Simulated annealing-wavelet neural network for vibration fault diagnosis of hydro-turbine generating unit

XIAO ZHIHUAI[a,*], SUN ZHAOHUI[a], SONG LIBO[a], ZHANG XIAOJING[a], O. P. MALIK[b]
[a]*Wuhan University, School of Power and Mechanical Engineering, Wuhan 430072,China*
[b]*University of Calgary, Department of Electrical and Computer Engineering, University of Calgary, Canada*

In view of non-linear characteristics between fault symptoms and fault types of hydro-turbine generating unit and defects of traditional wavelet neural network learning method, a wavelet neural network fault diagnosis model based on simulated annealing algorithm is designed and applied to the hydro-turbine fault diagnosis. Instead of gradient descent method, the simulated annealing algorithm is applied to optimize parameters of wavelet neural network. Example results show that the designed model has higher convergence precision and faster convergence speed compared with wavelet neural network and additional momentum BP neural network. The simulated annealing algorithm wavelet neural network can be effectively applied to hydro-turbine fault diagnosis, and it provides a new way for hydro-turbine fault diagnosis.

(Received April 15, 2015; accepted May 7, 2015)

Keywords: Simulated annealing algorithm; Wavelet neural network; Gradient descent; Hydro-turbine generating unit; Fault diagnosis

1. Introduction

Hydropower stations undertake the responsibility of frequency regulation, peak load and emergency reserve in electricity grid, and they play a significant role in improving the power supply quality, enhancing the economic performance and ensuring the safe and stable operation of electric power system. As the core of hydropower stations, hydropower units are being developed in the direction of large-scale, complicated and high-power. As a result, the structure of units becomes more complex and the degree of integration is improved, thus causing more potential safety problems. Efficient fault diagnosis of hydropower units has become an important subject in the field of hydropower units fault diagnosis. According to the statistics, over 80% of units' fault can be demonstrated in its vibration signals [1]. However, it's quite difficult to establish the one-on-one matching between the vibrating features. And the causes of faults for the vibration signals of hydro units are nonlinear signals that may affected by hydraulic, mechanical or electrical factors, hence makes it difficult to conduct fault diagnosis of hydropower units. In consideration of the complexity, coupling of hydro turbine vibration faults, nonlinear diagnostic models are commonly used to realize the effective mapping between feature set and the fault set. The development of computer and artificial intelligence gives the fault diagnosis more choices, like the artificial neural network [2], Bayesian network [3], support vector machine [4], etc. BP neural network is a widely used algorithm, however, the conventional BP network method has difficulties getting the optimal solution in the training process, for its disadvantages, such as slow convergence speed and easy to be caught in the local minimum point of the target function [5]; Bayesian network has such disadvantages that it's difficult to construct Bayesian network model and nodes' conditional probability tables and to obtain the root nodes' fault rates and fault probabilities accurately in Bayesian network method[6]; Support vector machine models are poor in the performance of classifier and generalization ability for the uncertainty of hyper-parameter selection[7], and they have the disadvantages of low diagnostic accuracy and slow computing speed.

Another rapidly developing signal processing method in recent years is the wavelet analysis technology. The free expansion and translation of wavelet basis function and its great time-frequency analyzing ability are the key factors that makes wavelet analysis a powerful tool in processing the non-stationary random signal. Combining with the artificial neural network, the wavelet analysis is able to provide a new method for the fault diagnosis field by constructing the wavelet neural network. The wavelet neural network is a new type of feed-forward network based on wavelet analysis that can adjust wavelet basis adaptively to realize the wavelet transform. Not only does it has the time-frequency localization property like the wavelet transform does, it also has the self-learning function of as conventional neural network provides, thus making the wavelet neural network a method with strong approximation capability, fault-tolerant ability and pattern classification ability. And wavelet neural network is widely used in many areas [8,9]. But there hasn't been a specialized training method for wavelet neural network, and its general training is conducted using gradient descent algorithm [10], which affects the precision and speed of network training. Worldwide experts have proposed a lot of intelligent algorithms in optimizing wavelet network. Literature [10] and [11] optimized the wavelet neural network using particle swarm optimization and genetic algorithm respectively. These algorithms were applied to the fault diagnosis of gear case and power transformer respectively, improving the fault diagnosis accuracy and speed. However, convergence doesn't perfectly happen in both particle swarm optimization and genetic algorithm [12]. For example, when there are several local extreme values, the function could be confused by the local minimum rather than finding the global optimal point. Therefore, it's essential to develop an algorithm with strong search capacity in both local and global region.

Simulated annealing (SA) algorithm is a stochastic optimization algorithm that can search for the optimal or near-optimal solution of the function throughout the global scope. The network can adjust in the direction of minimizing the target function and accepting the increase of the target function in a certain probability distribution according to the random changes of to-be-optimized variables, so as to prevent the network from falling into the local minimum and approach the global optimal solution [13]. Kirkpatrick [14] has successfully introduced it to the usage of combinatorial optimization, and it was widely used in fields of image processing, communication and economy [15-17]. We constructed the simulated annealing-wavelet neural network (SA-WNN) in this paper by combining simulated annealing and wavelet neural network together, and applied it to the fault diagnosis of hydropower units. Considering the spectrum signatures of hydropower units' vibration signals and the fault types of hydropower units as the inputs and outputs of SA-WNN respectively, we constructed the SA-WNN model of hydropower units vibrating fault diagnosis. By using simulated annealing algorithm instead of gradient descent to optimize the parameters of wavelet neural network, we can realize the vibration fault diagnosis of hydropower units. The diagnostic results show that the proposed SA-WNN algorithm reduces the number of iterations, improves the convergence precision in comparison with the conventional BP neural network and wavelet neural network, thus providing a new method for on-line vibrating fault diagnosis of hydropower units.

2. Principles of Simulated Annealing (SA) Algorithm

Simulated annealing (SA) algorithm derives from the solid annealing principle: First, heat the solid to a certain high temperature, at the same time particles inside the solid become disordered and the internal energy increases; Then, cool solid slowly and particles become orderly. During the period of cooling process, equilibrium state is achieved at each temperature. Finally, internal energy is reduced to minimum when the material is cooled to a specified low temperature. The energy function of SA algorithm is nonlinear, which contains a number of variables. The initial variables are randomly assigned. The variables are adjusted to minimize the energy function by searching back and forth. Due to the randomness of search process, SA algorithm will accept these points that increase energy function in accordance with the Boltzmann distribution to prevent falling into local minimum "trap".

SA algorithm can be decomposed into three parts: solution space, objective function and initial solution. Four steps of generation and acceptance of new solutions are given as follows:

step 1: Create a new solution located in the solution space by initial solution through generating function ; Usually, the way of generating new solution is to add random perturbation on basis of initial solution in order to shorten the time of subsequent calculation.

step 2: Calculate the increment of objective functions between initial and new solutions.

step 3：Determine the new solution is accepted or not.

step 4: If the new solution is accepted, let it be the initial solution and begin next iteration. Otherwise, begin next iteration on the basis of initial solution.

Basic processes of SA algorithm are shown as follows according to above principles:

(1) Initialize parameters. Set initial temperature T_0, initial solution S (S is the starting point of iteration), the value of iterations R to reach an equilibrium state at each temperature;

(2) Do step (3) to (6), for m=1,..., R;

(3)Generate new solution S';

(4) Calculate the increment of objective functions: $\Delta E= f(S')-f(S)$, f is defined as objective function;

(5) Judge whether the new solution can be accepted based on the Metropolis criterion; If $\Delta E<0$, $S=S'$, and S' is accepted. Otherwise, compute the probability of accepting S', $p=\exp(-\Delta E/kT)$, k is Boltzmann constant, T is current temperature;

(6) If the termination condition is satisfied, output S' as the optimal solution and end the program; Otherwise, go to step (3);

(7) Reduce T gradually, and then turn to step (2);

3. Wavelet Neural Network (WNN)

Wavelet neural network(WNN) is a kind of feed-forward neural network based on the theory of wavelet transform and combines the localization properties of wavelet transform and neural network of large-scale data parallel processing, self-study ability, thus has strong approximation ability and faster convergence speed. WNN can be divided as loose type and compact type. This paper adopts the compact type of WNN.

The structure of WNN is shown in Fig. 1, where M, n and N are number of input layer nodes, hidden layer nodes and output layer nodes, respectively. x (x_k , $k=1,2,\ldots \mathrm{M}$), O (o_j , $j=1,2,\ldots n$) and y (y_i , $i=1,2,\ldots \mathrm{N}$) are vectors of input layer, hidden layer and output layer, respectively. W_{jk} ($j=1,2,\ldots n$; $k=1,2,\ldots \mathrm{M}$)is weight parameters vector between input layer and hidden layer. V_{ij} ($i=1,2,\ldots \mathrm{N}$; $j=1,2,\ldots n$)is weight parameters vector between hidden layer and output layer. b (b_j , $j=1,2,\ldots n$)and a (a_j , $j=1,2,\ldots n$)are translation parameters vector and scaling parameters vector, respectively. h is the wavelet function.

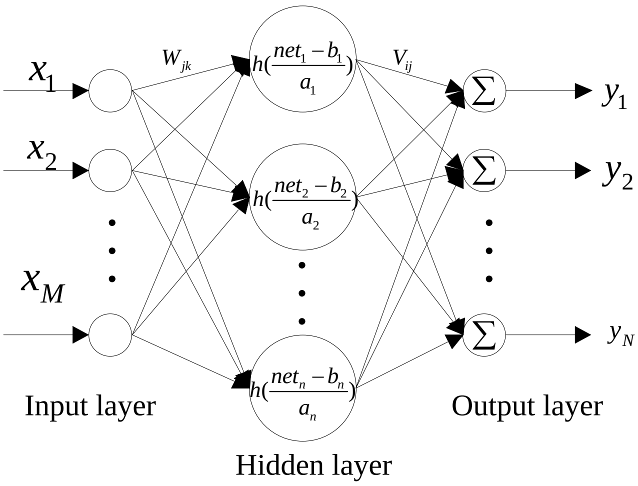

Fig.1 Structure of a WNN

Suppose that $net_j=\sum_{k=1}^{M}W_{jk}x_k$, the model of WNN can be expressed as equation (1):

$$y_i=\sum_{j=1}^{n}V_{ij}h(\frac{net_j-b_j}{a_j}) \tag{1}$$

The training of WNN is similar to that of BP neural network: error function of WNN should be defined firstly. In this paper, the error function of the WNN is shown as equation (2):

$$E=\frac{1}{2}\sum_{l=1}^{L}\sum_{i=1}^{N}(d_i^l-y_i^l)^2 \tag{2}$$

Here, L is the number of training samples, d_i ($i=1,2,\ldots \mathrm{N}$) is the object output vector.

4. Structure of simulated annealing-wavelet neural network (SA-WNN)

In this paper, a fault diagnosis model is proposed by combining simulated annealing algorithm and wavelet neural network (SA-WNN), the model structure is shown in Fig. 2. W_{jk}, V_{ij}, b, a are to-be-optimized solutions. The objective function of SA-WNN is the error function of WNN. As the replacement of gradient descent method, the SA algorithm is adopted to optimize the parameters of the fault diagnosis model.

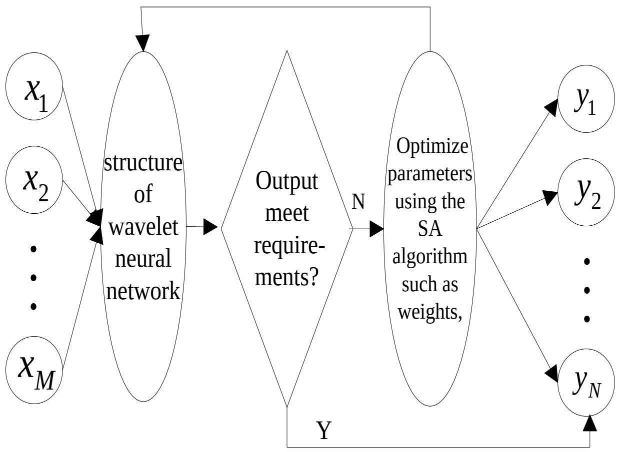

Fig.2 Structure of SA-WNN

4.1 Principles of SA algorithm for optimizing parameters

SA algorithm includes three functions (generating function, acceptance function and cooling function) and two principles (inner loop termination criteria and external loop termination criteria), which have great impacts on optimization performance of SA algorithm. In this paper, they are designed as follows:

(1) Generating function. The new solution is generated by applying random distortion which meets Cauchy distribution on initial solution. A control parameter is

given before the distortion so that local search could be realized by producing slight distortion with a larger probability and can walk out of the local minimum area by producing big disturbance appropriately.

(2) Acceptance function. The acceptance function is shown as equation (3):

$$p=\begin{cases}1 & ,\Delta E<0\\ e^{-\frac{\Delta E}{kT_p}} & ,\Delta E\geq 0\end{cases} \qquad (3)$$

If $\Delta E<0$ or $\Delta E\geq 0$ and $e^{-\frac{\Delta E}{kT_p}}>$ random [0,1],the new solution will be accepted, ΔE denotes the increment of objective functions, T_p denotes current temperature, k is Boltzmann constant.

(3) Cooling function. The temperature is updated using the formula (4) [18]:

$$T_p = T0 \big/ \left(1+\ln(p)\right) \qquad (4)$$

T0 is the initial temperature, p represents cool times.

(4) Inner loop termination criterion. It's used to determine searching times at each temperature. In this paper, we regarded that an equilibrium state could be reached at each temperature and the inner loop ends when inner loop iterate specified number *R*, where *R*=20.

(5) External loop termination criteria. This paper adopts two methods:

1) External loop iterate specified number *Z*, where *Z*=1000;

2) *E*, the value of objective function, reduces to λ, where λ =0.01; If external loop iterates more than 1000 times and *E*> λ, the optimization fails.

4.2 Steps of SA-WNN

Process of SA-WNN is shown in Fig. 3:

(1)Initialize solution and parameters, such as W_{jk}, V_{ij}, a_j, b_j, *T0,Z,R*. For *m*=1,⋯,*R*, do step(2)to(6).

(2)Calculate *E*, the value of objective function. If *E*< λ, output current solution and end the algorithm; Otherwise, turn to next step.

(3)Generate new solution. $W_{jk}*=W_{jk}+\alpha\Delta W_{jk}$, $V_{ij}*=V_{ij}+\alpha\Delta V_{ij}$, $a_j*=a_j+\alpha\Delta a_j$, $b_j*=b_j+\alpha\Delta b_j$, α (0< α <1)is a control parameter, calculate *E'*.

(4)Calculate the increment of objective functions: *ΔE=E-E'*.

(5)If $\Delta E<0$ or $\Delta E\geq 0$ and $e^{-\frac{\Delta E}{kT_p}}$ >random[0,1], accept new solution, $W_{jk}=W_{jk}*$, $V_{ij}=V_{ij}*$, $a_j=a_j*$, $b_j=b_j*$ Otherwise, go to step (3).

(6) Repeat step (2) to (5) if the equilibrium state at current temperature hasn't been reached.

(7)Cool down *T*, *p = p + 1*, and repeat step (2) to (6). When external loop termination criteria is satisfied, end the algorithm.

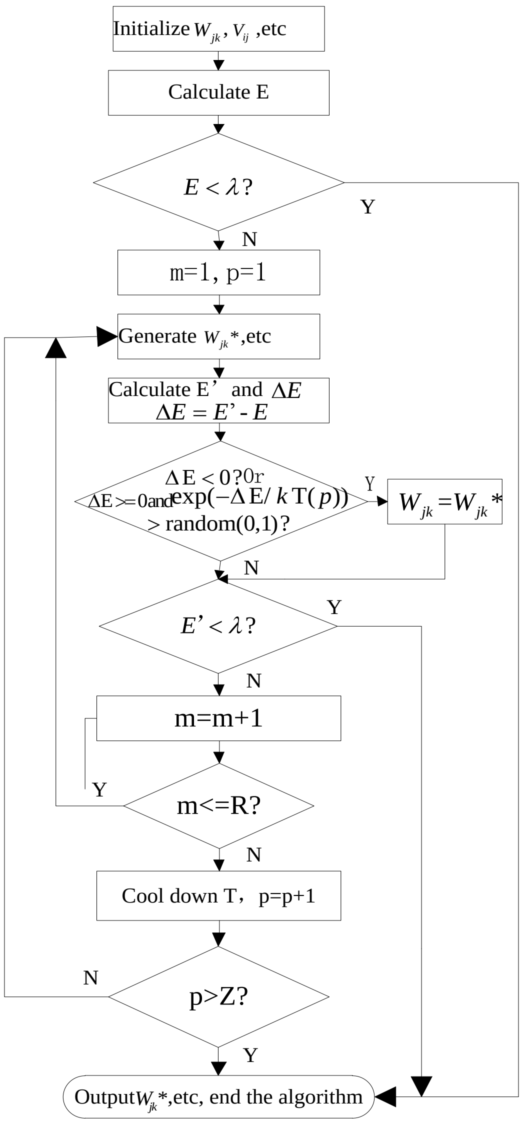

Fig.3 Steps of SA-WNN

5. Hydro-turbine generating fault diagnosis Example

5.1 Determination of input layer, output layer and hidden layer

Vibration fault of a hydro-turbine generating unit are the synthesized reflection of hydraulic, mechanical and electric factors. The method, which takes amplitude of vibration signals frequency components as feature vector and employs neural network to realize the mapping from vibration feature set to fault set, is a common method for hydro-turbine generating unit fault diagnosis. In this paper, the amplitudes of vibration signals frequency components $0.4 \sim 0.5f$, $1f$, $2f$, $3f$ and $>3f$ are chosen to form the feature vector of input layer, and 3 fault conditions (vortex with eccentric, unbalance and misalignment) and normal condition of hydro-turbine generating unit are taken as fault types waiting to be recognized. Here, the letter "*f*" represents the fundamental frequency of hydro-turbine generating unit. The model of SA-WNN is employed to diagnose vibration fault of the hydro-turbine generating unit. To make the structure simpler and the result of the neural network more intuitive, the multi-input and single-output structure is employed. According to the characteristics of the training samples, the number of input nodes is chosen as 5 which is equal to the number of feature parameters, the number of output nodes is 1. The number of hidden layer nodes is chosen as 8 according to our experience. The objective values 1, 2, 3, and 4 are defined as the values corresponding to vortex with eccentric, unbalance, misalignment and normal machinery conditions, respectively. The normalized training and testing feature samples are shown in Table 1 and 2, respectively.

Table 1 Training samples

Fault types	(0.4-0.5)*f*	1*f*	2*f*	3*f*	>3*f*	Object values
Vortex with eccentric	0.88	0.22	0.02	0.04	0.06	1
Vortex with eccentric	0.85	0.25	0.06	0.02	0.01	1
Unbalance	0.04	0.98	0.10	0.02	0.02	2
Unbalance	0.03	0.96	0.12	0.04	0.03	2
Misalignment	0.02	0.41	0.43	0.34	0.15	3
Misalignment	0.02	0.45	0.42	0.28	0.29	3
Normal	0.01	0.02	0.01	0.05	0.04	4
Normal	0. 10	0.03	0.02	0.03	0.04	4

Table 2 Testing samples

Fault types	(0.4-0.5)*f*	1*f*	2*f*	3*f*	>3*f*	Object values
Vortex with eccentric	0.82	0.28	0.05	0.04	0.03	1
Unbalance	0.02	0.91	0.08	0.01	0.02	2
Misalignment	0.01	0.48	0.48	0.36	0.20	3
Normal	0.10	0.03	0.02	0.03	0.04	4

5.2Network training and analysis

The wavelet basis function is set as the commonly used Morlet wavelet function $h(t)=\cos(1.75t)\exp(-\frac{1}{2}t^2)$, and the objective error E is set as 0.01. The value of error function is reduced to expectation error when the SA-WNN iterates 84steps, which consumes 1.771 s. Training curve is shown in Fig. 4.

In order to prove the effectiveness of SA-WNN diagnosis model, WNN and additional momentum BP network are chosen as comparison methods. The training method of WNN takes gradient descent algorithm, and the learning factor of WNN weights is set as 0.04. The structure of additional momentum BP neural network is 5-8-1; It can be seen from these figures that the WNN needs 206 training times to reach the objective error and consumes 3.693s. The additional momentum BP network needs 542 training times to reach the objective error and consumes 10s. Training curves are shown in Fig. 5 and 6, respectively.

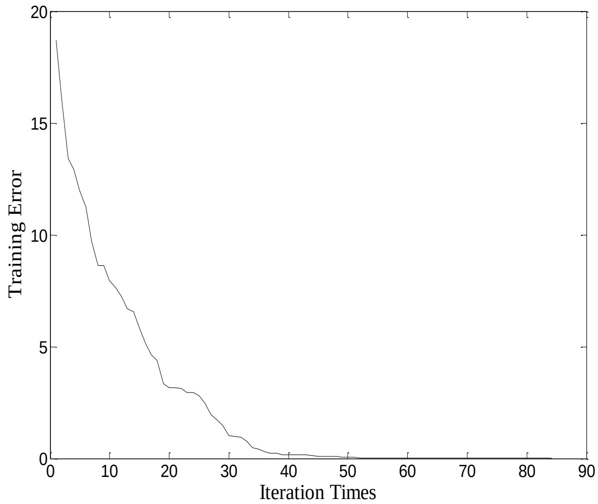

Fig.4 Training curve of SA-WNN

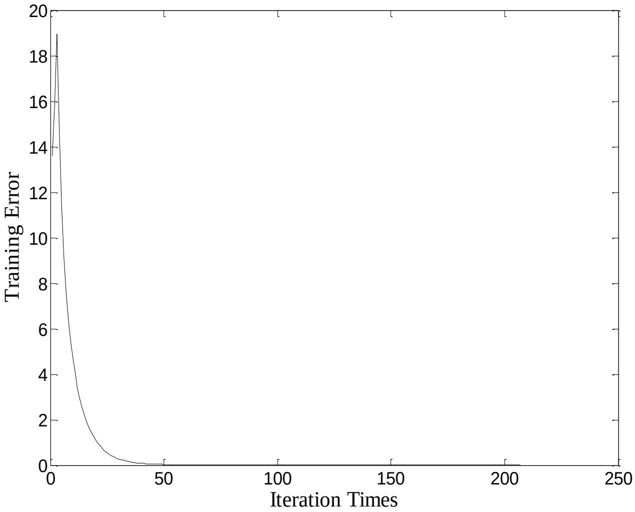

Fig.5 Training curve of WNN

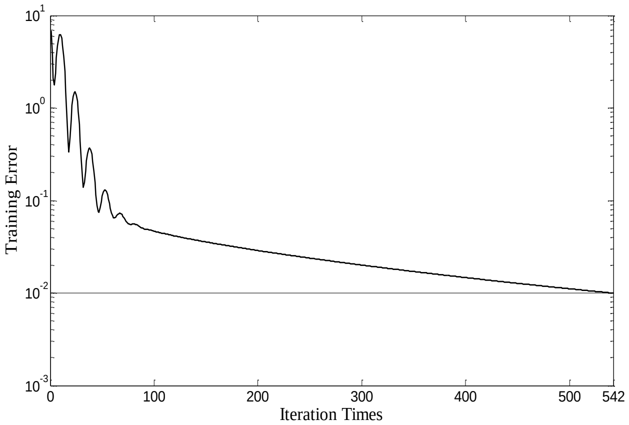

Fig.6 Training curve of additional momentum BP

The testing samples in Table 2 are used to test the obtained SA-WNN, WNN and the additional momentum BP network. Table 3 shows the test results. Table 4 demonstrates the diagnosis error and training time of the three kinds of networks.

Table 3 Diagnosis Results of Three Methods

Fault types	SA-WNN	WNN	BP	Object values
Vortex with eccentric	1.0960	0.9896	1.0996	1
Unbalance	2.0421	2.1618	2.1069	2
Misalignment	2.9945	2.9291	2.8070	3
Normal	4.0011	3.9944	3.8236	4

Table 4 Error and Training Time of Three Methods

Diagnosing methods	SA-WNN	WNN	BP
Diagnosing error	0.0055	0.0156	0.0445
Training time/s	1.771	3.693	10

As can be seen from table 3, the diagnosis results of three methods are all close to the target value aiming at the same fault, all three kinds of neural networks can recognize fault types of the hydro-turbine generating units. From Figures 4-5 and Tables 3-4, convergence time and diagnostic accuracy are significantly different in fault diagnosis of hydropower units among three methods. The additional momentum BP network is easy to fall into local minimum area, which makes it consume longest time, so it is difficult to meet the requirements of real-time diagnosis of hydropower units. Because of using gradient descent method to train the network, WNN iterates 206 steps and consumes 3.693s. After only 84 iterations, the value of error function is reduced to 0.01 with minimal training time and highest diagnostic accuracy by using SA-WNN. The results shows that the using SA algorithm to optimize the parameters of WNN can significantly improve the efficiency and convergence capacity of the network, and SA-WNN has stronger generalization ability than WNN and additional momentum BP network.

6. Conclusions

Considering the drawbacks of gradient descent algorithm of WNN, such as low convergence speed, falling into local minima easily, a SA-WNN diagnostic model is proposed in this paper for vibration fault diagnosis of a hydro-turbine generating unit. A real vibration fault diagnosis case result of a hydro-turbine generating unit shows that the proposed model has faster convergence speed and higher diagnostic accuracy than WNN and additional momentum BP neural network and turns out to be a good method for fault diagnosis of hydro-turbine generating units.

Acknowledgement

The authors acknowledge the financial support from the National Natural Science Foundation of China under Grant No.51379160

References

[1] Zhao Daoli, Ma Wei, Liang Wuke, Luo Xingqi. Proceedings of the CSEE, **25**(20), 137 (2005) (in Chinese).
[2] Fu Xiangqian, Liu Guang-Lin, Jiang Jin. Engineering Journal of WuHan University, , **35**(1), 24 (2002) (in Chinese).
[3] Hua Bin, Zhou Jianzhong, Yu Jing. Journal of North China Electric Power University **05**, 33 (2004) (in Chinese).
[4] Zou Min, Zhou Jianzhong, Liu Zhong. China Rural Water and Hydropower, **01**, 114 (2008) (in Chinese).
[5] Peng Wenji, Luo Xingqi. Journal of Hydroelectric Engineering, **26**(1), 123 (2007) (in Chinese).
[6] Yao Chengyu, Chen Dongning, Wang Bin. Journal of Mechanical Engineering, **02,** 193 (2014). (in Chinese).
[7] Cao Chongfeng, Yang Shixi, Zhou Xiaofeng. Journal of Vibration, Measurement & Diagnosis, **29**(3), 270 (2009) (in Chinese).
[8] Zhimin Du, Xinqiao Jin, Yunyu Yang, Applied Energy, **86**(9), 1624 (2009).
[9] Nengmou Wang, Hojjat Adeli,. Engineering Applications of Artificial Intelligence, **41,** 249 (2015).
[10] Mao Hongwei, Pan Hongxia, Liu Wenli. Journal of Vibration and Shock, **26**(5), 133 (2007) (in Chinese).
[11] Pan Chong, Chen Weigen, Yun Yuxin, Du Lin, Sun Caixin. Automation of Electric Power Systems, **31**(13), 88 (2007) (in Chinese).
[12] Lei Haoxia, Liu Nian, Cui Dongjun,Ma Tiejan, Xu Haiba. Power System Protection and Control, **39**(22), 52 (2011) (in Chinese).
[13] Zeng Liping,Wang Haibo, He Yuehua. Computer Simulation, **07**, 316 (2011) (in Chinese).
[14] S Kirkpatrick, C D Gelatt, M P. Vecchi Science, **220**(4598), 671 (1983).
[15] F.-C. Jeng, J.W. Woods IEEE Transactions on Information Theory **36**(1):94 (1990).
[16] Chung-Ju Chang, Chen-Hsiang Wu. IEEE Transactions on Vehicular Technology, **42**(2), 205 (1993)
[17] K.P. Wong, C.C. Fung. IEEE proceedings. Generation, Transmission, and Distribution, **140**(6), 509 (1993)
[18] Li Laiqiang, Wang Shulin, Li Shengjuan, Xu Bo. Journal of Vibration and Shock, **29**(3), 89, (2010) (in Chinese).

***Corresponding author: xiaozhihuai@126.com**