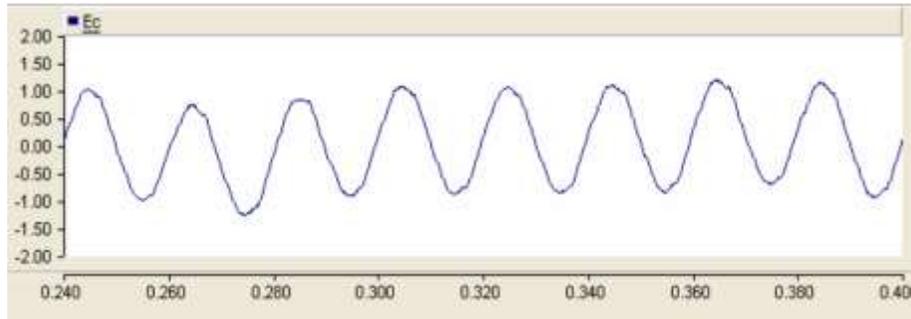
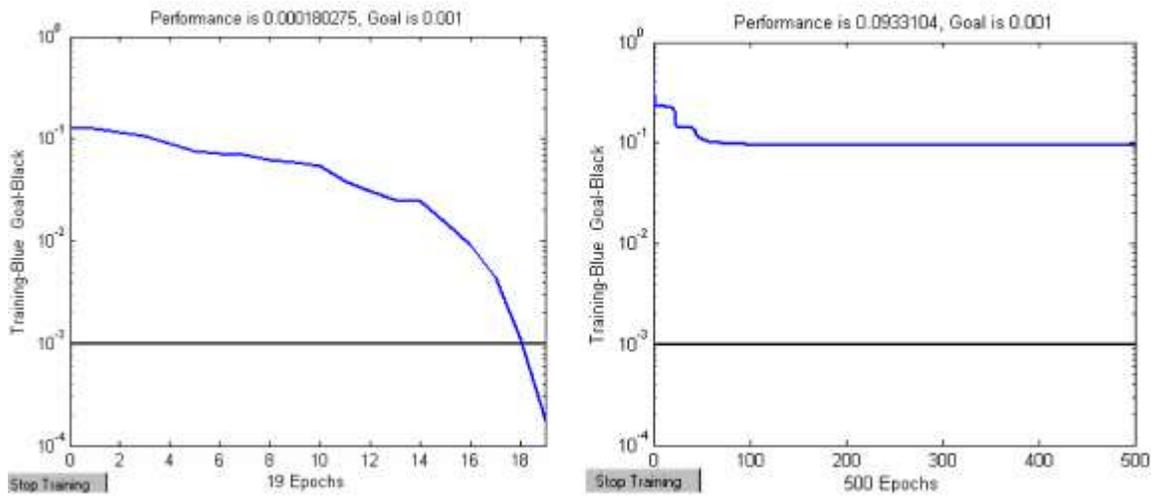


Fig 2: The voltage fluctuation results

The training error curve and times of PSO-BP network are shown in Fig. 3 (a). The training error curve and times of BP network without PSO algorithm are shown in Fig. 3 (b). It can be seen from Figure 3 that after PSO algorithm searches for the near optimal initial weight and threshold of BP network, the training convergence of BP network is fast and can be easily met by the training requirements for the target. The BP network without PSO algorithm needs a long time to meet the convergence requirements because of the strong randomness of initial weight and threshold. Sometimes it doesn't even converge and can't meet the training objectives.



(a) PSO-BP network training curve

(b) BP network training curve

Fig 3: The training curve of two networks

V. CONCLUSION

Firstly, the power quality problem is introduced, and the existing power quality disturbance classification methods are deeply analyzed and studied. On this basis, the combination of

wavelet transform and artificial neural network is used to classify power quality disturbances. This method is mainly divided into two aspects: disturbance feature extraction and neural network classification. Aiming at the shortcomings of the previous methods of directly extracting wavelet coefficients for classification (such as large input and slow learning speed), a two-stage feature extraction method based on wavelet coefficients PCA + KPCA is proposed; Aiming at the shortcomings of BP network widely used in neural network classification (slow convergence speed, easy to fall into local minimum, etc.), a classification method using PSO-BP neural network is proposed.

ACKNOWLEDGEMENTS

This research was supported by Training Plan for Young Backbone Teachers in Henan Higher Vocational Schools of China (No. 2020GZGG076).

REFERENCES

1. Liu Xiangyang, Wang Dan. Design and Research of Electrical Automation Control System Based on Intelligent Technology. *Electronic Design Engineering*, 2019, V.27; No.414 (16): 72-75
2. Wen Wenhao. Research on Electrical Automation Control System Based on Artificial Intelligence. *Information and Computer (theoretical Edition)*, 2020, V.32; No.446 (04): 152-153
3. Liu Xiaoyan. Application of Artificial Intelligence Technology in Electrical Automation Control System. *Rural Economy and Technology*, 2019, 030 (022): 254-255
4. Guo Yanan, Wang Yanfei. Application of Artificial Intelligence Technology in Electrical Automation Control. *Architectural Engineering Technology and Design*, 2018, 16: 4184
5. Huang Qiang. Application of Artificial Intelligence Technology in Electrical Automation Control. *Consumer Guide*, 2019, 10: 120
6. Yin Zhangsheng. Research on Electrical Automatic Control Technology Under Artificial Intelligence Technology. *Development Orientation of Building Materials*, 2019, 017 (009): 363
7. Wang Li. Application of Artificial Intelligence in Electrical Control System. *Full Text Edition: Engineering Technology*, 2016 (6): 175-175
8. Jiang Fkuan, Zhou Rui. Discussion on the Application of Artificial Intelligence Technology in Electrical Automation Control. *Dual Use Technologies and Products*, 2018, 6: 96