ANN with the Error Contracting Gradually Algorithm and Its Application in Generator Fault Diagnosis

Shuting Wan[1,2] Yonggang Li[1,2]

[1] College of Mechanical Engineering, North China Electric Power University, Baoding 071003, P. R. China
[2] Key Laboratory of Condition Monitoring and Control for Power Plant Equipment of Ministry of Education , Baoding 071003, P. R. China

Abstract

Based on analysis of conventional back-propagation (BP) network, the causes of error curve oscillation and excessive learning are first proposed. Next, a new BP with the error contracting gradually algorithm is put forward, through setting up neuron error threshold function, only when neuron's error is bigger than the error threshold, the neuron's parameters can be adjusted, otherwise the neuron which error is smaller than the error threshold can't be adjusted. The proposed algorithm can avoid the excessive learning and learning error oscillation. Finally, two fault diagnosis models based on the new BP algorithm is set up respectively, which are turbine-generator set vibration fault diagnosis model and rotor winding inter-turn short circuit fault diagnosis model. The results of verification show that the model has faster speed and higher diagnosis precision.

Keywords: Artificial neural network, The error contracting gradually algorithm, Turbine-generator set vibration fault, Rotor winding inter-turn short circuit fault, Fault diagnosis

1. Introduction

Turbine-generator set belongs to large-scale complicated rotating machinery, the status monitoring and diagnosis are more importance for safe and steady operation. Artificial neural network (ANN) has the ability to perform nonlinear mappings and recognize patterns, it is applied widely to intelligent diagnosis of turbine-generator set. Back-propagation (BP) network is one of the ANN applied widely. The conventional BP algorithm has larger applied potential, but it has the following shortcomings: (1) lower learning and converging speed; (2) when learning rate and momentum factors are selected improperly, error curve will oscillate, and affects learning speed, precision and stability. For improving BP network capability, the conventional back-propagation network was optimized by the method of layer by layer[1]. to overcome the defection of divergent phenomenon, when the selected original parameter and network structure is not suitable, POWELL optimization algorithm was applied to the turbine-generator set fault diagnosis and realized the neural network weight and threshold rapid calculation.[2]

This article improves the conventional BP network, and presents a new fast BP algorithm. Through setting up neuron error threshold function, only when neuron's error is bigger than the error threshold, the neuron's parameters will be adjusted, the new algorithm can avoid the excessive learning phenomenon of neural network and learning error oscillation, and increase BP network learning speed. The new fast BP algorithm is applied successfully to fault diagnosis and identification of the turbine-generator set vibration fault and the rotor winding inter-turn short circuit fault. The results of verification show that the method can assure steady learning process, avoid error curve oscillating, and has faster learning speed and higher diagnosis precision.

2. Topology structure and algorithm of BP neural network

Fig.1 is BP neural network structure. N, L, M is the number of input layer, hidden layer and output layer neuron respectively. [3]
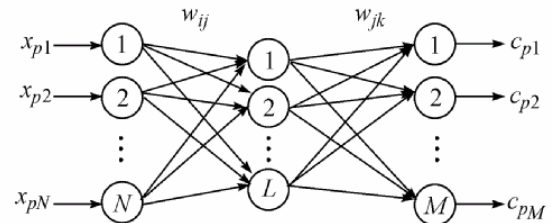

Fig.1: Structure of BP neural network.

The neuron function is Sigmoid function:

$$f(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

The parameter calculation formulae are follow:

$$\begin{cases} s_{pj} = \sum_{i=1}^{N} w_{ij} x_{pi} - \theta_j \\ b_{pj} = f(s_{pj}) \\ t_{pk} = \sum_{j=1}^{L} w_{jk} x_{pj} - r_k \\ c_{pk} = f(t_{pk}) \end{cases} \quad (2)$$

where x_{pi}, s_{pj}, t_{pk} is the input of input layer neuron i, hidden layer neuron j and output layer neuron k respectively; b_{pj} and c_{pk} is the output of hidden layer neuron j and output layer neuron k respectively; θ_j and γ_k is the threshold of hidden layer neuron j and output layer neuron k respectively; w_{ij} is the weight between input layer neuron i and hidden layer neuron j; w_{jk} is the weight between hidden layer neuron j and output layer neuron k, $i = 1,2,...,N$, $j = 1,2,...,L$, $k = 1,2,...,M$.

For the given learning sample set $(x_{p1}, x_{p2},..., x_{pN})$, $(y_{p1}, y_{p2},... y_{pM})$, $p = 1,2,...,P$, P is the sum of sample, the error function is:

$$E_p = \frac{1}{2} \sum_{k=1}^{M} [y_{pk} - c_{pk}]^2 \quad (3)$$

where: y_{pk} is the ideal output of output layer neuron k, c_{pk} is the real output of output layer neuron k.

The adjusting equation of the weight w_{jk} is (with momentum factor):

$$w_{jk}(t+1) = w_{jk}(t) - \eta \frac{\partial E_p}{\partial w_{jk}(t)} + \alpha (w_{jk}(t) - w_{jk}(t-1)) \quad (4)$$

where: η is learning rate, α is momentum factor.

3. The oscillatory characteristic of BP network error curve

3.1. The oscillatory characteristic of error convergence curve

Learning ratio η and momentum factor α have great influence on ANN learning speed. It is difficult to determine η and α, if it is selected too large, the learning speed is fast, but the error is easily oscillated, on the opposite, if it is too little, the learning or converging speed is too slow. Therefore, increasing learning speed and avoiding learning error oscillation is ambivalent. Reference [4] gave the theoretical equation of learning rate η, but the equation is complicated relatively, and its application is limited. Reference [5] also discussed the problem, but the neurons with less error are neglected in error function. so the learning error couldn't reflect the sum error of neuron network exactly. In this paper, the following equations are adopted[4]:

$$\eta(t+1) = \eta(t) + \Delta\eta(t)$$
$$\Delta\eta(t) = -\beta\Delta E(t) = -\beta[E(t) - E(t-1)] \quad \beta \in (0,1) \quad (5)$$

When error increment $\Delta E(t)$ is positive, learning rate will be decreased gradually, otherwise it will be increased gradually. The results of verification show that it can improve the network learning speed effectively. But it will cause error curve oscillate when momentum factors are selected improperly. Selects data in Tab.1 as training samples sets, the structure of network:9-12-9, the initial parameters: $\eta = 0.75$, $\beta = 0.15$, $\alpha = 0.85$. Fig.2 is the error curve.
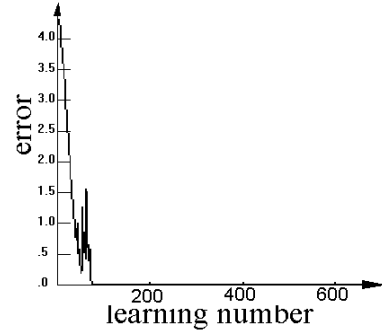

Fig.2: Error curve of BP with adaptive learning coefficient.

3.2. Analysis of error curve oscillation

At the beginning of ANN training, the error of each output layer neuron and total error of ANN are bigger relatively, when selects bigger learning rate and momentum factor, the ANN error curve can convergent rapidly. At training deeper time, although the error of some output layer neuron is larger still, a lot of output layer neuron error became less. If adopting the momentum factor of training beginning at this time, which will cause the output layer neuron with less error excessive learning, and the error of the neuron will appear fluctuation, further leads to total error of ANN appear fluctuation and oscillation.

Fault	Fault manifestations(f is working frequency)								
	0.0~0.39f	0.40~0.49f	0.50f	0.51~0.99f	1f	2f	3~5f	Odd$f$	>5f
Unbalance	0	0	0	0	0.90	0.05	0.05	0	0
Radial rub	0.10	0.10	0.10	0.10	0.20	0.10	0.10	0.10	0.10
Misalignment	0	0	0	0	0.40	0.50	0.10	0	0
Eccentricity	0	0	0	0	0.80	0.20	0	0	0
Rotor crack	0	0	0	0	0.40	0.60	0	0	0
Coupler fault	0.10	0.20	0	0.10	0.20	0.30	0.10	0	0
Subharmonic resonance	0	0	0	1.00	0	0	0	0	0
Oil whirl	0	1.00	0	0	0	0	0	0	0
Bearing seat looseness	0	0.70	0.20	0	0	0	0	0.10	0

Table 1: The training samples sets.

4. The learning algorithm of the new fast BP network

Selecting smaller momentum factor can avoid the oscillation of total error, but it can't improve network convergence speed effectively. Considering that error curve oscillation is caused by excessive learning of the output layer neuron with less error, this article presents a new learning algorithm. Through setting up neuron error threshold function, only adjusts the neuron which error is bigger than the error threshold during each training, doesn't adjust the neuron which error is smaller than the error threshold. It can avoid the excessive learning of the neuron with less error, so can avoid total error curve fluctuation.

Supposes the threshold function is:

$$\text{step}\left(\frac{y_{pk}-c_{pk}}{E_R}\right) = \begin{cases} 1 & |y_{pk}-c_{pk}| \geq E_R \\ 0 & |y_{pk}-c_{pk}| < E_R \end{cases} \quad (6)$$

where: E_R is threshold.

Error function adopts the Eq.(3). The adjusting equation of the weight between hidden layer and output layer is:

$$\begin{cases} w_{jk}(t+1) = w_{jk}(t) - \eta(t)\text{step}\left(\frac{y_{pk}-c_{pk}}{E_R}\right)\frac{\partial E_p}{\partial w_{jk}(t)} + \\ \qquad \alpha\text{step}\left(\frac{y_{pk}-c_{pk}}{E_R}\right)[w_{jk}(t)-w_{jk}(t-1)] \\ \frac{\partial E_p}{\partial w_{jk}(t)} = -(y_{pk}-c_{pk})c_{pk}(1-c_{pk})b_{pj} \end{cases} \quad (7)$$

The adjusting equation of the threshold of output layer is:

$$\begin{cases} r_k(t+1) = r_k(t) - \eta(t)\text{step}\left(\frac{y_{pk}-c_{pk}}{E_R}\right)\frac{\partial E_p}{\partial r_k(t)} + \\ \qquad \alpha\text{step}\left(\frac{y_{pk}-c_{pk}}{E_R}\right)[r_k(t)-r_k(t-1)] \\ \frac{\partial E_p}{\partial r_k(t)} = -(y_{pk}-c_{pk})c_{pk}(1-c_{pk}) \end{cases} \quad (8)$$

The adjusting equation of the weight between input layer and hidden layer is:

$$\begin{cases} w_{ij}(t+1) = w_{ij}(t) - \eta(t)\frac{\partial E_p}{\partial w_{ij}(t)} + \alpha[w_{ij}(t)-w_{ij}(t-1)] \\ \frac{\partial E_p}{\partial w_{ij}(t)} = [\sum_{k=1}^{M}(-d_{pk})w_{jk}]f'(s_{pj})x_{pi} \end{cases} \quad (9)$$

The adjusting equation of the threshold of hidden layer is:

$$\begin{cases} \theta_j(t+1) = \theta_j(t) - \eta(t)\frac{\partial E_p}{\partial \theta_j(t)} + \alpha[\theta_j(t)-\theta_j(t-1)] \\ \frac{\partial E_p}{\partial \theta_j(t)} = [\sum_{k=1}^{M}(-d_{pk})w_{jk}]f'(s_{pj}) \end{cases} \quad (10)$$

where $\eta(t)$ is learning ratio; α is momentum factor;

$d_{pk} = \text{step}\left(\frac{y_{pk}-c_{pk}}{E_R}\right)(y_{pk}-c_{pk})c_{pk}(1-c_{pk})$;

$\eta(t) = \eta(t-1) - \beta[E_p(t-1)-E_p(t-2)] \quad \beta \in (0,1)$;

$i = 1,2,...,N; j = 1,2,...L; k = 1,2,...,M$

With above adjusting equations, when the neuron's error between ideal output y_{pk} and real

output c_{pk} is bigger than threshold E_R, the parameters of the neuron is adjusted. When the neuron's error between ideal output y_{pk} and real output c_{pk} is smaller than threshold E_R, the parameters of the neuron isn't adjusted, and back propagation error of hidden layer neuron is zero.

How to select threshold E_R is very important. If it is too smaller, then needed adjusting neuron will be more, which will affect training speed and total error curve smooth. If it is too bigger, then needed adjusting neuron will be fewer, which will cause that ANN can't perform nonlinear mappings and recognize patterns. For simple fault diagnosis problems, because their output layer neuron are less, E_R can be selected smaller. For complex fault diagnosis problems, because output layer neuron are more, E_R can be selected bigger, then E_R will decrease with error decreasing.

Selects data in Tab.1 as training samples sets, the structure of network: 9-12-9, the initial parameters: $\eta = 0.75$, $\beta = 0.15$, $\alpha = 0.9$, $E_R = 0.022$. Fig.3 is the error curve, which shows that ANN convergence speed is fast and hasn't oscillation.

To avoid entering into the saturation range of the Sigmoid curve, when the practical output is smaller than 0.01 or bigger than 0.99, supposes its output value is 0.01 or 0.99 during network learning.
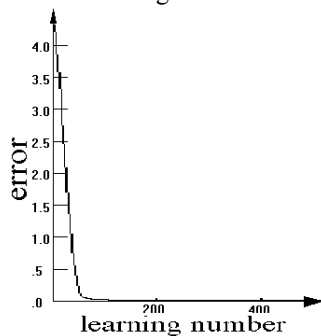

Fig.3: Error curve of fast BP algorithm.

5. Turbine-generator set vibration fault diagnosis based on the new BP algorithm

To verify the effective of the above algorithm, selects the nine faults in Tab.1. The network structure is 9-12-9, the initial parameters: $\eta = 0.75$, $\beta = 0.15$, $\alpha = 0.9$, $E_R = 0.022$. Fig.3 is the network learning convergence curve.

The vibration signal of the third bearing of the turbine-generator set in power plant is normalized in nine frequency range after FFT processing, and the Fault symptom eigenvector is [0,0,0,0,0.574, 0.346, 0.08,0,0], the fault diagnosis result is in Tab.2. According to Tab.2, we can diagnosis the fault is misalignment and eccentricity, the result has been confirmed by practical maintenance.

Tab.2 shows that the fast BP network identification performance is well than conventional BP network.

Fault	Result of the fast BP algorithm in this paper	Result of the self-adaptive BP algorithm in [4]
Unbalance	0.96e-5	0
Radial rub	0.080	0
Misalignment	0.982	0.850
Eccentricity	0.891	0.696
Rotor crack	0.006	0.054
Coupler fault	0.001	0.011
Subharmonic resonance	4.9e-4	1.2e-3
Oil whirl	2.6e-5	0
Bearing seat looseness	1.1e-4	4.0e-3

Table 2: The diagnosis result.

6. The rotor winding inter-turn short circuit fault diagnosis based on the new BP algorithm

Turbine-generator rotor winding inter-turn short circuit fault is a familiar electrical fault, and analyzing the mechanics and the on-line diagnosis method has important actual significance. Reference [6] analyzed electromagnetic characters, and revealed that reactive power reduces relatively whereas exciting current increases, and selected the relative error between theoretical value and measured value of exciting current as the fault criterion. But it is difficult to calculate exciting current exactly, because need for knowing all the generator parameter of model while the generator parameter is changing as operation method and condition. In order to solve this problem, a rotor winding inter-turn short circuit fault diagnosis method based on ANN is developed[7]-[,9]. Firstly, collects exciting current, active power and reactive power on normal condition as sample data to establish the generator neural network model, then inputs the measured active power and inactive power into the neural network that can identifies theoretical value of exciting current.

To increase fault diagnosis speed and accommodate to the requirement of fast diagnosis, the rotor winding inter-turn short circuit fault diagnosis based

on the error contracting gradually BP algorithm is established. Then linking to the actual running condition of power plant, develops the generator rotor winding short-circuit fault diagnosis system based on C/S mode which functions involve data acquisition, data base design and fault diagnosis etc..

5.1. Implementation of generator data acquisition system

Generator model number: QFS-300-2, rating data: S_N=353MW 、 U_N=18kV 、 I_f=1.85kA 、 f_N=50Hz 、 P=300MW、Q=186.035MVar.

(1) Hardware electric circuit of data acquisition system and implementation of Visual Basic

For implementing on-line data acquisition, changes dial instrument into a digital watch with RS485 digital output, an NPort is connected between computer and digital watch. Fig.4 is data acquisition system and background database, and Fig.5 is Nport installment and connection diagram.

(2) Data acquisition of generator exciting current

Collect generator exciting current in digital output port of exciting current meter. Query command data format of exciting current meter: Addr 04 00 00 00 03 CRC_{HI} CRC_{LO}, eight bits in all, where Addr is address number of apparatus; 04 is data query command; 00 00 is data initial address; 00 03 is polling word length, six bits in all; CRC_{HI} CRC_{LO} are separately high bits and low bits of CRC16 check code.

Response command data format of exciting current meter: Addr 04 02 $Data_{HI}$ $Data_{LO}$ CRC_{HI} CRC_{LO}, seven bits in all, where Addr is address number of apparatus; 04 is data query command; 02 is data length; $Data_{HI}$ $Data_{LO}$ are separately high bits and low bits of data. Data transmitted is hexadecimal number, which decimalizes at first, and then divides by the ratio that is the video data on dial plate, for an ampere meter, the ratio is 16^4.

(3) Data acquisition of generator power, phase current and line voltage etc.

Data of generator power, phase current and line voltage collect from digital output port of power meter. Query command data format of power meter: Addr 04 00 00 00 03 CRC_{HI} CRC_{LO}, eight bits in all, where Addr is address number of apparatus; 04 is data query command; 00 00 is data initial address; 00 08 is polling word length which query U_{AB}, U_{CB}, I_A, I_c, P, Q, cosα, F, eight words(sixteen bits) in all; CRC_{HI} CRC_{LO} are separately high bits and low bits of check code.

Response command data format of power meter: Addr 04 10H U_{ABHI} U_{ABLO} U_{CBHI} U_{CBLO} I_{AHI} I_{ALO} I_{cHI} I_{cLO} P_{HI} P_{LO} Q_{HI} Q_{LO} $cosα_{HI}$ $cosα_{LO}$ F_{HI} F_{LO} CRC_{HI} CRC_{LO}, twenty one bits in all, where P_{HI} P_{LO} are

separately high bits and low bits of data, other are consistent to P. Data transmitted decimalizes at first, and then multiplies ratio that is the value in accord with dial plate. The specific ratio is $U/16^2$, $F/16^2$, $P/16$, $cosα/16^4$.

Using Output attribute of MSComm control to send query command to power meter, instruction format: address、command、initial code、length and CRC, for example, 3B04000000011496. In order to check whether data has mistakes in the course of transmission, source and receiving terminal all adopt CRC16 to check data. In the same way to check whether data has mistakes in the course of transmission, source and receiving terminal both adopt CRC16 to check data. After power meter receives the right instruction, send data that is as alike as exciting current meter.
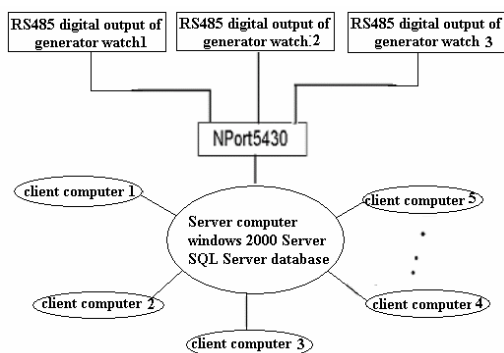

Fig.4: Data acquisition system.

Fig.5: Nport installment and connection diagram.

5.2. Diagnosis system of generator rotor winding inter-turn short circuit fault based on ANN

After generator overhauling (considering the generator of normal operation here), collect 21120 data, and distills 312 samples as training samples of the neural network according to the distribution of the collected

data. Fig.6 is dispersing spot diagram of sample. Active power changes from 230MW to 320MW, and inactive power changes from 120Mvar to 190MVar. According to annual operation data of generator, cover the normal operation extent of generator basically.

ANN structure is 2-30-1, two input layer neurons represent separately active power and reactive power, output layer neuron represents exciting current. Adopt the error contracting gradually training algorithm, training rate original value $\eta = 0.75, \beta = 0.15$; momentum factor $\alpha = 0.9$, E_R =0.002; training times is 5000; the whole system error is 5.3034701e-002. Then look the neural network as generator mathematic model, enter generator operation measured data into the neural network that can distinguish the corresponding theoretical value of exciting current which compare with measured exciting current, judge whether rotor excitation winding inter-turn short-circuit fault happen according to the relative error. Tab.3 is a part of results that measured data distinguished. According to running condition actually, the threshold quantity is selected 4%. The on-line diagnosis system of rotor winding short circuit fault is successfully applied to ＃3 generator of certain power plant for nearly a year which run steadily, the fig.7 is the interphase of the system.
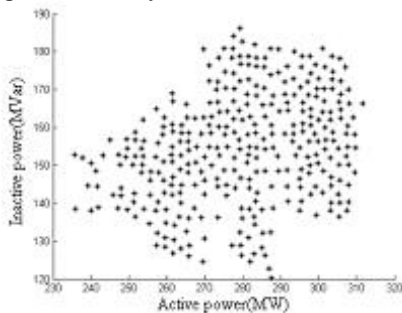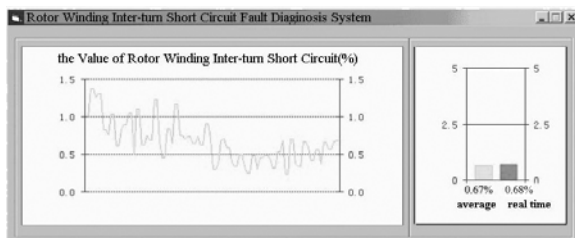

Fig.6: The disperse spot diagram of sample.

Fig.7: The interphase of the system.

6. Conclusions

(1) In conventional BP network, when learning rate and momentum factor are selected improperly, will cause learning error curve oscillation and affect learning stability and precision. The reason is that each output neuron error is different during network learning. If adopting the same learning rate and momentum factor, which will cause the neuron with less error excessive learning, and it's error will oscillate, and causes total error fluctuation and oscillation.

(2) A new fast BP algorithm is presented. Through setting up neuron error threshold function, only when neuron's error is bigger than the error threshold, the neuron's parameters were adjusted, the algorithm can avoid the excessive learning of neural network and learning error oscillation, and increase BP network learning speed.

(3) The new BP algorithm is applied successfully to the turbine-generator set vibration fault diagnosis and the rotor inter-turn short-circuit fault diagnosis. The results of verification show that the algorithm can assure steady learning process, avoid error curve oscillating, and has faster learning speed and higher diagnosis precision.

Order	Active power (MW)	Inactive power (Mkvar)	Identification exciting current i_{fo} (theoretical value)(kA)	Measured exciting current i_{fc} (kA)	Relative error (%) $\frac{\lvert i_{fc} - i_{f0} \rvert}{i_{f0}}*$
1	311.692	166.293	1.514	1.511	0.198
2	309.540	158.267	1.461	1.451	0.684
3	309.127	162.638	1.454	1.484	2.063
4	307.986	166.280	1.492	1.500	0.536
5	307.898	170.076	1.481	1.514	2.228
6	305.875	162.512	1.480	1.479	0.068
7	305.552	140.207	1.347	1.379	2.376

Table 3: The results of signal sample and training of the generator.

Acknowledgements

References

[1] C.Z. Chen, Q. Li, Y.F. Liu, and G. Ya, Intelligent fault diagnosis method for turbo-generator unit. *proceedings of the CSEE*. 5:121-124,2002.

[2] P.X. Yuan, Y.X. Xu, E.D. Liu, Comparison between POWELL algorithm and BP algorithm in training neural network. *proceedings of the CSEE*, 2:103-106,2002.

[3] W. Wang, *Artificial neural network theory----introduction and application*. Beijing University of Aeronautics & Astronautics Press.1995.

[4] Y. Li, H.Y. Li, R.X. Ye, BP network algorithm with self-adaptive learning rate and its use in the fuzzy diagnosis of turbo-generator failures. *Journal of Engineering for Thermal Energy and Power*, 6: 455-458,1997.

[5] Z.Y. Wang, Z.H. Hong, The improved BP algorithm and its application on pattern recognition. *Journal of Nanjing University of Aeronautics & Astronauts*, supplement: 216-223,1994.

[6] Y.G. Li, H.M. Li, H. Zhao, The new criterion on inter turn short-circuit fault diagnose of steam turbine generator rotor winding. *Proceedings of the CSEE*,6:112-116,169,2003.

[7] M. Vilaragut, A. Costa, X.M.L. Fernandez, Diagnostic of turbo-generator rotor short-circuit turns utilizing artificial neural network. *15th International Conference on Electrical Machines*, 2002.

[8] Marcos F. S. V. D'Angelo; Pyramo P. Costa, Jr. Detection of shorted turns in the field winding of turbogenerators using the neural network MLP . *IEEE interantional Conference on Systems, Man, and Cybernetics,* pp.1930-1935,2001.

[9] J.Y. Ding, Y.P. Chen. On-line diagnosis for shorted field-turns of synchronous generator based on artificial neural networks. *International Conference on Power System Technology*,pp. 1675-1678,2000.