Fault Diagnosis Based on Fuzzy C-means Algorithm of the Optimal Number of Clusters and Probabilistic Neural Network

Qing Yang[1,2*]**, Jingran Guo**[3]**, Dongxu Zhang**[2]**, Chang Liu**[2]

[1] *College of Optical and Electronical Engineering, Changchun University of Science and Technology, Changchun 130022, China*
[2] *School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110159, China*
[3] *School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China*

* *Corresponding author's Email: yangqingxp@126.com*

Abstract: Fault diagnosis is essential for the reliable, safe, and efficient operation of the plant and for maintaining quality of the products in industrial system. This paper presents an ensemble fault diagnosis algorithm based on fuzzy c-means algorithm (FCM) with the Optimal Number of Clusters (ONC) and probabilistic neural network (PNN), called FCM-ONC-PNN. In clustering methods, the estimation of the optimal number of clusters is significant for subsequent analysis. As a simple clustering method, FCM has been widely discussed and applied in pattern recognition and machine learning, but FCM could not guarantee unique clustering result because initial cluster number is chosen randomly. As the number of clusters is randomly chosen, the iterative amount is large and the result of the classification is unstable. In this paper, firstly subtractive clustering is proposed to find the optimal number of clusters and the clustering results of the FCM are compared with random initialization method, and then PNN is used to classify the clustering data of FCM. The experiments show that the modified initial cluster number of FCM algorithm can improve the speed, and reduce the iterative amount. At the same time, FCM-ONC-PNN approach can make classification more stable and have higher precision.

Keywords: Fuzzy c-means algorithm; Initial cluster number; Subtractive clustering; Probabilistic neural network; FCM-ONC-PNN

1. Introduction

With an increasing requirement for dynamic systems to be more secure and more reliable, fault detection and diagnosis (FDD) in a control system is becoming more and more critical and important. The main task of fault diagnosis is to detect and isolate the occurring fault in order to avoid overall failure of the monitored system and any catastrophes involving human fatalities or material damage.

There are many methods for fault diagnosis [1-5], which can be grouped into the following two categories: 1) Model-based FDD; 2) Data-driven FDD including knowledge-based FDD [6]. In the early days, model-based FDD constituted the main stream of research. However, mode-based FDD uses mathematical system models to estimate the system state and parameters, and in general these methods can only be applied to low dimensional systems. Alternatively, data-driven FDD can deal with high dimensional data, and data dimension reduction techniques [7-9] are generally used to highlight important information in data sets.

Fault diagnosis involves two stages [10]: Preprocessing and classification. The preprocessing includes feature extraction and selection. The classification stage maps a pattern from the feature space to a decision.

This mapping is done by a classifier which generates a class membership function in order to classify unlabelled incoming patterns into one of the predefined classes. Depending on the information available for classifier training, one can distinguish between supervised learning [11-14], such as PNN and SVM, and unsupervised learning [15-17], also called clustering.

Although single fault diagnosis approach can be used for fault classification, but each fault diagnosis technique has its disadvantages, and its effect of fault classification is limited. Therefore, ensemble methods incorporating several techniques for fault detection and diagnosis seem attractive.

Some researchers deal with an ensemble method of testing analog VLSI circuits [18], using wavelet transform for analog circuit response analysis and arti?cial neural networks (ANN) for fault diagnosis. Recognition system for totally unconstrained handwritten characters for south Indian language of Kannada is proposed [19]. The proposed feature extraction technique is based on Fourier Transform and well known Principal Component Analysis (PCA). The system trains the appropriate frequency band images followed by PCA feature extraction scheme. For subsequent classification technique, Probabilistic Neural Network (PNN) is used.

Some articles add some clustering algorithms before PNN classification [1, 20]. Seok-Beom Roh [21] introduces a new category of fuzzy models called a fuzzy ensemble of parallel polynomial neural network, which consists of a series of polynomial neural networks weighted by activation levels of information granules formed with the use of fuzzy clustering. The two underlying design mechanisms of the proposed networks rely on information granules resulted from the use of fuzzy C-means clustering (FCM) and take advantage of polynomial neural networks (PNN). They consider polynomial neural networks, which exhibit highly nonlinear characteristics when being viewed as local learning models. They use FCM to form information granules and to overcome the high dimensionality problem. They adopt PNN to find the optimal local models, which can describe the relationship between the input variables and output variables within some local region of the input space.

A landslide susceptibility analysis is performed by means of Artificial Neural Network (ANN) and Cluster Analysis (CA) [22]. This kind of analysis is aimed at using ANN to model the complex non linear relationships between mass movements and conditioning factors for susceptibility, in order to identify unstable areas. CA is adopted to improve the selection of training, validation, and test records from data.

There are two types of evaluation index to estimate the optimal number of clusters: one is an external measure based on a priori or empirically obtained biological information, and the other is an internal measure based on statistical indices. Different internal measures consider the compactness of each cluster, the separation of clusters, or a combination of both. In fuzzy k-means clustering, indices, such as the partition coefficient (PC) and fuzzy hyper volume (FHV) consider compactness, and indices such as the Xie-Beni index (XB) and the PBM index consider both compactness and separation. These techniques [23-29] could estimate the number of clusters in several benchmarks; however, they are not universally applicable for many practical situations.

To solve the optimal number of clusters problem, this paper adds the subtractive clustering [30, 31] to determine the number of clusters, and presents an ensemble approach based on FCM of the optimal number of clusters and PNN.

This paper is organized as follows. The second section briefly introduces fault diagnosis strategy based on FCM-ONC-PNN. To validate the performance and effectiveness of the proposed scheme, the application of FCM-ONC-PNN approach to simulation benchmark of TE process is illustrated in Section 3. Finally, concluding remarks are made in Section 4.

2. FCM-ONC-PNN Algorithm

2.1 FCM

Fuzzy C-means (FCM) is one of the most popular unsupervised fuzzy clustering algorithms, which has been widely used for image segmentation. The general fuzzy C-means clustering algorithm was proposed by Bezdek [32, 33].

The FCM algorithm assigns pixels to each category by using fuzzy memberships. Let $X = \{x_i, i = 1, 2, \ldots, N | x_i \in R^d\}$ denote an image with N pixels to be partitioned into c classes (clusters), where x_i represents feature data. The algorithm is an iterative optimization that minimizes the objective function defined as follows:

$$J_m = \sum_{k=1}^{c} \sum_{i=1}^{N} u_{ki}^m \parallel x_i - v_k \parallel^2 \tag{1}$$

with the following constraints:

$$\sum_{k=1}^{c} u_{ki} = 1, \qquad \forall i;$$

$$0 \le u_{ki} \le 1, \qquad \forall k,i; \qquad (2)$$

$$\sum_{i=1}^{N} u_{ki} > 0, \qquad \forall k$$

where u_{ki} represents the membership of pixel x_i in the kth cluster, υ_k is the kth class center, $\| \cdot \|$ denotes the Euclidean distance, $m > 1$ is a weighting exponent on each fuzzy membership. The parameter m controls the fuzziness of the resulting partition. The membership functions U and cluster centers V are updated by the following expressions:

$$u_{ki} = \frac{1}{\sum_{l=1}^{c} \left(\frac{\|x_i - \upsilon_k\|}{\|x_i - \upsilon_l\|} \right)^{2/(m-1)}} \qquad (3)$$

and

$$\upsilon_k = \frac{\sum_{i=1}^{N} u_{ki}^m x_i}{\sum_{i=1}^{N} u_{ki}^m} \qquad (4)$$

In implementation, matrix υ is usually randomly initialized, and then U and V are updated through an iterative process using (3) and (4), respectively.

2.2 Subtractive Clustering

Part of the problems faced by the FCM is determining the number of clusters needed prior to learning. This is usually inputted by the user through a series of trial and error values. Also the usage of random initialization does not provide deterministic results. Subtractive clustering can overcome these problems.

Subtractive clustering [30, 31] is, essentially, a modified form of the Mountain Method. Thus, let Z be the set of data points obtained by concatenation of N value of the $X(k-1)$ and t, where $t = \lfloor t_1, \ldots t_{n_y} \rfloor$ is the vector of output target values, corresponding to the $\widehat{y}(k)$. In the algorithm, each point is seen as a potential cluster center, for which some potential measures are assigned according to (5):

$$p_i = \sum_{j=1}^{N} e^{-\alpha[z_i - z_j]^2} \qquad (5)$$

where $\alpha = 4/r_a^2$ and $r_a > 0$ define the neighborhood radius for each cluster center. Thus, the potential associated with each cluster depends on its distance to all the points, leading to clusters with high potential where neighborhood is dense. After calculating potential for each point, the one with higher potential is

selected as the first cluster center. Let z_1^* be the center of the first group and p_1^* its potential. Then the potential p_i for each z_i is reduced according to Eq.(6), especially for the points closer to the center of the cluster:

$$p_i = p_i - p_1^* e^{-\beta[z_l - z_1^*]^2} \qquad (6)$$

where $\beta = 4/r_b^2$ and $r_b > 0$ represent the radius of the neighborhood for which significant potential reduction will occur. The radius for reduction of potential should be to some extent higher than the neighborhood radius to avoid closely spaced clusters. Typically, $r_b = 1.25 r_a$. Since the points closer to the cluster center will have their potential strongly reduced, the probability for those points to be chosen as the next cluster is lower. This procedure (selecting centers and reducing potential) is carried out iteratively until stopping criteria is satisfied. Additionally two threshold levels are defined, one above which a point is selected for a cluster center and the other below which a point is rejected.

2.3 Probabilistic Neural Network

The probabilistic neural network (PNN) used in this study, shown in Fig.1, mainly includes a radial basis layer and a competitive layer. The radial basis layer contains the same number of neurons as that of the train set. Each neuron is responsible to calculate the probability that an input feature vector is associated with a specific class.

The radial basis layer biases are all set as

$$b = \frac{[-\log(0.5)]^{1/2}}{spread} \qquad (7)$$

where $spread$ is extended coefficient of RBF.

With an input victor X, the radial basis neuron compares it with the neuron weight W_{ji} and multiplies with a bias b to calculate the probability,

$$O_j = radbas(v_j) \qquad (8)$$

where $radbas$ can be selected from any of several candidate radial functions. In this paper, this radial function is selected as a Gaussian function.

$$O_j = e^{-v_j^2} \qquad (9)$$

$$v_j = \| W_{ji} - X \| \cdot b \qquad (10)$$

where $\| \cdot \|$ denotes the Euclidean distance.

Consequently, as the distance between W_{ji} and X decreases, the output O_j increases and reaches the maximum of 1 when $W_{ji} = X$. The sensitivity of the radial
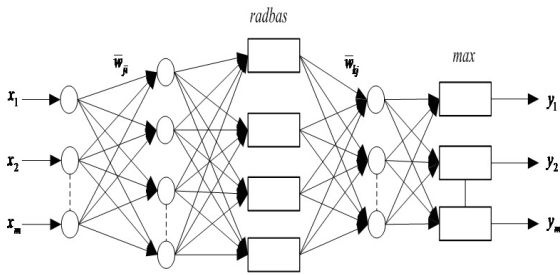
Figure 1 The sketch of PNN

basis neurons can be adjusted by varying the value of b through the extended coefficient: *spread*. The competitive lay then determines the maximum in the probabilities and assigns 1 to the associated class, 0 to the others.

2.4 FCM-ONC-PNN

This algorithm is proposed to improve the diagnose accuracy and speed. At the same time, it is a combination of algorithm about FCM and PNN. Fig.2 shows the sketch of the algorithm, which is briefly described as follows:

Step 1: Get test data. For example, the data in this paper come from Tennessee Eastman process.

Step 2: Chose the initial cluster number randomly or determine it by subtracting .

Step 3: Cluster the data sets using FCM algorithm.

Step 4: Compare the clustering results of FCM-ONC with random initialization method, and then evaluate clustering effectiveness.

Step 5: Diagnose faults to the clustering data of FCM-ONC by PNN.

3. Case Study

To validate the performance and effectiveness of the method, FCN-ONC-PNN algorithm is applied to TE Process (as showed in Fig.3) [34]. The test data of fault 1fault 2 and fault 8 are seen as Example data set (as showed in Fig.4, Fig.5 and Fig.6).

TE process is a benchmark problem in process engineering. Downs and Vogel presented this particular process at an AICHE meeting in 1990 as a plant-wide control problem. The simulator of the Tennessee Eastman process consists of five major unit operations: a reactor, a product condenser, a vapor-liquid separator, a recycle compressor, and a product stripper. Two products are produced by two simultaneous gas-liquid exothermic reactions, and a byproduct is generated by two additional exothermic reactions. The process has
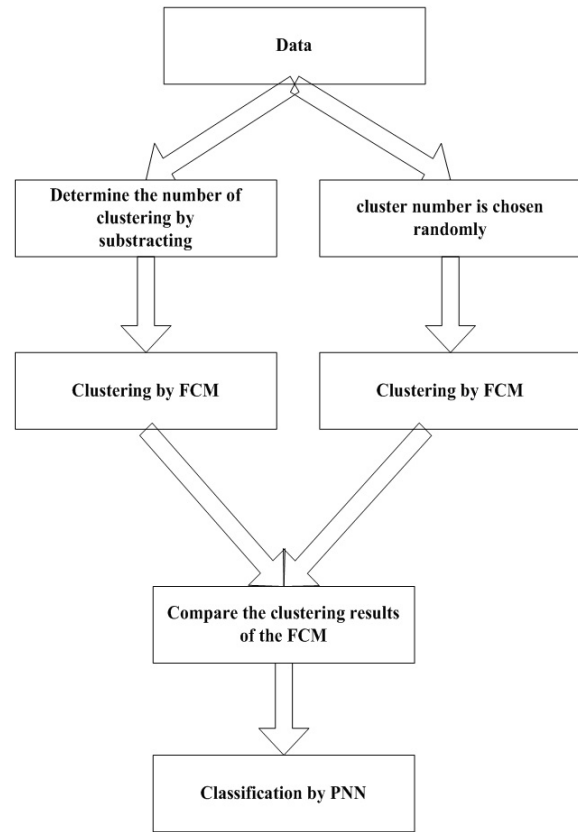

Figure 2 The sketch of the process

12 manipulated variables, 22 continuous process measurements, and 19 compositions. The simulator can generate 21 types of different faults, listed in Table 1. Once the fault enters the process, it affects almost all state variables in the process.

3.1 Evaluation method

To evaluate the efficiency of clustering algorithms, *PC* and *PE* have been employed in the following experiments. Bezdek has defined a performance measure based on minimizing the overall content of pair wise fuzzy intersection in U, the partition matrix. He has also proposed cluster validity index for fuzzy clustering [30]: partition coefficient (*PC*). The index is defined as

$$PC = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^2 \qquad (11)$$

where the *PC* index indicates the average relative amount of membership sharing done between pairs of fuzzy subsets in U, by combining into a single number, the average contents of pairs of fuzzy algebraic products. The index values range in $1/C$, where C is the number of clusters.

The more the *PC* is close to 1, the more distinct the result is; on the contrary the more the *PC* is close to $1/C$, the vaguer the result is.

Bezdek has proposed the partition entropy (*PE*) [33], which is defined as

$$PE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij} \cdot \log_a(u_{ij}) \qquad (12)$$

where *a* is the basis of the algorithm. The *PE* index is a scalar measure of the amount of fuzziness in a given *U*. The index is computed for values of *C*, greater than 1, and its values vary between $[0, \log_a C]$.
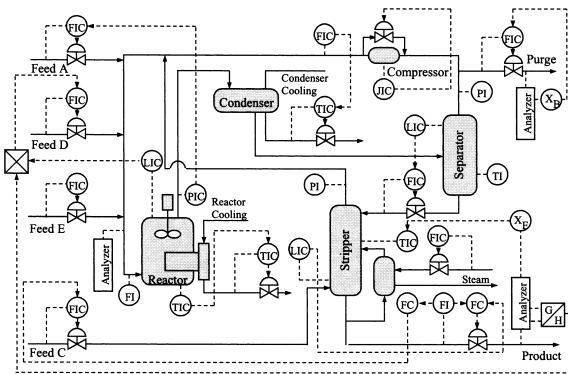
Figure 3 Control system of the Tennessee Eastman process
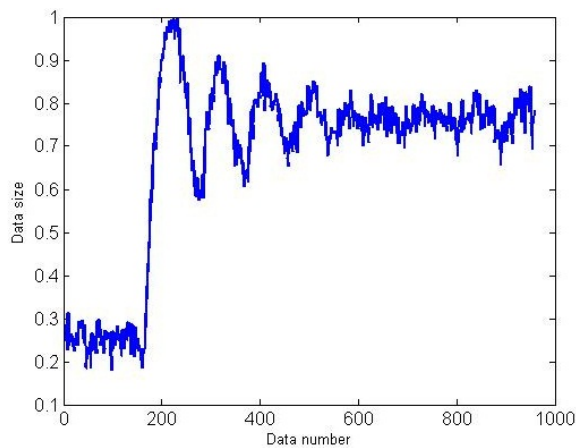
Figure 4 Example data set of fault 1

The lesser *PE*, the more distinct clustering result of FCM On the contrary, the closer to the $\log_a C$ for *PE*, the vaguer clustering result of FCM.

3.2   Evaluation on clustering effectiveness

In order to test the effectiveness of the FCM-ONC method, the data set of TE Process is used.

Supposing the random number of clusters is 3. Fig.7, Fig.8 and Fig.9 show the results of the FCM algorithm

Table 1 Process faults for the Tennessee Eastman process

| Variable	Disturbances	Tpye
1	A/C feed ratio, B composition constant	Step
2	B composition, A/C ration constant	Step
3	D feed temperature	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A feed loss	Step
7	C header pressure loss-reduced availability	Step
8	A, B, C feed	Random variation
9	D feed temperature	Random variation
10	C feed temperature	Random variation
11	Reactor cooling water inlet temperature	Random variation
12	Condenser cooling water inlet temperature	Random variation
13	Reaction kinetics	Slow drift
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16-20	Unknown	Unknown
21	The valve for Stream 4 was fixed at the steady state position	Step constant position
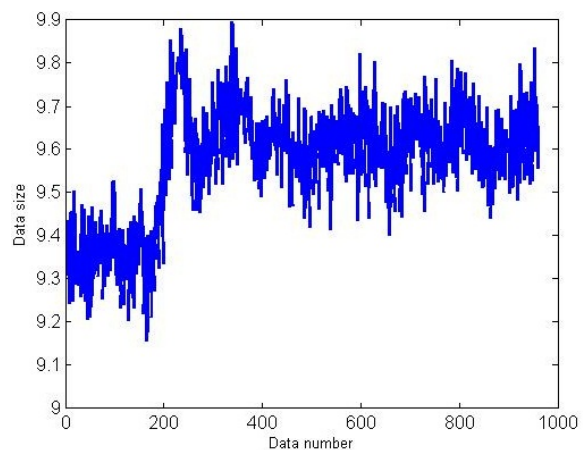
Figure 5 Example data set of fault 2
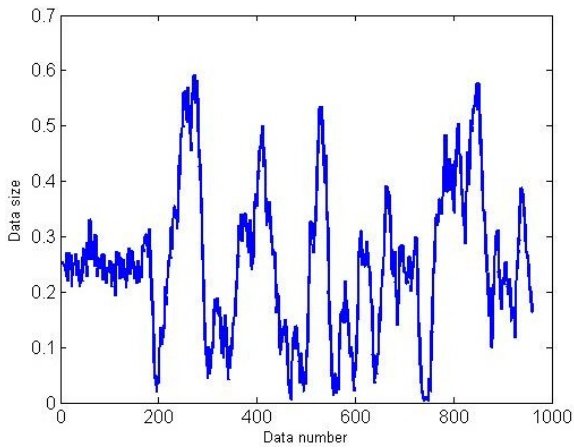
Figure 6 Example data set of fault 8

Figure 8 Clustering result of FCM on fault 2 data set ($C = 3$)

with the initial $C = 3$. From Fig.7, Fig.8 and Fig.9 we find that the random number of clusters algorithm can not accurately cluster the data set.

However, by subtracting clustering we know that the initial number of clusters is 2, Fig.10, Fig.11 and Fig.12 show the clustering result of the FCM algorithm with the ONC initialization method on the example data ($C = 2$). From Fig.10, Fig.11 and Fig.12, we find that the FCM-ONC algorithm can accurately discover the initial cluster number.

Furthermore, Table 2 illustrates the influence of the two evaluation methods on clustering and the comparison results of two initialization methods on the example data.

Figure 9 Clustering result of FCM on fault 8 data set ($C = 3$)

Figure 7 Clustering result of FCM on fault 1 data set ($C = 3$)

3.3 Faults diagnosis of the clustering data of FCM through the PNN

PNN is a feed-forward neural network with supervised learning which uses Bayes decision rule and Parzen
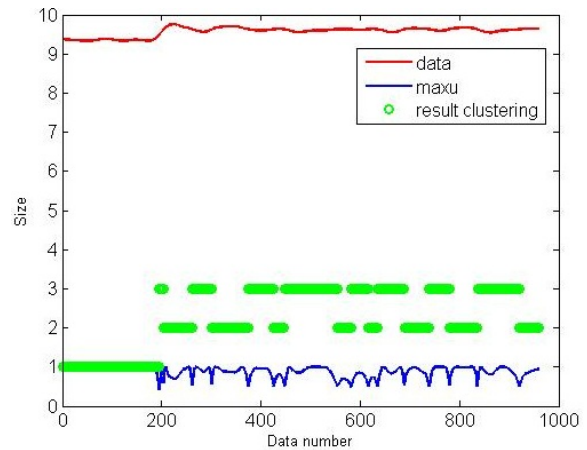

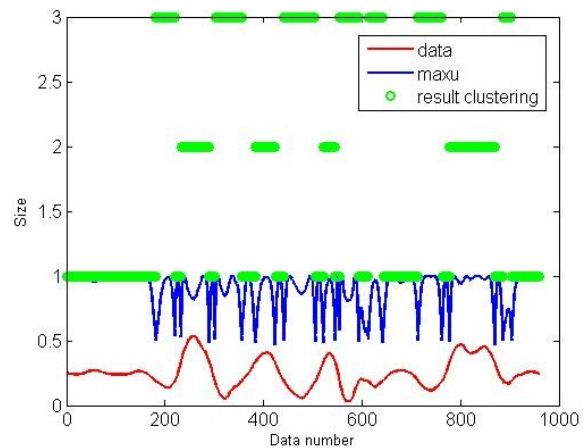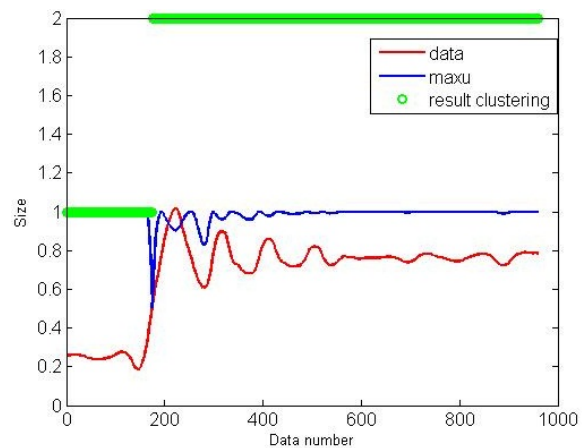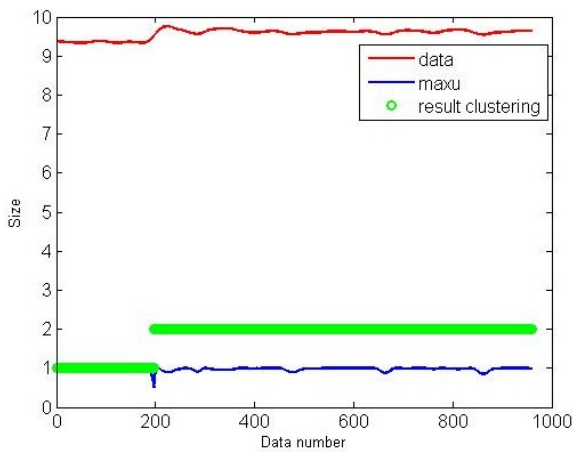Figure 10 Clustering result of FCM-ONC on fault 1 data set ($C = 2$)

Figure 11 Clustering result of FCM-ONC on fault 2 data set ($C = 2$)
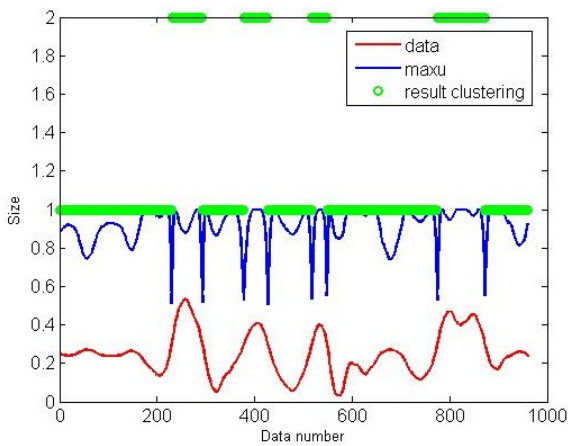

Figure 12 Clustering result of FCM-ONC on fault 8 data set ($C = 2$)

Table 2 The Comparison with different initialization methods

TE Fault	Evaluation method	The proposed method ($C = 2$)	Random method ($C = 3$)
01	PC	0.97184	0.89076
01	PE	0.082754	0.2935
02	PC	0.95784	0.82378
02	PE	0.12547	0.43962
08	PC	0.85747	0.83544
08	PE	0.35521	0.43653

window. In order to deal with the effectiveness and precision of fault diagnosis, the PNN is used. The initial data sets come from result of clustering through FCM (as showed in Fig.13 and Fig.14) show the initial data sets about fault 1. It is clustered two classes by FCM. Fig.13 ($C = 1$) is a kind of data. Fig.14 ($C = 2$) is another kind of data. Compared fault 1 ($C = 1$ and $C = 2$) with the normal data by PNN, Fig.15 shows the classifying result of the two kinds of data set ($C = 1$ and $C = 2$). The first 160 data are normal data, the last 800 data are fault data. The experimental results show that the FCM-ONC-PNN method outperforms the other initialization methods. At the same time, the experiments show that the modified initial cluster number of FCM algorithm can improve the speed (as showed in Fig.16 and Fig.17), and reduce the iterative amount.
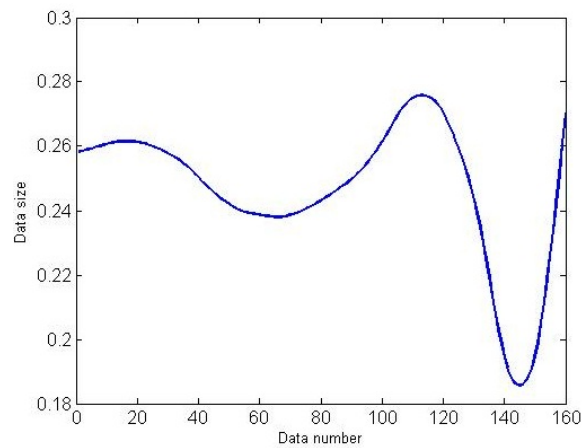

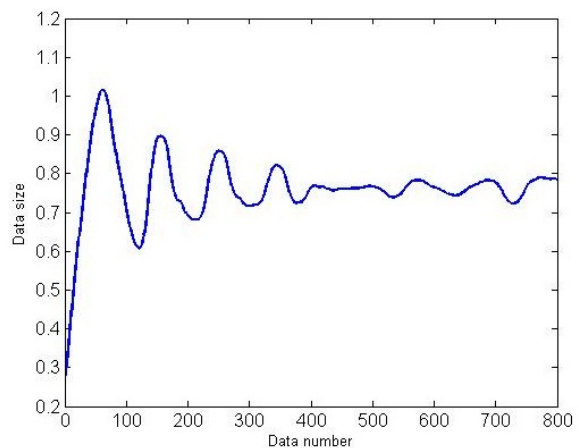Figure 13 Fault 1 data set by FCM-ONC clustering ($C = 1$)

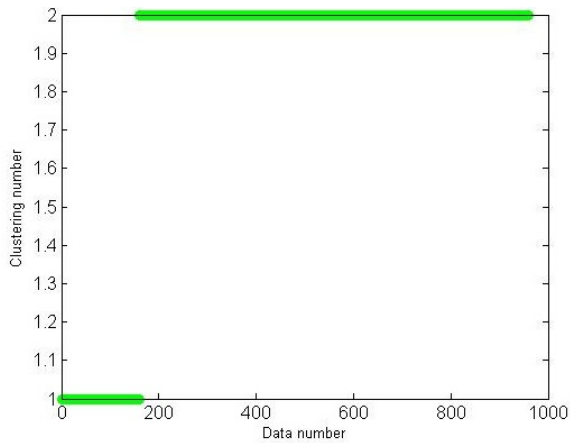Figure 14 Fault 1 data set by FCM-ONC clustering ($C = 2$)

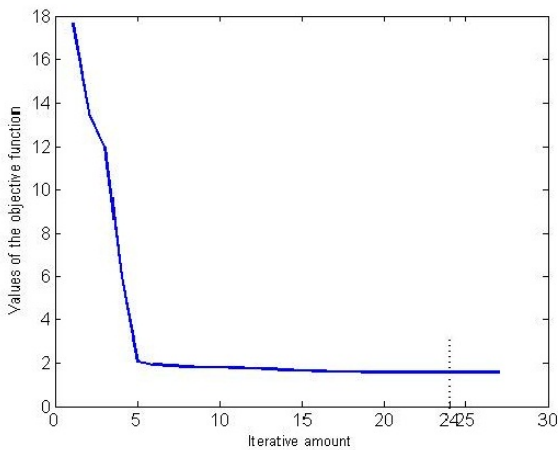Figure 15 Classifying result of the fault 1 data by FCM-ONC-PNN

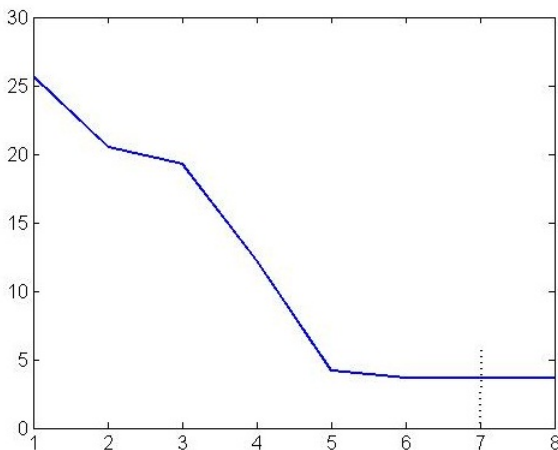Figure 16 Iterative amount of FCM on fault 1 data

Figure 17 Iterative amount of FCM-ONC on fault 1 data

4. Conclusion

An ensemble fault diagnosis method based on FCM with optimal number of clusters and PNN, called FCM-ONC-PNN, is presented in this paper. Simulation studies show that the proposed algorithm not only provides an accepted degree of accuracy in fault classification under different fault conditions, but it is also a reliable, fast and computationally efficient tool.

5. Acknowledgments

References

[1] R. Eslamloueyan, "Designing a hierarchical neural network based on fuzzy clustering for fault diagnosis of the Tennessee-Eastman process", *Applied Soft Computing*, Vol.11, pp.1407-1415, 2011.

[2] Q. Wu and Z.H. Ni, "Car assembly line fault diagnosis based on triangular fuzzy support vector classifier machine and particle swarm optimization", *Expert Systems with Applications*, Vol.38, pp.4727-4733, 2011.

[3] S.C. Du and L.F. Xi, "Fault diagnosis in assembly processes based on engineering-driven rules and PSOSAEN algorithm", *Computers & Industrial Engineering*, Vol.60, pp. 77-88, 2011.

[4] S. Ding, P. Zhang, E. Ding, S. Yin, A Naik, P. Deng, and W. Gui, "On the application of PCA technique to fault diagnosis", *Tsinghua Science & Technology*, Vol.15, pp.138-144, 2010.

[5] Y. Gao, X.S. Wang, Y.H. Cheng, and J. Pan, "Fault diagnosis using a probability least squares support vector classification machine", *Mining Science and Technology*, Vol.20, pp.917-921, 2010.

[6] H. Wang, T.Y. Chai, J.L. Ding, and M. Brown, "Data driven fault diagnosis and fault tolerant control: some advances and possible new directions", *Acta Automatica Sinica*, Vol.35, pp.739-747, 2009.

[7] K. Bunte, B. Hammer, A. Wismller, and M. Biehl, "Adaptive local dissimilarity measures for discriminative dimension reduction of labeled data", *Neurocomputing*, Vol.73, pp.1074-1092, 2010.

[8] E. Gutirrez, S. Lozano, M.T. Moreira, and G. Feijoo, "Assessing relationships among life-cycle environmental impacts with dimension reduction techniques", *Journal of Environmental Management*, Vol.91, pp.1002-1011, 2010.

[9] P. Poncela, J. Rodrłguez, R. Snchez-Mangas, and E. Senra, "Forecast combination through dimension reduction techniques", *International Journal of Forecasting*, Vol.27, pp.224-237, 2011.

[10] M.S. Mouchaweh, "Semi-supervised classification method for dynamic applications", *Fuzzy Sets and Systems*, Vol.161, pp.544-563, 2010.

[11] H.Q. Wang and P. Chen, "Intelligent diagnosis method for rolling element bearing faults using possibility theory and neural network", *Computers & Industrial Engineering*, Vol.60, pp. 511-518, 2011.

[12] S.R. Samantaray, P.K. Dash, and S.K. Upadhyay, "Adaptive Kalman filter and neural network based high impedance fault detection in power distribution networks", *Electrical Power & Energy Systems*, Vol.31, pp.167-172, 2009.

[13] V. Sugumaran and K.I. Ramachandran, "Effect of number of features on classification of roller bearing faults using SVM and PSVM", *Expert Systems with Applications*, Vol.38, pp.4088-4096, 2011.

[14] L.M.R. Baccarini, V.V.R. e Silva, and B.R. de Menezes, and W.M. Caminhas, "SVM practical industrial application for mechanical faults diagnostic", *Expert Systems with Applications*, Vol.38, pp. 6980-6984, 2011.

[15] C.T. Yiakopoulos, K.C. Gryllias, and I.A. Antoniadis, "Rolling element bearing fault detection in industrial environments based on a K-means clustering approach", *Expert Systems with Applications*, Vol.38, pp.2888-2911, 2011.

[16] B. Chen and C.Z. Zang, "A hybrid immune model for unsupervised structural damage pattern recognition", *Expert Systems with Applications*, Vol.38, pp.1650-1658, 2011.

[17] F. Corona, M. Mulas, R. Baratti, and J. A. Romagnoli, "On the topological modeling and analysis of industrial process data using the SOM", *Computers & Chemical Engineering*, Vol.34, pp.2022-2032, 2010.

[18] P. Kalpana and K. Gunavathi, "Wavelet based fault detection in analog VLSI circuits using neural networks", *Applied Soft Computing*, Vol.8, pp.1592-1598, 2008.

[19] V.N. Manjunath Aradhy, S.K. Niranjan, and G. Hemantha Kumar, "Probabilistic neural network based approach for handwritten character recognition", *Special Issue of IJCCT*, Vol.1, pp.9-13, 2010.

[20] S.-K. Oh, W. Pedrycz, and S.-B. Roh , "Hybrid fuzzy set-based polynomial neural networks and their development with the aid of genetic optimization and information granulation", *Applied Soft Computing*, Vol.9, pp.1068-1089, 2009.

[21] S.-B. Roh, S.-K. Oh, and W. Pedrycz, "A fuzzy ensemble of parallel polynomial neural networks with information granules formed by fuzzy clustering", *Knowledge-Based Systems*, Vol.23, pp.202-219, 2010.

[22] C. Melchiorre and M. Matteucci, A. Azzoni, and A. Zanchi, "Artificial neural networks and cluster analysis in landslide susceptibility zonation", *Geomorphology*, Vol.94, pp.379-400, 2008.

[23] M.C. Naldi, R.J.G.B. Campello, E.R. Hruschka, and A.C.P.L.F. Carvalho, "Efficiency issues of evolutionary k-means", *Applied Soft Computing*, Vol.11, pp.1938-1952, 2011.

[24] M.K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters", *Pattern Recognition*, Vol.37, pp.487-501, 2004.

[25] M.K. Pakhira, S. Bandyopadhyay, and U. Maulik, "A study of some fuzzy cluster validity indices, genetic clustering and application to pixel classification", *Fuzzy Sets and Systems*, Vol.155, pp.191-214, 2005.

[26] X.L. Xie and G. Beni, "A validity measure for fuzzy clustering", *IEEE Trans. Pattern Anal. Mach. Intell*, Vol.13, pp.841-847, 1991.

[27] M. Eftekhari and S.D. Katebi, "Extracting compact fuzzy rules for nonlinear system modeling using subtractive clustering, GA and unscented filter", *Applied Mathematical Modelling*, Vol.32, pp.2634-2651, 2008.

[28] I. Gath and A.B. Geva, "Unsupervised optimal fuzzy clustering", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.11, pp.773-781, 1989.

[29] R. K. Y. Chang, C. K. Loo, and M.V.C. Rao, "A global k-means approach for autonomous cluster initialization of probabilistic neural network ", *Informatica*, Vol.32, pp.219-225, 2008.

[30] S.L. Chiu, "Fuzzy model identification based on cluster estimation", *Journal of Intelligent and Fuzzy Systems*, Vol. 2, pp.267-278, 1994.

[31] M. Eftekhari, S.D. Katebi, M. Karimi, and A.H. Jahanmiri, "Eliciting transparent fuzzy model using differential evolution", *Applied Soft Computing*, Vol. 8, pp.466-476, 2008.

[32] J.C. Bezdek, R.Ehrlich, and W.Full, "FCM: The fuzzy c-means clustering algorithm", *Computers & Geosciences*, Vol.10, pp.191-203, 1984.

[33] J.C. Bezdek, "Cluster validity with fuzzy sets", *J. Cybernet.*, Vol.3, pp.58-73, 1974.

[34] T.J. Mcavoy and N. Ye, "Base control for the Tennessee Eastman problem", *Computers & Chemical Engineering*, Vol.18, pp.383-413, 1994.