International Journal of Applied Mathematics in Control Engineering 2 (2018) 194-199

Contents lists available at YXpublications

International Journal of Applied Mathematics in Control Engineering

Journal homepage: http://www.ijamce.com

Application Research of Combining Gaussian Mixture of ART2 Network in Transformer Fault Diagnosis

Guping Zheng, Biao Ma*, Gang Li

Department of Control and Computer Engineering, North China Electric Power University, 071000, China

ARTICLE INFO

Article history:
Received 3 August 2018
Accepted 21 October 2018
Available online 25 December 2018

Keywords:
adaptive resonance neural network
pattern recognition
Gaussian hybrid clustering
transformer fault diagnosis

ABSTRACT

Since the adaptive resonance neural network (ART2) solves the dilemma of other artificial neural networks in terms of "adaptation" and "stability", it has been widely used in pattern recognition. However, the ART2 network uses the “hard competition” method for classification, which leads to the degradation of classification accuracy. The Gaussian Mixture Clustering (GMM) algorithm is proposed to “soften” the clustering results of the ART2 neural network. The simulation results of a transformer operation data show that the GMM-ART2 model has a correct rate of fault diagnosis of 90%, which is significantly improved compared with the traditional ART2 neural network classification results. It is verified that the model constructed in this paper can provide a new effective means for transformer fault diagnosis.

1. Introduction

The transformer is an important device of the power system, and its operating state directly affects the safety level of the system. With the development and gradual promotion of transformer condition monitoring, it is quite necessary and feasible to establish a fault diagnosis system based on online monitoring information of transformer status. Since the composition and content of dissolved gases in transformer oil can reflect the operating state of the transformer to a large extent, Dissolved Gas Analysis (DGA) has become an effective method for fault diagnosis of oil-immersed transformers[1-3]. Based on this, traditional methods such as three-ratio method and Rogers method are formed[4],and artificial neural networks (ANN)[5-6], support vector machine (SVM)[7],Deep learning (DL)[8]Supervised intelligent methods applied to power transformer fault diagnosis. The emergence of intelligent methods enables the computer to quickly judge and diagnose unknown information, but when new fault types appear or fault data is missing, the above intelligent methods cannot be effectively and quickly identified. This makes it difficult to get an effective application based on intelligent diagnosis methods with supervised learning.

Adaptive resonance theory is a learning mechanism of a non-instructor-competitive neural network proposed by Professor Stephen Grosserg of the Adaptive Systems Center of Boston University in the United States based on a large amount of biological research[9].The neural network mainly includes ART1, ART2, ART3, ARTZA, FART and PART. Because ART2 function is powerful and computationally intensive, it is widely used in pattern clustering and recognition, data mining, process identification and data fusion[10],However, in some application scenarios, the ART2 neural network needs to be improved to suit specific needs[11].

In order to make up for the lack of clustering of sample data by “hard competition” in the traditional ART2 neural network, the literature [13] is based on the classification method of Finnish scholar Luukka using the Yu norm for similarity measure analysis. Combined with the ART2 neural network and applied to the diagnosis of bearing faults, good results have been achieved[12]. Literature [14] proposed to use C-means clustering to “soften” the clustering results of ART2 neural network, and applied the improved network to wind turbine gearbox fault diagnosis, which is superior to traditional ART2 neural network. Clustering results. However, there is no neural network application for the "hard clustering" improvement in ART2 in transformer fault diagnosis. For this reason, an unsupervised classification method of ART2 neural network (GMM-ART2) combined with Gaussian mixture clustering is proposed in ART2. The initial parameters of the Gaussian mixture model are calculated on the result of the class, and the probability that each sample data belongs to each Gaussian model "component" is obtained by the Gaussian mixture model, and the most probable probability is selected as the final classification

* Corresponding author.
E-mail addresses: mabiao1993@126.com (B. Ma)

result. This can effectively solve the problem that the classification accuracy of the ART2 neural network is reduced due to the internal use of the "hard competition" mechanism[14].

2. Adaptive resonance neural network combined with Gaussian mixture clustering(GMM-ART2)

2.1. Adaptive Resonance Neural Network (ART2)

It can be seen from the topology diagram of ART2 neural network[15] (as shown in Fig. 1) that the whole network includes three parts, which are comparison layer F_1, identification layer F_2 and decision subsystem. The A layer contains six sub-layers (w,x,v,u,p,q) and three gain control units (shown as black solid dots in the figure).

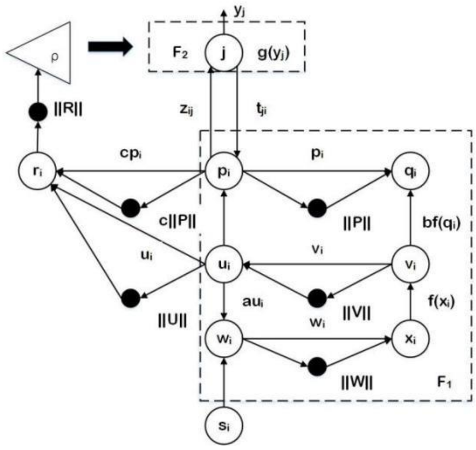

Fig.1. schematic diagram of ART2 neural network topology

In the F_1 layer, the pre-processed input signal s_i, through vector normalization and nonlinear transformation $f(x)$, obtains a stable intermediate layer mode U, and the enhanced information is sent to the F_2 layer by the upper layer mode P. Considering the case of the Gth node i, the following dimensionless equation characterizes the STM -variable $p_i, s_i, u_i, v_i, w_i, q_i$ calculation method:

$$p_i = u_i + \sum g(y_i)t_{ji} \quad q_i = \frac{p_i}{e+\|P\|} \quad u_i = \frac{v_i}{e+\|V\|}$$

$$v_i = f(x_i) + bf(q_i) \quad w_i = I_i + au_i \quad X_i = \frac{w_i}{e+\|W\|}$$

where $\|P\|$, $\|V\|$ and $\|W\|$ are the L_2 norms of P, V and W, respectively; a,b is the constant of the neural network, which characterizes the internal feedback of the F_1 layer, affecting the speed at which the intermediate layer U approaches the input mode; t_{ji} is the network from the top. The connection weight of the lower layer F_2 to the p layer, that is the LTM coefficient, y_j is the variable of the first j node in the F_2 layer.

The nonlinear signal transformation function is $f(x) = \begin{cases} 0 & 0 \le x \le \theta \\ x & x \ge \theta \end{cases}$. The above equations represent the loop iterative process of the F_1 layer. After the processing of this layer, the noise in the input signal is suppressed and the characteristic signal is enhanced.

In the ART2 model, the key role of the F_2 layer is to improve the contrast of the F_1 to F_2 layer filtered input mode and issue a reset signal. Among them, contrast enhancement is achieved through competition. The F_1 layer outputs a signal to the F_2 layer, wherein the input to the first j node in the F_2 layer is $T_j = \sum_i p_i z_{ij}$, where z_{ij} is the LTM coefficient connecting the F_1 and F_2 layers from bottom to top.

If the j th node in the F_1 layer is activated and the other nodes are in the suppressed state, then there are:

$$g(y_j) = \begin{cases} d & T_i = \max\{T_i\} \\ 0 & other \end{cases} \tag{1}$$

where j is the j th F_2 -layer node has not been reset in the current experiment; d is the constant of the F_2 -layer feedback back to the F_1 layer.

The learning of the weight coefficients z_{ij} and t_{ji} between the F_1 and F_2 layers is a continuous iterative process. The operation process is as follows:

From bottom to top（$F_1 \to F_2$）: $\frac{d}{dt} z_{ij} = g(y_{ij})[p_i - z_{ij}]$

From top to bottom （$F_2 \to F_1$）: $\frac{d}{dt} t_{ji} = g(y_j)[p_i - t_{ji}]$

The matching degree r_i and the similarity measure R between the STM mode at F_1 and the activated LTM mode are:

$$r_i = \frac{u_i + cp_i}{e + \|U\| + \|cP\|} \tag{2}$$

$$R = \frac{[\|U\|^2 + 2c\|U\| \cdot \|P\| \cos(U,P) + c\|P\|^2]^{1/2}}{\|U\|^2 + c\|P\|} \tag{3}$$

where $\frac{\rho}{e+\|R\|} < 1$, indicating that the new sample and the Class j similarity in the F_2 layer meet the requirements, the matching is successful; On the other hand, if $\frac{\rho}{e+\|R\|} > 1$, the similarity between the new sample and the j th in the F_2 layer does not meet the requirements, then F_2 resets, the system resets, re-finds the pattern class or defines a new class. In the formula, ρ is a warning parameter, the value range is 0~1, and the closer the ρ is to the value 1, the finer the classification, and the closer to 0, the coarser the classification.

2.2. Gaussian mixture model(GMM)

The Gaussian mixture model means that in the mixed model, the data to be clustered is regarded as a mixed probability distribution from multiple normal distributions, which represent different classes (Fig. 2), so that The correlation parameters (mean and covariance matrix) of each normal distribution are used as class prototypes[16].

Suppose the d -dimensional random variable $X=(x_1^T,x_2^T,\cdots,x_n^T)^T$ has a finite mixed normal distribution. If x is one of the samples, its probability density function is:

$$f(x\mid\psi)=\sum_{j=1}^{G}\pi_j f_j(x\mid\theta_j)$$

where

$$f_j(x\mid\theta_j)=(2\pi)^{-\frac{d}{2}}\left|\sum{}_j\right|^{-\frac{1}{2}}\exp\left(-\frac{(x-u_j)^T\sum_j^{-1}(x-u_j)}{2}\right)$$

$$j=1,2,\cdots$$

G the probability density function of the first j branch, that is, the probability density function of the d -dimensional normal distribution obeyed by the data contained in the class j, μ_j is the mean, and $\sum_j$ is the covariance matrix. π_j is the blending ratio, which satisfies: $\pi_j\geq 0,\ \sum_{j=1}^{G}\pi_j$, which describes the ratio of the amount of data contained in the class j to the total amount of data. The parameter space is:

$$\Theta=\left\{\begin{array}{l}\left(\pi_1,\pi_2,\cdots,\pi_G,\theta_1^T,\theta_2^T,\cdots,\theta_G^T\right)^T: \\ \sum_{j=1}^{G}\pi_j=1,\pi_j\geq 0,\ j=1,2,G\end{array}\right\},\psi\in\Theta \quad (4)$$

The parameter θ_j consists of the mean μ_j and the covariance matrix $\sum_j$. π_j is the number of mixed branches (number of components), and each component corresponds to a class of clusters for clustering purposes.

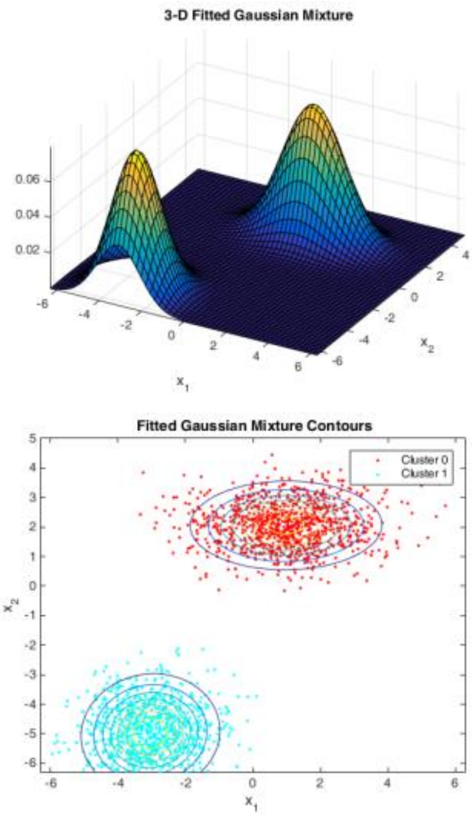

Fig. 2. Mixed Gaussian distribution

The posterior probability that the observed value of the eigenvector x_i belongs to the j component is: $\tau_j(x_i)=\frac{\pi_j f_j(x_i)}{f(x_i)}$, which is the largest component of the posterior density of each mixed data component, and then it is assigned to that component, which can be divided into G Class $C_1,C_2,...,C_G$ that does not overlap. Therefore, the j th cluster contains observation data assigned to the j th group, that is, C_j contains these observation data x_i, and x_i satisfies $\hat{z}_{ij}=(\hat{z}_i)_j=1$. Here:

$$\hat{z}_{ij}=\begin{cases}1 & if\ \hat{\tau}_j(x_j)\geq\hat{\tau}_h(x_j),(h=1,2,\cdots,G;\ h\neq j)\\ 0 & other\end{cases}$$

To complete the clustering of the dataset, it is only necessary to estimate the parameters of each branch of the hybrid model. Here, this paper does not further develop the specific solution process.

2.3. Warning line dynamic adjustment of ρ value

The selection of cordon ρ has a great influence on the accuracy of ART2 classification. In section 1.1, the value of ρ is in the range of $0\sim1$. The closer the value of ρ is to 1, the more sensitive the network is, the finer the classification is, but the network is in a noisy input environment. Memory capacity will quickly run out. Conversely, the closer the value of ρ is to 0, the coarser the classification, and it is easy to classify the different categories of patterns into one category. Therefore, for the researcher to choose the appropriate ρ value, the network will have a certain fault tolerance to the external input distortion mode and noise mode. In order to solve a given ρ value, no matter how the external input environment changes, the classification ability of the network will not change. Therefore, this paper adopts a method of dynamic adjustment of ρ value [17], which adaptively changes the ρ value according to the classification accuracy of the current network, so that the network can obtain relatively optimal classification under constant iteration. ρ is adaptively adjusted as follows:

$$\rho(t+1)=\rho(t)+g\cdot S(x)$$

which

$$S(x)=\begin{cases}1 & x\geq a\\ 0 & b<x<a\\ -1 & x\leq b\end{cases}$$

g is the gain factor, x is the class with the largest number of patterns in all classes after class, and a,b is the class with the largest and smallest number of patterns in all classes after the specified classification ($a\geq b+2$). When $x>a$, it indicates that the classification is too thick, and it is necessary to increase the ρ value appropriately. When $x\leq b$, it indicates that the classification is too fine, and it is necessary to appropriately reduce the ρ value. $\rho(0)$ may be an empirical value.

2.4. GMM-ART2 algorithm flow

From the formula (1) (2) (3), the ART2 neural network uses a "hard competition" classification method, that is, through a "hard competition" way to determine a certain neuron win, and then activate the largest output neuron, and only the activated neurons are tested for similarity, and the final result is uniquely determined. However, the neurons whose output values are ranked behind the winning neurons may be tested for the warning value. In this way, when the pattern features are similarly classified, Causes misclassification, resulting in reduced classification accuracy. In order to solve this problem, based on the existing ART2 neural network, combined with Gaussian hybrid clustering algorithm, the ART2 neural network classification results are appropriately modified. Firstly, the ART2 neural network is used to classify the samples, determine the initial number j, and calculate the cluster center μ_j, the covariance matrix $\sum_j$ and the mixed Gaussian branch weight coefficient π_j of each class, and then according to the Gaussian mixture clustering algorithm. The classification results of the ART2 neural network are corrected, and the main process is shown in the figure.

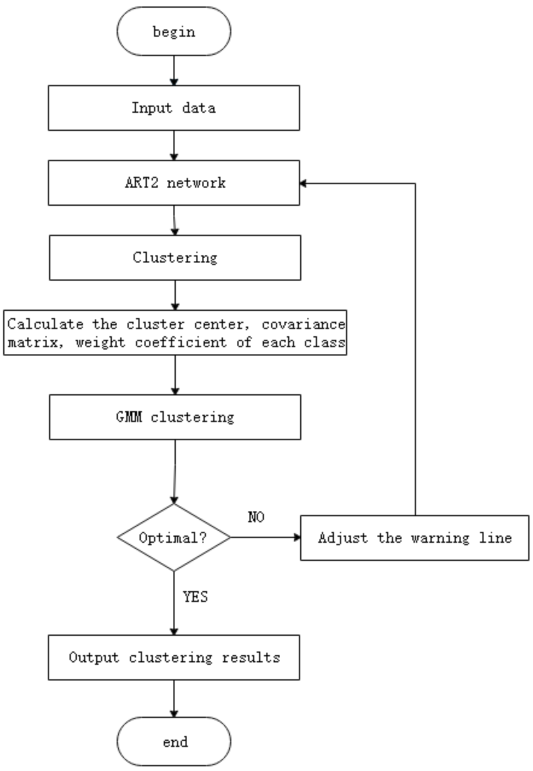

Fig. 3. GMM-ART2 algorithm flow chart

GMM-ART2 specific implementation steps:

(1) Initialize the parameters of the ART2 neural network, input the model to be classified into ART2 for preliminary classification, obtain j classification results, and calculate the cluster center μ_j, covariance matrix $\sum_j$ and mixed Gaussian branch weight coefficient of each class. π_j.

(2) Calculate the $m, \mu_j, \Sigma_j, \pi_j$ result calculated in step (1) as the initialization parameter of the Gaussian mixture model, and calculate the posterior probability $\tau_j(x_i)$ of the observation value of x_i belonging to the j component, and the posterior probability of which mixed component each observed data point belongs to. If the estimate is the largest, assign it to that component and judge whether the classification result is optimal. If it is optimal, the iteration is terminated and the final clustering result is output.

(3) If the classification is not optimal, dynamically adjust the warning line ρ of the ART2 neural network and perform steps (1) and (2).

(4) Repeat the above steps until the classification is optimal, the iteration is terminated, and the final clustering result is output. Optimal classification (no ultra-small super-class)

3. GMM-ART2 network experiment

3.1 Handling network input data

The gas in the transformer oil mainly contains seven kinds of gases $H_2, CH_4, C_2H_6, C_2H_4, C_2H_2, CO, CO_2$. In this paper, the gas sample data is encoded by the three-ratio method (C_2H_2/C_2H_4, CH_4/H_2, C_2H_4/C_2H_6), and 20 sets of network input are obtained. The coding rules of the three-ratio method are shown in Table 1.

Tab.1. Three-ratio method coding rules

Gas ratio range	Ratio range coding		
	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
<0.1	0	1	0
$\geq 0.1 and < 1$	1	0	0
$\geq 1 and < 3$	1	2	1
≥ 3	2	2	2

Diagnose the faults of five types of transformers: low temperature overheating(LTO), medium temperature overheating(MTO), high temperature overheating(HTO), low energy discharge(LED) and high energy discharge(HED). The identification and coding combination of faults are shown in Tab. 2.

Tab.2. Transformer fault type

Fault type	LTO	MTO	HTO	LED	HED

symbol	F_1	F_2	F_3	F_4	F_5
Coding combination	0	0	0	1	2
	2	2	(0,1,2)	(0,1)	(0,1)
	0	1	2	(0,1,2)	(0,1,2)

(1) Set the initial parameters of the GMM-ART2 network $\rho = 0.995, a = 10, b = 10, c = 0.1,$ $d = 0.90, \mathrm{e} = 0.000001, \theta = 0, \mathrm{g} = 0.001$

(2) Enter 20 sets of fault data into the GMM-ART2 network, and get $\rho = 0.990$ after classification. The specific classification results of the network are shown in Table 3.

3.2 Experimental verification

Tab.3. GMM-ART2 network transformer fault classification results

Actual ratio			Coded input			STM-F1 vector P			Classification result	Actual fault	Result
0.0247	4.3429	0.5845	0	2	0	0	10.000	0	1	F_1	YES
0.0000	1.6727	0.2597	0	2	0	0	10.000	0	1	F_1	YES
0.0000	4.6000	0.7209	0	2	0	0	10.000	0	1	F_1	YES
0.0000	1.4475	0.6829	0	2	0	0	10.000	0	1	F_1	YES
0.0051	5.6345	1.1800	0	2	1	0	0.8944	0.4472	2	F_2	YES
0.0032	3.3333	1.5000	0	2	1	0	8.9443	4.4721	2	F_2	YES
0.0000	2.2838	2.2669	0	2	1	0	8.9443	4.4721	2	F_2	YES
0.0714	1.4475	2.5143	0	2	1	0	8.9443	4.4721	2	F_2	YES
0.0075	5.1071	9.6667	0	2	2	0	7.071151	7.071151	3	F_3	YES
0.0000	0.8534	4.8444	0	0	2	0	0	1.0000	4	F_3	NO
0.0079	2.8600	3.3591	0	2	2	0	7.071151	7.071151	3	F_3	YES
0.0395	1.6556	15.000	0	2	2	0	7.071151	7.071151	3	F_3	YES
0.9474	0.3333	19.000	1	0	2	0.4472	0	0.894451	4	F_4	YES
0.1517	0.1427	9.3684	1	0	2	4.4721	0	8.9443	4	F_4	YES
0.8914	0.1253	1.1218	1	0	1	0.707194	0	0.0711	5	F_4	NO
1.2308	0.2485	13.000	1	0	2	4.4721	0	8.9443	4	F_4	YES
4.3931	0.2085	20.714	2	0	2	7.071194	0	7.0711	5	F_5	YES
4.6918	0.4494	112.3077	2	0	2	7.071194	0	7.0711	5	F_5	YES
3.5211	0.2927	3.9444	2	0	2	7.071194	0	7.0711	5	F_5	YES
22.456	0.2124	32.568	2	0	2	7.071194	0	7.0711	5	F_5	YES

It can be seen from Table 3 that two of the 20 sets of data are incorrectly classified, and the correct classification rate of the GMM-ART2 model proposed in this paper reaches 90%.

4. Conclusion

In this paper, the fault diagnosis method of GMM-ART2 is proposed for the fault diagnosis of transformers. The method has adaptive characteristics in fault diagnosis. The simulation results based on transformer data show that the proposed method has good feasibility and comparison high accuracy. The research results of transformer faults have certain reference significance and engineering application value.

References

Chen W, Song Y, Sun J, 2000 Concept and present situation of condition monitoring of power system equipment,J. Power System Technology.

Yan L, Wang W J, Zhou D H, 2003.A Survey on Oil Chromatogram Analysis for Fault Diagnosis of Transformers,J. Control Engineering of China.

Tang W H,Wu Q H,2011.Condition monitoring and assessment of power transformers using computational intelligence, M. New York: Springer-Verlag Press.

W. Du, Zhao X., Du H., and Lv F., Fault Diagnosis of Analog Circuit based on Extreme Learning Machine. International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 9-14.

H. Liu, Lv X., and Xiao J., The Research Status and Prospect of Particle Swarm Optimization Algorithm for Water Environment Quality Assessment.

International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 23-30.

Su P, Wang H, 2003.Technology analysis of state monitoring and fault diagnosis in electic power systems,J. Automation of Electric Power Systems.

Wagh N, Deshpande D, 2011.Transformer Incipient Fault Diagnosis Using Artificial Neural Network,M// Computational Intelligence and Information Technology. Springer Berlin Heidelberg.

Ling R, Xie Q, Gao S, et al, 2014.Application of Artificial Neural Network and Information Fusion Technology in Power Transformer Condition Assessment,J. High Voltage Engineering.

Zhao W Q, Zhu Y L, Zhang X Q,2008,.Combinational Forecast for Transformer Faults Based on Support Vector Machine,J. Proceedings of the Csee, 28(25):14-19.

Shi X, Zhu Y, 2015.Application of Deep Learning Neural Network in Fault Diagnosis of Power Transformer,J. Electric Power Construction.

Li F,Zhan J1994:.Fuzzy adapting vigilance parameter of ART-II neural nets,C//IEEE International Conference on Neural Networks. New York: IEEE,1680−1685.

Lv X, Yao G, Zhao Y, et al 2006.An improved architecture based on typical ART-2 neural network,C// International Symposium on Systems and Control in Aerospace and Astronautics. IEEE,3 pp.-1121.

Zhang H, Guan W, Guan G,2008.On-Line Diagnosis of Faulty Insulators Based on Improved ART2 Neural Network,M// Advances in Neural Networks - ISNN 2008. Springer Berlin Heidelberg, 465-472.

Luukka P，2007. Similarity classifier using similarity mea-sure derived from Yu’s norms in classification of medical data sets,J.Computers in Biology and Medicine，37(8):1133-1140.

Hailin M U, Wang Z, Zengbing X U, 2017.Fault Diagnosis Method of Bearings by Yu's Norm ART Based on Soft Competition,J. China Mechanical Engineering,.

Zhuang L,Yibing L , Zhiyong M , Wei T 2015.Application of ART2 Neural Network Combined with C-means Clustering in Fault Diagnosis of Wind Turbine Gearbox,J.Journal Of Chinese Society Of Power Engineering，35(8):647-665.

Gail A. Carpenter,Stephen Grossberg,1987.ART 2: self-organization of stable category recognition codes for analog input patterns,J. Applied Optics,26(23):4919-4930.

Sun G L, Tang X L, 2004. A Semi-Supervised Learning Algorithm Based on a Hierarchical GMM,J. Journal of Computer Research & Development, 41(1):156-161.

Zhang M, Dong G, 2000.The Improvement on ART-2 Neural Network Algorithm for Pattern Recognition,J. Control Theory & Applications, 17(1).

X. Luo, Ge S., Wang J., and Guan, Time Delay Estimation-based Adaptive Sliding-Mode Control for Nonholonomic Mobile Robots. International Journal of Applied Mathematics in Control Engineering, 2018. 1(1): p. 1-8.

Zheng Guping (1960-), male, Baoding, Hebei, Ph.D., professor. The main research direction is smart grid and big data (E-mail: zhengguping@126.com)

Ma Biao (1993-), male, Bozhou, Anhui, master student, the main research field is electric power information technology (E-mail: mabiao1993@126.com);

Li Gang (1980-), male, Hebei Zaoqiang, Ph.D., associate professor. Mainly engaged in research on smart grid and big data, information physical energy systems, fault prediction and health management (E-mail: ququ_er2003@126.com);