# Power transformer diagnostic prediction research based on quantum neural networks and evidence theory

**Qiang Song[1*], Ai-min Wang[2]**

[1]*School of Mechanical Engineering, Anyang Institute of Technology, Henan, Anyang City 455000, China*

[2]*School of Computer and Information Engineering, Anyang Normal University, Henan, Anyang City 455000, China*

*Received 1 March 2014, www.cmnt.lv*

**Abstract**

Aiming at the fault of power transformer fault information diversity and uncertainty, a large amount of data and no regularity characteristics, a new fault diagnosis method of quantum neural network based on information fusion. In order to accurately and effectively identify transformer fault model, combining the quantum neural network and evidence theory combination of transformer fault diagnosis. A quantum neural networks to collect data on the macroscopic, microscopic quantum corrections in the interval of fuzzy intersection data according to a certain proportion of the rational allocation of the associated mode, so as to improve the accuracy of pattern recognition; use of the evidence theory can improve the convergence speed of quantum neural networks. The results were compared with the diagnosis and BP neural network input, that this method has a higher accuracy in transformer fault pattern recognition.

*Keywords:* transformer diagnostic prediction, combination forecasting, quantum neural networks, evidence theory

## 1 Introduction

Transformer plays an essential role in power system. A lack of coherent maintenance strategy may result in accidence of power system and reduce the reliability. Online monitoring for power transformer is important for safe operating. In order to keep transformer in good condition, diagnosis has become increasingly essential [3-5]. Although there are different methods used for detecting transformer fault. Dissolved gas in transformer oil is widely used as a reliable approach. In this paper, we present the RBF networks with self-adjustable number of hidden neurons for transformer faults detection. A new type of SOM RBF network together with its training algorithm is proposed [6, 7].

As a result, this has made the implementation of an effective neural network used in transformer fault detection system easy. The advantages of the proposed RBF neural networks are twofold. First, the best possible network architecture is determined by E-mail: author@domain.com the proposed training algorithm according to the input data. It does not require many trial tests. Second, the outputs of the neural network are able to not only perform fault detection, but also indicate the extent of the fault. In the first advantage, a cell-splitting grid (CSG) neural network, which is an extended SOM, is used to automatically determine the centres and the number of hidden neurons of the RBF networks. After completion of the training, the learned network is able to detect different types of faults. Our obtained results indicate that the proposed neural network approach is promising for diagnosing transformer faults via analysing the dissolved gases in transformer.

Transformer is an indispensable equipment of power system, whose fault state directly affects power supply reliability and system's operation. As the power network develops towards highly automation, growing demand has been put forward to the liability of power system. It's an urgent requirement to improve present equipment maintenance system. Therefore, increasing attention is being paid to state maintenance system based on online condition monitoring and fault diagnosis. It is a trend for this state maintenance system to replace the preventive maintenance system. Measurement and studies show that, the use of gas chromatography (Dissolved Gas Analysis) for transformer latent faults analysis is one of the most effective measures to ensure transformer operation, which has been the most effective way for equipment fault diagnosis at home and abroad [1]. Currently, Improved three-ratio method (formerly improved electrical committee agreements) is the main method to diagnose internal faults of transformer.

Among the various transformer fault diagnosis measures, dissolved gas analysis (DGA) is based on the principle that different types of transformer faults correspond to different dissolved gas concentrations. It detects transformer's latent fault by analysing the concentration of fault characteristic gases ($H_2$, $CH_4$, $C_2H_2$, $C_2H_6$, $C_2H_4$ etc.). Besides, DGA method can make the diagnosis under energized condition. Thus, it can periodically make internal diagnosis of transformer during its operation, without interference from external

[*]***Corresponding author*** e-mail: songqiang01r@126.com

electromagnetic field. So this method has been commonly and widely used in fault diagnosis in electrical equipment's [2].

## 2 Quantum neural networks

Quantum Neural Network appeared in 1990s. With the idea of quantum mechanics introduced into neural network study, it can overcome the flaw and deficiency of conventional neural network. Quantum neural network is the extension traditional neural network. By taking some advantages of quantum computations, parallel computation especially, quantum neural network has more parallel processing power, capable of handling much larger data sets. Therefore, quantum neural network has unprecedented advantages in data processing.

Currently, quantum neural network has been applied to fault diagnosis for power electronic circuit [6], soft fault diagnosis of tolerance analogue circuits [7], speech noise reduction [7] and so on. Compared with traditional neural network, quantum neural network has the following advantages: Set the input value of the system as $X=(X_1,X_2,...,X_n)$, output value as $Y=(Y_1,Y_2,...,Y_n)$. The action function of three-layer $\sigma$ is Sigmoid function, and its weights and number neurons intervals, $\theta_v(v=1,2,3,\ldots,ns)$ is quantum interval. Its size is the same as the number of fault modes to be diagnosed, namely, the same as the number of fault components. *sf* is the steepness factor.

a) Exponential memory capacity and memory speed [8,9];

b) high-speed learning and information processing capabilities;

c) capable of eliminating catastrophic amnesia, with no mutual interference between modes.

High stability and reliability. Quantum neural network is the combination of quantum computation and traditional neural network. Generally, there are two integrated forms:

a) to introduce quantum computing theory in the structure and training process of neural networks;

b) to design topology and training algorithm of neural network by directly borrowing some principles and definitions of quantum theories.

The structure and learning algorithm of quantum neural network are described below.

### 2.1 STRUCTURE OF QUANTUM NEURAL NETWORK

The structure of quantum neural network is shown is Figure 1, which includes the input layer, the first hidden layer, the second hidden layer and the output layer respectively $\omega 1_{k,n}$, $\omega 2_{m,k}$, $\omega 3_{s,m}$, $K$, $M$, $S$.

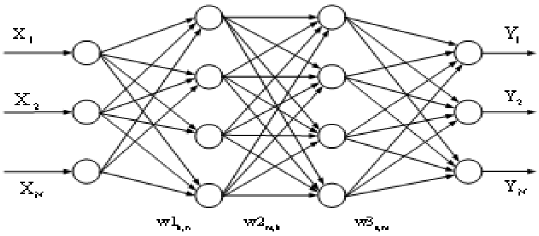


FIGURE 1 The structure of four-layer quantum neural network

Based on its network structure, the network output is:

$$Y_s=\sigma\left(\sum_{n=1}^{N}\omega 3_{s,m}\sigma\left(\sum_{k=1}^{K}\left(\omega 2_{m,k}\times\frac{1}{ns}\times\sigma\left(sf*\sum_{n=1}^{N}\omega 1_{k,n}x_n-\theta_v\right)\right)\right)\right), \quad (1)$$

where $s=1,2,3,\ldots,S$; $k=1,2,3,\ldots,K$; $m=1,2,3,\ldots,M$. Set the total number of input samples as P, output expected value as $\hat{Y}_s^p$, based on Equation (1), the *s-th* node, *P-th* sample of network output layer is note as $Y_P^S$, then the error energy function expression can be defined as:

$$e=\frac{1}{2}\sum_{s=1}^{S}\sum_{p=1}^{P}\left(\hat{Y}_s^p-Y_s^p\right)^2, \quad (2)$$

where the partial derivative of $e$ with respect to $\omega 1_{k,n}$, $\omega 2_{m,k}$, $\omega 3_{s,m}$ can be derived as:

$$\frac{\partial e}{\partial w3_{s,m}}=-\sum_{p}\left(Y_S^P-\hat{Y}_S^P\right)\times\sigma'\times \sigma\left[\sum_{k=1}^{K}w2_{m,k}\times\frac{1}{ns}\times\sigma\left[sf\sum_{n=1}^{N}w1_{k,n}x_n-\theta_v\right]\right], \quad (3)$$

$$\frac{\partial e}{\partial w2_{m,k}}=-\sum_{p}\sum_{s}\left(Y_S^P-\hat{Y}_S^P\right)\times\sigma'\times w3_{s,m}\times\sigma'\times\frac{\sigma}{ns}\left[sf\sum_{n=1}^{N}w1_{k,n}x_n-\theta_v\right], \quad (4)$$

$$\frac{\partial e}{\partial w1_{k,n}}=-\sum_{p}\sum_{s}\sum_{m}\left(Y_S^P-\hat{Y}_S^P\right)\times\sigma'\times w3_{s,m}\times\sigma'\times w2_{m,k}\times\frac{sf}{ns}\times\sigma' x_n^P. \quad (5)$$

### 2.2 LEANING ALGORITHM OF QUANTUM NEURAL NETWORKS

In network learning algorithms, the weight update between neurons is same as the conventional BP algorithm, which obtains changes and error back propagation by gradient descent method. In addition, to accelerate the training speed and avoid trapping into the local minimum, additional momentum and adaptive learning rate method are introduced, which enables the avoidance of network's local minimum and rapid convergence. So we have:

$$\begin{cases} w = \{w1_{k,n}, w2_{m,k}, w3_{s,m}\} \\ \Delta w = \{\Delta w1_{k,n}, \Delta w2_{m,k}, \Delta w3_{s,m}\} \end{cases}, \tag{6}$$

where $\Delta w$ is the weight's update amount. Then the improved momentum equation of back propagation can be shown as:

$$w(t+1) = w(t) - \Delta w(t), \tag{7}$$

$$\Delta w(t) = mc \cdot \Delta w(t-1) - (1-mc) \cdot lr \cdot \frac{\partial e}{\partial w}, \tag{8}$$

where $w(t-1)$ is the weight before training and $w(t+1)$ after training; mc, lr are respectively momentum factor and learning rate. So they can be updated as below:

If $e(t) > e(t-1) \times em$ then:

$$mc = mc_1, \; lr = lr \times d_{1r}. \tag{9}$$

If $e(t) < e(t-1)$ then:

$$mc = mc_2, \; lr = lr \times i_{1r}, \tag{10}$$

where $em$ is the maximum error rate; $mc_1$, $mc_2$ are momentum factors; $d_{1r}$, $i_{1r}$ are decrement and increment of learning rate respectively And the update equation of quantum interval is:

$$\langle o_{i,q} \rangle = \frac{1}{|c_q|_{x_k \in C_q}} \sum_{x_k \in C_q} o_{i,k}, \; v_{i,k,s} = o_{i,k,s}(1 - o_{i,k,s}),$$

$$\langle v_{i,q,s} \rangle = \frac{1}{|c_q|_{x_k \in C_q}} \sum_{x_k \in C_q} v_{i,k,s},$$

where, $\Delta\theta_v$ is the update amount of quantum interval; $o_{i,k}$ is the output value of the *i-th* neuron of the first hidden layer when the input vector is $x_k$; $o_{i,k,s}$ is the output value of the *i-th* neuron of the first hidden layer when the input vector is $x_k$, the quantum interval as *s*.

### 3 Evidence theory

Evidence theory, proposed by Dempster in 1967 and developed by Shafer in 1976, is also known as D/S evidence theory. It can be used to deal with the uncertainty caused by the Unknown. With belief function as as a metric, it can build belief function by relaxing probability restriction of event, with no concern for the precise and inaccessible functions. Once the constraint is limited to strict probability, this theory becomes probability theory. Due to its strong ability of handling uncertain information, D/S evidence theory has always been the important means for multi-sensor information fusion. However, its foundation, BPA (basic probability assignment) is not easily determinable, which constrains the application of Evidence theory. This study proposed a fault diagnosis based on the information confusion with the integration of various intelligence theories and DA evidence theory. The respective evidence body is composed by SVM, grey relational entropy grey relation entropy, and D/S evidence theory is used to complete evidence integration for partial diagnosis of diagnosis module groups. Finally, the decision diagnostic results can be get. The organic combination of such intelligence theory as support vector machine and Evidence theory has improved the comprehensive diagnosis performance and fault diagnostic effects of complex systems.

D/S evidence theory, proposed by Dempster and then developed by his students Shafer, is an extension of the classical probability theory. According to this theory, Collect all the evidences might influence the Assumption, and then divide them into some relative independent Meta evidence (evidence ingredient with single factor), thus forming an evidence space. Next, assign all the possible combinations of meat evidences with a value satisfying certain constraints weaker than probability restriction. Finally, a function defined in power set of evidence space can be obtained, called basic probability assignment function.

Since Subset of evidence space is not independent and the constraint condition is weaker than probability, basic probability assignment function is not probability. In order to obtain reliability similar to probability, another function based on basic probability assignment function is designed, namely the genus probability function. It meets the constraint condition similar to probability, can be used to represent the reliability of the evidence. Meanwhile, the Unknown and Uncertainty can be distinguished by basic probability assignment function, which is more profound and detailed description of Inaccuracy.

Set *U* as a recognition framework, which contains all the possible results known by people. *A* is a subset of *U*. Then function $m: 2 \to [0,1]^U$ meets the following conditions:

1. $m(\phi(0)) = 0$;
2. as $\sum_{A=U}(A) = 1$, *A* is said to be the basic probability assignment of *m*(*A*), and function $Bel :\to [0,1]$, $Bel(A) = \sum_{B \subset A}(B)(\forall A \subset U)$, the belief function on *U*.

If $m(A) \geq 0$, then *A* is termed focal element of belief function *Bel*, the union of all focal elements are called kernel. For combination of evidence groups, Dempster-Shafer has provides the following rules. Set $m_1, m_2$ as two basic probability assignments defined in the same recognition framework *U*, $Bel_1$ and $Bel_2$ are the belief functions in *U*, and the focal elements are $A_1,\ldots,A_k$ and $B_1,\ldots,B_r$ and their combinations can get a new basic belief assignment [11-16], when $A_i \cap B_j = \phi$, $K = \sum_{i,j} m_1(A_i) m_2(B_j)$, $M(P) = \sum_{i,j} m_1(A_i) m_2(B_j)$, $\forall A_i \cap B_j = P$, $P \neq \phi$, when $P = \phi$, $m(P) = 0$

Taking into account the mentioned above fusion equation, separate evidence from different sources can be combined to obtain more accurate information.

Plausibility function: the plausibility function of proposition $A$ is: $PL(A) = 1 - Bel(\overline{A}) = \sum_{A \cap B = \phi} (B)$.

Plausibility function is also called upper limit function, which represents uncertainty measure for trust degree of proposition. Belief function and plausibility function constitute belief interval of proposition $A$.

## 4 Combination forecasting based on quantum neural networks and evidence theory

To compensate for the shortcomings of single diagnostic method, this study combines quantum neural networks with information fusion to make a more accurate comprehensive diagnosis.

For combination of quantum neural networks and information fusion, the key point is to take the output value of quantum neural networks as an evidence, and then make a comprehensive diagnosis by DS evidence theory. The specific implementation method is: firstly, quantum neural networks fusion is processed to get the preliminary fusion results; then normalization for these results are carried out; next take the output value as basic probability assignment for proposition in D/S theory recognition framework; finally, apply the make a comprehensive diagnosis by combination rules. Figure 2 shows repeated fault diagnosis mode for transformer, which is also the overall research program structure of this study.

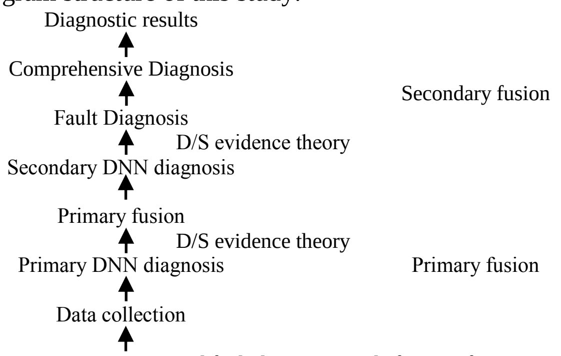


FIGURE 2 Repeated fault diagnosis mode for transformer

This is a secondary fusion structure. On the basis of the primary fusion structure which makes transformer fault diagnosis with quantum neural networks and Evidence theory, this secondary structure makes repeated fault diagnosis with the two fusion algorithms.

In the primary fault diagnosis, the input information combining D/S theory and quantum neural networks is concentration of $H_2$, $CH_4$, $C_2H_6$, $C_2H_4$ $C_2H_2$, expressed in 5 neurons. There are two kinds of input forms: input directly various gas concentrations or normalize samples. While the former easily leads to excessive large sample space, which in turn, leads to oversized network scales, affecting network's normal training and diagnosis. So this study selects the latter form, and takes the ratio of gas concentration to the sum of 5 gas concentrations as the input information. The optimal number of the hidden layer nodes is sought on actual situation. There are 6 neurons in the output layer, representing 6 fault types that networks can diagnosis, i.e. Low-temperature overheat, Middle-temperature over, High-temperature overheat, partial discharge, spark discharge and arc discharge.

Based on the primary fusion, the input information of second fault location diagnosis combining D/S theory and quantum neural networks is respectively: low-temperature overheat, middle-temperature over, high-temperature overheat, partial discharge, spark discharge and arc discharge; the optimal number of the hidden layer nodes is sought on actual situation.

## 5 Experimental study

The encoding and decoding problem is one of the basic problems of neural network training. In this study, a 10-23-10 codec is trained to test the convergence properties of improved quantum neural network mode. There are 10 different input modes, each mode with only one bit as 0 and the rest as 1. The input mode is required to be the network's output value. Set the target literacy accuracy as 0.001, replacement learning rate of weight and threshold value as 0.9, replacement learning rate of quantum interval as 0.7. The initial weight and threshold value are selected randomly. Below is the convergence curve (Figure 3).

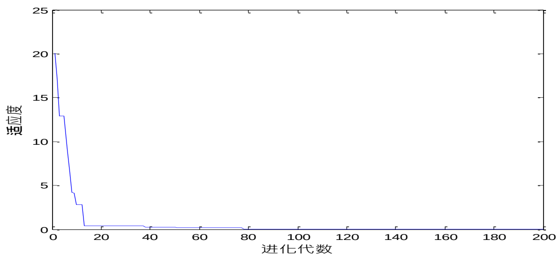


FIGURE3 Convergence curve of quantum neural networks.

As it is seen in Figures 4 and 5, 200 groups of data are selected as samples, among which 33 are test data. After network training, the training data is put into the trained quantum neural networks as input values. As a result, only two samples are misjudged. Moreover, when verified with training samples, also only two samples - two types of transformer fault - are misjudged, a diagnosis precision of 93.7%. The resulting quantum neural network can be used for more sample forecasts.

To select fault features of neural network modes, it is required the maximal fault information be concluded in fault feature samples. Therefore, depth analysis of fault mechanism and fault information transfer relationship is needed. Furthermore, select the best indication of fault characteristics and dismiss irrelevant ones, to ensure the generation of the smallest quantum neural networks.

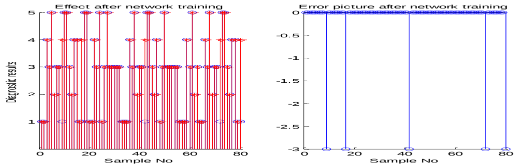


FIGURE 4 Transformer fault classification effect based on quantum neural networks (training sample data)

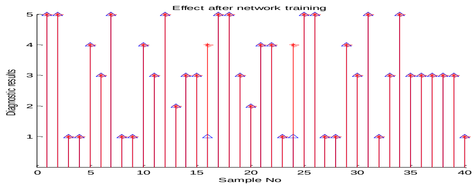


FIGURE 5 Transformer fault classification effect based on quantum neural networks(test data)

In this study, for the same fault, BP neural network is applied to analyse transformer fault diagnosis on the same samples. It includes tree network structures: the input lay, the hidden layer and the output layer. 5 input nodes and 6 output nodes are used based on original data.

After repeated training with BP network gradient descent algorithm of BP network, the result shows that 17 is the optimal number of nodes for the hidden layer, with faster convergence and smaller network errors.

## 5 Conclusions

By using quantum computation theory, quantum neural networks have more parallel processing power and larger storage capacity than ANN. Quantum theory is the outcome of classical physics to the micro-level. Quantum system is the microscopic system basis of all the physical process, also the basis of biological and psychological ones, which has the similar dynamic characteristics as biological neural networks. By combing quantum theory with ANN, Quantum neural networks can better simulate and interpret human brain's information process, which is a quantum extension and evolution of ANN.

## Acknowledgements

For the completion of my BA thesis, first, I wish to express my deepest gratitude to my supervisor – professor Yang Cheng-zhi, who has given me the most valuable suggestions and advice, and made necessary corrections. Then I am greatly indebted to professor Wang Ai-min for his advice on the format and help with my computerization and I also owe a lot to professor Wang Ai-min who has shown much consideration for my composition and has provided me with some useful materials. Finally, I would like to express my thanks to my colleagues Wu Yao-chun, who have generously offered their help with my computerization.

## References

[1] Cao D 2005 Gas Analysis Diagnostics and Troubleshooting in Transformer Oil *Beijing: China Electric Power Press* 2005 *(in Chinese)*
[2] Zhou D, Ye Y 2000 Modern Fault Diagnosis and Fault Tolerant Control *Beijing: Tsinghua University Press (in Chinese)*
[3] Wang Dg, Zhang Y 2004 Intelligent Fault Diagnosis System *Beijing: Metallurgical Industry Press (in Chinese)*
[4] Jiao L 2005 Application and Implementation of Neural Networks *Xi'An: Xidian University Press (in Chinese)*
[5] Zheng R 2007 Fault Diagnosis Technology of Power Transformer Based on Grey System Theory *Jilin: Jilin University (in Chinese)*
[6] Gao Y 2007 Power Transformer Fault Diagnosis and Life Prediction Method *Xi'An: Xidian University (in Chinese)*
[7] Yin T 2006 Research on Transformer Fault Diagnosis Based on Grey Neural Network *Electrotechnical Application* **25**(10) 15-17 *(in Chinese)*
[8] Ventura D, Martinez T 1997 An artificial neuron with quantum mechanical properties *Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms* 482-5
[9] Behrman E C, Steck J E, Skinner S R 1999 A spatial quantum neural computer *Proceedings of the International Joint Conference on Neural Networks* Washington D.C.

**Authors**

**Qiang Song, Xintai City, Shandong Province, China**

**Current position, grades:** associate professor at Anyang Institute of Technology.
**University studies:** Master's degree at Kunming University of science and technology in 2006.
**Scientific interests:** intelligence control theory and rough set theory.

**Ai-min Wang, born in 1957**

**Current position, grades:** professor at Anyang normal college.
**University studies:** Doctor's degree at Wuhan institute of Technology.
**Scientific interests:** SVM and rough set of intelligent decision-making research.