

## A New Ensemble Fault Diagnosis Method Based on K-means Algorithm

Qing Yang<sup>1,2\*</sup>, Chang Liu<sup>1</sup>, Dongxu Zhang<sup>1</sup>, Dongsheng Wu<sup>1,2</sup>

<sup>1</sup> School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110159, China

<sup>2</sup> College of Optical and Electronical Engineering, Changchun University of Science and Technology, Changchun 130022, China

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

---

**Abstract:** A new ensemble algorithm based on K-means clustering and probabilistic neural network called K-means-PNN for classifying the industrial system faults is presented. The proposed technique consists of a preprocessing unit based on K-means clustering and probabilistic neural network. Given a set of data points, firstly the K-means algorithm is used to obtain K-temporary clusters, and then PNN is used to diagnose faults. To validate the performance and effectiveness of the proposed scheme, K-means-PNN is applied to diagnose the faults in TE Process and compared with K-means clustering and back-propagation neural network called K-means-BP, algorithm. Simulation studies show that the proposed K-means-PNN algorithm compared with K-means-BP algorithm not only improves the accuracy in fault classification but also is a reliable and computationally efficient tool.

**Keywords:** Fault diagnosis; K-means algorithm; Probabilistic neural network; Ensemble algorithm; K-means-PNN

---

### 1. Introduction

There has been an increasing interest in fault diagnosis in recent years, as a result of the growing demand for higher performance, efficiency, reliability and safety in control systems. Industrial fault may cause inevitable economic losses, seriously environmental pollution and even fatal injuries. Quick fault diagnosis method can help avoid abnormal event progression and minimize the quality and productivity offsets.

There are many methods for fault diagnosis [1-9]. Usually, fault diagnosis methods are classified into three general categories: quantitative model-based methods, qualitative model-based methods, and process data-based methods [10-12]. In the early days, many researchers diagnose faults by the model-based method. However, model-based FDD uses mathematical system models to estimate the system state and parameters, and in general these methods can only solve the

problem of low dimensional systems. Alternatively, data-driven FDD can deal with high dimensional data, and data dimension reduction techniques [13-17] are generally applied to many areas of information.

In general, fault diagnosis involves a two-step sequential process of symptom extraction and actual diagnosis. Many fault diagnosis algorithms have been proposed in the literature [18], such as statistics, geometric, neural network, expert system, K-means and fuzzy methods, with applications to traffic systems, chemical processes, electrical systems, and rotating machinery [19, 20]. Depending on the information available to classify training, there are two kinds of classification ways. One is supervised learning [21-24], such as PNN and SVM. The other is unsupervised learning [25-27], also called clustering.

As each fault diagnosis technique has its disadvantages, the ensemble method makes traditional classification method improve. The ensemble method can solve many complex problems in fault diagnosis area.

More and more researchers pay attention to the ensemble method.

Coal mining requires various kinds of machinery. The fault diagnosis of this equipment has a great impact on mine production. The problem of incorrect classification of noisy data by traditional support vector machines is addressed by a proposed Probability Least Squares Support Vector Classification Machine (PLSSVCM) [28]. The introduction of a probability value gives the classification results from traditional LSSVCM a quantitative evaluation to match their qualitative explanation. The PLSSVCM is applicable to binary and multi-class problems. Diagnostic results using roller bearings show that PLSSVCM performs well with small-scale fault diagnosis problems. It has a perfect generalization property even with uncertain, noisy samples. If the number of samples is larger than large computation, loads will be needed to use the cross-validation method of determining the regularization factor and the width of a Gaussian kernel. Therefore, future research will focus on the selection of optimal kernel parameters for PLSSVCM that will improve its accuracy even more.

Some articles add some clustering algorithms before ANN classification [29]. R.Eslamloueyan has proposed a duty-oriented hierarchical neural network (DOHANN) for isolating the faults of a relatively complex process. The concept behind the suggested method is to appoint a specific fault diagnostic agent for a particular set of similar fault patterns. The performance of DOHANN has been evaluated and compared to some other methods of fault diagnosis through using the test data of TE process. According to the results, the DOHANN recognizes the process faults much better than SNN. Training of each network in the DOHANN is carried out more conveniently because its neural networks are structurally simpler than that of the SNN model.

Many researchers classify faults by the K-means algorithm. M. Laszlo and S. Mukherjee [30] are motivated by the observation that the popular k-means method for clustering is very sensitive to the initial set of centers with which it is seeded. The new method employs a simple chromosomal representation which is easily implemented and yields partitions of better quality. They have demonstrated that their GA works as well as published methods on the small data sets considered in the literature, and that it produces near optimal partitions for data sets of much larger size and dimensionality. Jim Z.C. Lai, Tsung-Jen Huang and Yi-Ching Liaw [31] present a fast k-means clustering

algorithm (FKMCUCD) using the displacements of cluster centers to reject unlikely candidates for a data point. The computing time of the proposed algorithm increases linearly with the data dimension  $d$ , whereas the computational complexity increases exponentially with the value of  $d$ . Theoretical analysis shows that the method can reduce the computational complexity of full search by a factor of SF and SF is independent of vector dimension.

A key problem in fault detection and diagnosis is the issue of misclassification due to noises or outliers in the raw data. For this reason a combined approach based on K-Means clustering and PNN is presented in this paper.

This paper is organized as follows. The second section briefly introduces fault diagnosis strategy based on K-means-PNN. And to validate the performance and effectiveness of the proposed scheme, the fault diagnosis approach using K-means-PNN based on simulation benchmark of TE process is illustrated in Section III. Finally, concluding remarks are made in Section VI.

## 2. K-means and PNN Algorithm

### 2.1 K-means

The K-means algorithm is a simple and efficient clustering method that has been applied to many engineering problems.

For almost all center-based clustering algorithms, the structure is:

Input:

$K$ : Number of clusters

$X$ : Data set matrix

$N$ : Number of data samples (patterns)

$d$ : Data dimension

Output:

$C$ : Cluster centers matrix

$M$ : Membership matrix

Variable

$Z$ : The k-clustering performance

$d_{ij} = \|x_i - c_j\|$ : Distance from data sample  $x_i$  from center  $c_j$  and the steps are

- Initialization;
- Data membership calculation;
- Center recalculation;
- Convergence check.

K-means algorithm was first proposed by J. B. MacQueen in 1967 [32]. The k-means clustering algorithm partitions data points into  $k$  clusters  $S_j$  ( $j =$

$1, 2, \dots, k$ ) associated with representatives (cluster centers)  $C_j$ . Denote the set of data points as  $S = \{X_m\}$ , where  $m = 1, 2, \dots, N$  and  $N$  is the number of data points in the set  $S$ . Let  $d(X, Y)$  be the distortion between any two vectors  $X$  and  $Y$ . In this paper,  $d(X, Y)$  is defined as the squared Euclidean distance between  $X$  and  $Y$ .

Let  $C_{nm}$  be the nearest cluster center of  $X_m$  and  $d_m = d(X_m, C_{nm})$ . The goal of k-means clustering algorithm is to find a set of cluster centers  $SC = \{C_I\}$  so that the distortion  $J$  defined below is minimized, where  $I = 1, 2, \dots, k$  and  $k$  is the number of clusters.

$$J = \sum_{m=1}^N d_m \quad (1)$$

The major process of k-means clustering is mapping a given set of representative vectors into an improved one through partitioning data points. It begins with an initial set of cluster centers and repeats this mapping process until a stopping criterion is satisfied. The Lloyd iteration for the K-means clustering is given as follows [33].

partition of  $S$  which is divided into  $k$  clusters  $S_I$ , where  $I = 1, 2, \dots, k$  and  $S_I = \{X | d(X, C_I) \leq d(X, C_j)\}$  for all  $I \neq J$

Compute the centroid for each cluster to obtain a new set of cluster representatives  $SC_{p+1}$ .

The k-means clustering algorithm is briefly described as follows.

Step1. Begin with an initial set of cluster centers  $SC_0$ . Set  $p = 0$ .

Step2. Given the set of cluster centers  $SC_p$ , perform the Lloyd iteration to generate the improved set of cluster representatives  $SC_{p+1}$ .

Step3. If it is changed by a small enough amount since the last iteration, then stop. Otherwise set  $p + 1 \rightarrow p$  and go to step2.

The nearest cluster center is determined by computing the Euclidean distances between each cluster center and all data points. The Euclidean distance between a data point  $X = (X_1, X_2, \dots, X_m)^d$  and cluster center  $C = (C_1, C_2, \dots, C_m)^d$  is defined as

$$d(X, C) = \left[ \sum_{k=1}^m |X_m - C_k|^2 \right]^{0.5} \quad (2)$$

Figure 1 shows the sketch of K-Means algorithm.

## 2.2 Probabilistic neural network

Probabilistic neural network (PNN) is a special type of radial basis-function networks [34]. The structure of the PNN classifier [35] is illustrated in Figure 2. It consists of an input layer, pattern layer, summation layer and an output layer. The input layer is merely a distribution layer. No computation is performed in this layer. Neurons in the pattern layer utilize multi-dimensional kernels to estimate the probability density function for classification. On receiving a pattern  $x$  from the input layer, the neuron  $x_{ij}$  of the pattern layer computes its output using the following equation:

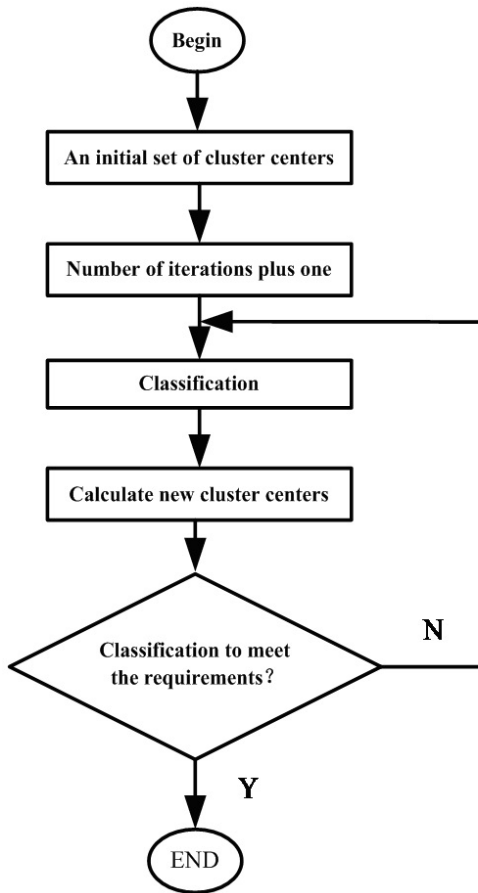


Figure 1 The sketch of K-Means algorithm

Given a set of cluster centers  $SC_P = \{C_j\}$ , find the

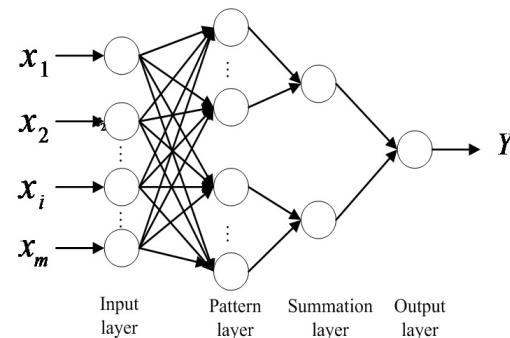


Figure 2 The architecture of the PNN

$$\varphi_{ij}(x) = \frac{1}{(2\pi)^{a/2} \sigma^a} \exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right] \quad (3)$$

where  $a$  denotes the dimension of the pattern vector  $x$ ,  $\delta$  is the smoothing parameter and  $x_{ij}$  is the neuron vector.

Each neuron in the summation layer computes the maximum likelihood of pattern  $x$  being classified into class  $c$  by summarizing and averaging the output of all neurons that belong to the same class:

$$P_i(x) = \frac{1}{(2\pi)^{a/2} \sigma^a} \frac{1}{N_i} \sum_{j=1}^{N_i} \exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right] \quad (4)$$

where  $N_i$  denotes the total number of samples in class  $c_i$ . If the priori probabilities for each class are the same and the losses associated with making an incorrect decision for each class are the same, the output layer classifies the sample  $x$  accordance with the Bayes's decision rule based on the outputs of all the summation layer neurons and the final result of output is  $Y$ .

### 2.3 K-means-PNN

This algorithm is proposed to improve the degree of accuracy on fault diagnosis. At the same time, it is a combination of algorithm about K-means and PNN. Figure 3 shows the sketch of the algorithm, which is briefly described as follows:

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

Step 2: Clustering the data sets using K-means algorithm.

Step 3: Diagnosing the fault data of K-means clustering by PNN and BP.

Step 4: Comparing the clustering results of the K-means with two different methods.

## 3. Case Study

### 3.1 K-means algorithm clusters the data of Iris flower

The Iris flower data set or Fisher's Iris data set is a multivariate data set introduced by Sir Ronald Aylmer Fisher (1936) as an example of discriminant analysis [36]. It is sometimes called Anderson's Iris data set because Edgar Anderson collected the data to quantify the geographic variation of Iris flowers in the Gaspé Peninsula [37].

The dataset consists of 50 samples from each of three species of Iris flowers (Iris setosa, Iris virginica

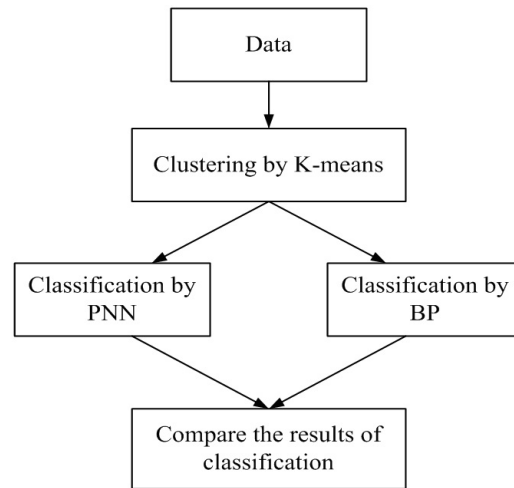


Figure 3 The sketch of the process

and Iris versicolor). Four features are measured from each sample which are the length and the width of sepal and petal. Based on the combination of the four features, Fisher developed a linear discriminant model to distinguish the species from each other. Based on Fisher's linear discriminant model, this data set became a typical test case for many classification techniques in machine learning, such as support vector machines.

This data set in cluster analysis is uncommon, since the data set only contains two clusters with rather obvious separation. One of the clusters contains the Iris setosa species, while the other cluster contains Iris virginica and Iris versicolor, and is not separable without the species information Fisher used. This makes the data set a good example to explain the difference between supervised and unsupervised techniques in data mining: Fisher's linear discriminant model can only be obtained when the object species are known.

In order to test the accuracy of the K-means algorithm, the data set of Iris flower is used. Figure 4 shows that the K-means algorithm divides Iris flowers into three species based on Iris flowers length of sepal feature. The red is virginica, the blue is setosa, and the green is versicolor.

Experiment shows that the k-means algorithm can cluster Iris flower data distinctly.

### 3.2 K-means-PNN diagnoses the fault data of TE

An example of the application of the proposed strategy is presented, and a combined approach based on K-Means clustering and PNN was analyzed under the TE process (see Figure 5). TE process [38-39] is a benchmark problem in process engineering. Downs and

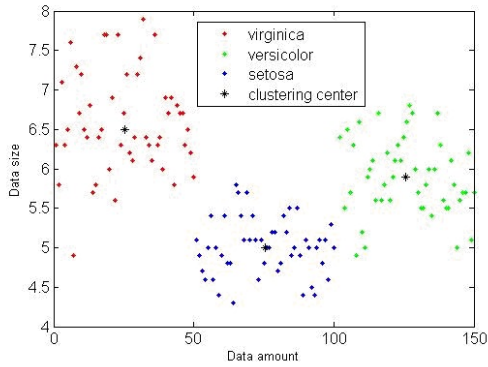


Figure 4 Clustering result of Iris flower

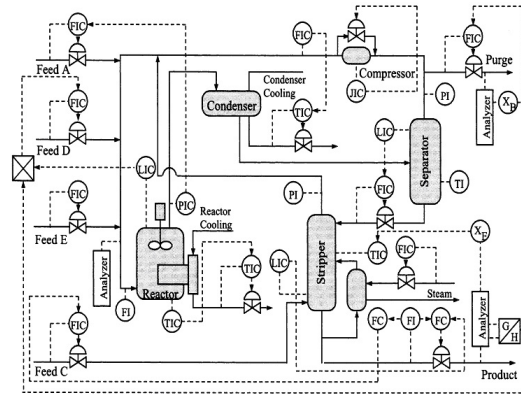


Figure 5 Control system of the Tennessee Eastman process

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 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.

First we divide the TE process data into six species by K-means algorithm. The following fault data are obtained by wavelet processing. Figure 6 shows the clustering result.

The faults clustered by K-means algorithm are as follows:

- Cluster No.1: { F1, F2, F12}
- Cluster No.2: { F3, F9, F13, F14}
- Cluster No.3: { F11, F15, F16, F21}
- Cluster No.4: { F7, F10, F17, F19}
- Cluster No.5: { F4, F5, F20}
- Cluster No.6: { F6, F8, F18}

PNN is a feed-forward neural network with supervised learning, which uses Bayes decision rule and Parzen window. In this paper, PNN is used to diagnose faults in each cluster.

Figure 7 shows the fault types divided by K-means-PNN. In Figure 7, each segment contains 300 data and cluster No.1 is classified into three kinds of types by K-means-PNN. Label 1 is normal data, label 2 is fault 1, label 3 is fault 2, and label 4 is fault 12. In Figure 7, normal data segment contains 300 data, each fault type also contains 300 data, and the total data is 1200.

Table 1 Process faults for the Tennessee Eastman process

| Variable | Disturbances  | Type                   |
|----------|---|------------------------|
| 1        | A/C feed ratio, B composition constant                        | Step                   |
| 2        | B composition, A/C ratio 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 |

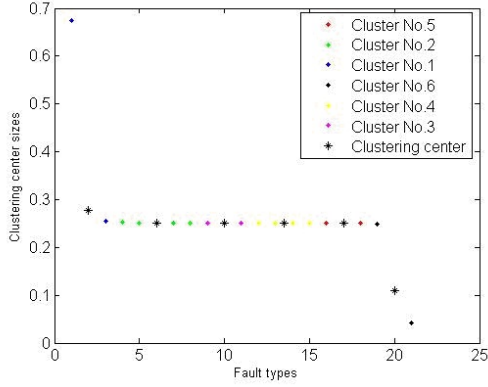


Figure 6 Clustering result of all fault types

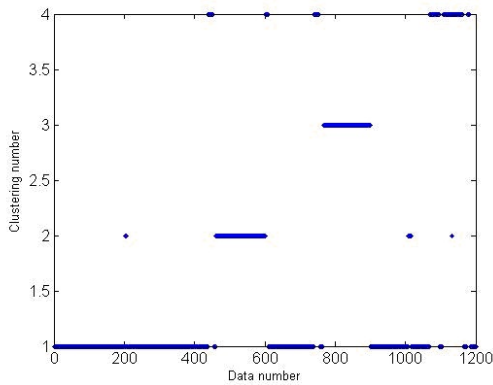


Figure 7 Clustering result of K-means-PNN on No.1

In each fault type segment, the first 160 data are the normal data and the last 140 are the fault data. Figure 7 shows that the correct percentage of classification by K-means-PNN is 91.33%.

Experiment result indicates that K-means-PNN is a kind of efficient classification method.

### 3.3 K-means-BP diagnoses the fault data of TE

BP neural network is a kind of feed forward, whole link, and multi-layers network. Its characteristics include the association ability, memory ability, self-flexibility, self-learning and so on. It provides a new method to solve the prediction and control problem for the non-linear system [40].

Currently, BP algorithm is widely used for training neural networks [41, 42]. Supposing a set of  $P$  training samples is available, the problem can be characterized as the process of minimizing the following sum squared error:

$$J(W) = \frac{1}{2} \sum_{s=1}^P \sum_{i=1}^N (d_{s,i} - y_{s,i})^2 \quad (5)$$

where  $d_{s,i}$  and  $y_{s,i}$  are the  $i$ th target and actual outputs corresponding to the  $s$ th training pattern,  $W$  is a vector composed of all the weights and biases involved in the network, and  $N$  is the number of output units. In this scheme, an initial weight vector  $W_0$  is iteratively adapted according to the following recursion to find out an optimal weight vector. The positive constant  $\eta$  is the learning rate.

$$W_{k+1} = W_k - \eta \frac{\partial J(W)}{\partial W} \quad (6)$$

In general, the BP algorithm [43] includes the forward course and the backward course. In the forward course, a vector is added to the input layer, which is then spread along the network; finally, an output vector is obtained as a response of the input vector, in which the synaptic weights can not be changed. In the backward course, an error signal will be obtained by comparing the output signal with the desired output; the error signal is forward-spread to modify the weight from one output layer to another. The modified network will output the signal that is closer to the desired output. The forward course and the backward course alternate and constantly circulate, and the output will be convergent with the desired output in some states.

Figure 8 shows that No.1 is classified to three kinds of types by K-means-BP. Label 1 is the normal data, label 2 is the fault 1, label 3 is the fault 2, and label 4 is the fault 12. In Figure 8, each segment contains 300 data, and the total data is 1200. In each fault type segment, the first 160 data are the normal data and the last 140 are the fault data. Figure 8 shows that the correct percentage of K-means-BP is 60%.

Comparing Figure 7 with Figure 8, experiment proves that the K-means-PNN method has a better classification result than the K-means-BP method.

## 4. Conclusion

An ensemble fault diagnosis method based on K-means and PNN called K-means-PNN is presented in this paper. Simulation studies show that the proposed algorithm provides an accepted degree of accuracy in fault classification under different fault conditions and the result is also reliable.

## 5. Acknowledgments

This work was supported by Liaoning Science and Technology Plan Projects, China, under Grant 201022-2005.

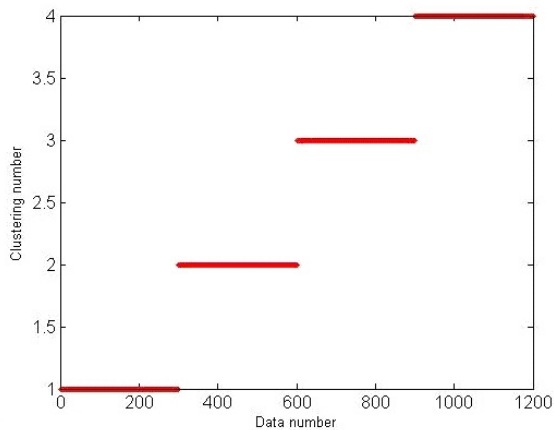


Figure 8 Clustering result of K-means-BP on No.1

## References

- [1] B. Han, and S.J. Lee, "A genetic algorithm approach to measurement prescription in fault diagnosis", *Information Sciences*, Vol.120, pp.223-237, 1999.
- [2] F. Li, B.R. Upadhyaya, and L.A. Coffey, "Model-based monitoring and fault diagnosis of fossil power plant process units using Group Method of Data Handling", *ISA Transactions*, Vol.48, pp.213-219, 2009.
- [3] R. Vaez-Ghaemi, W. Godbersen, H. Schwetlick, and D. Filbert, "Some aspects of knowledge-based fault diagnosis in electronic devices", *Measurement*, Vol.10, pp.2-7, 1992.
- [4] Y.G. Lei, Z.J. He, and Y.Y. Zi, "EEMD method and WNN for fault diagnosis of locomotive roller bearings", *Expert Systems with Applications*, Vol.38, pp.7334-7341, 2011.
- [5] M. F. White, "Expert systems for fault diagnosis of machinery", *Measurement*, Vol.9, pp.163-171, 1991.
- [6] G.M. Xian, and B.Q. Zeng, "An intelligent fault diagnosis method based on wavelet packer analysis and hybrid support vector machines", *Expert Systems with Applications*, Vol.36, pp.12131-12136, 2009.
- [7] 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.
- [8] Z. Czaja, "A method of fault diagnosis of analog parts of electronic embedded systems with tolerances", *Measurement*, Vol.42, pp.903-915, 2009.
- [9] J. Huang, X.G. Hu, and X. Geng, "An intelligent fault diagnosis method of high voltage circuit breaker based on improved EMD energy entropy and multi-class support vector machine", *Electric Power Systems Research*, Vol.81, pp.400-407, 2011.
- [10] 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.
- [11] Y.W. Chang, and Y.L. Chen, "Mining associative classification rules with stock trading data - A GA-based method", *Knowledge-Based Systems*, Vol.23, pp.605-614, 2010.
- [12] S. Brofferio, G. Mastronardi, and V. Rampa, "A migrating data-driven architecture for multidimensional signal processing", *Signal Processing: Image Communication*, Vol.3, pp.249-257, 1991.
- [13] I. Budak, M. Sokovic, and B. Barisic, "Accuracy improvement of point data reduction with sampling-based methods by Fuzzy logic-based decision-making", *Measurement*, Vol.44, pp.1188-1200, 2011.
- [14] D. Sudhanva, and K.C. Gowda, "Dimensionality reduction using geometric projections: A new technique", *Pattern Recognition*, Vol.25, pp.809-817, 1992.
- [15] J.D. Tubbs, W.A. Coberly, D.M. Young, "Linear dimension reduction and Bayes classification with unknown population parameters", *Pattern Recognition*, Vol.15, pp.167-172, 1982.
- [16] F.S. Tsai, "Dimensionality reduction techniques for blog visualization", *Expert Systems with Applications*, Vol.38, pp.2766-2773, 2011.
- [17] J. Goddard, G. Schlotthauer, M.E. Torres, and H.L. Rufiner, "Dimensionality reduction for visualization of normal and pathological speech data", *Biomedical Signal Processing and Control*, Vol.4, pp.194-201, 2009.
- [18] K.O. Kima, and M.J. Zuo, "Two fault classification methods for large systems when available data are limited", *Reliability Engineering & System Safety*, Vol.92, pp.585-592, 2007.
- [19] R. Isermann, and P. Ballé, "Trends in the application of model-based fault detection and diagnosis of technical processes", *Control Eng Practice*, Vol.5, pp.709-719, 1997.
- [20] S. Leonhardt, and M. Ayoubi, "Methods of fault diagnosis", *Control Eng Practice*, Vol.5, pp.683-692, 1997.
- [21] J.D. Wu, P.H. Chiang, Y.W. Chang, and Y.J. Shiao, "An expert system for fault diagnosis in internal combustion engines using probability neural network", *Expert Systems with Applications*, Vol.34, pp.2704-2703, 2008.
- [22] J.D. Wu, and P.H. Chiang, "Application of Wigner-Ville distribution and probability neural network for scooter engine fault diagnosis", *Expert Systems with Applications*, Vol.36, pp.2187-2199, 2009.

- [23] 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.
- [24] L.M.R. Baccharini, V.V.R. e Silva, 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.
- [25] 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.
- [26] 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.
- [27] H.S. Wang, and B.W. Fei, "A modified fuzzy C-means classification method using a multiscale diffusion filtering scheme", *Medical Image Analysis*, Vol.13, pp.193-202, 2009.
- [28] 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.
- [29] 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.
- [30] M. Laszlo, and S. Mukherjee, "A genetic algorithm that exchanges neighboring centers for k-means clustering", *Pattern Recognition Letters*, Vol.28, pp.2359-2366, 2007.
- [31] J.Z.C. Lai, T.J. Huang, and Y.C. Liaw, "A fast k-means clustering algorithm using cluster center displacement", *Pattern Recognition*, Vol.42, pp.2551-2556, 2009.
- [32] J. Macqueen, "Some methods for classification and analysis of multivariate observation", *Proc. 5th Berkeley Symp. Math. Statist.*, Vol.1, pp.281-297, 1967.
- [33] A. Gersho, and R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston MA, USA, 1991.
- [34] P.D. Wasserman, *Advanced Methods in Neural Computing*, VNR Press, New York, 1993.
- [35] S.N. Yu, and Y.H. Chen. "Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network", *Pattern Recognition Letters*, Vol.28, pp.1142-1150, 2007.
- [36] R.A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems", *Annals of Eugenics*, Vol.7, pp.179-188, 1936.
- [37] E. Anderson, "The irises of the Gasp Peninsula". *Bulletin of the American Iris Society*, Vol.59, pp.2-5, 1935.
- [38] J.J. Downs, and E.F. Vogel, "A plant-wide industrial-process control problem," *Computers & Chemical Engineering*, vol.17, pp.245-255, 1993.
- [39] T.J. Mcavoy, and N. Ye, "Base control for the Tennessee Eastman problem," *Computers & Chemical Engineering*, Vol.18, pp.383-393, 1994.
- [40] Q.M. Zhu, "A back propagation algorithm to estimate the parameters of non-linear dynamic rational models", *Applied Mathematical Modelling*, Vol.27, pp.169-187, 2003.
- [41] Y.J. Sun, S. Zhang, C.X. Miao, and J.M. Li, "Improved BP Neural Network for Transformer Fault Diagnosis", *Journal of China University of Mining and Technology*, Vol.17, pp.138-142, 2007.
- [42] F. Marini, *Neural Networks*, Comprehensive Chemometrics, Italy, 2009.
- [43] X.Y. Zhao, K.S. Lai, and D.M. Dai, "An Improved BP Algorithm and Its Application in Classification of Surface Defects of Steel Plate", *Journal of iron and steel research, international*, Vol.14, pp.52-55, 2007.