CHEMICAL ENGINEERING TRANSACTIONS
VOL. 46, 2015

A publication of
AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Peiyu Ren, Yancang Li, Huiping Song

ISBN 978-88-95608-37-2; **ISSN** 2283-9216

DOI: 10.3303/CET1546054

A Fault Diagnosis Based On Combination Model Of VPRS And PSO-BP Neural Network

XiaolingNiu, Jun Wang*, Zihui Ren

School of Information and Electrical Engineering, China University of Mining and Technology,Xuzhou, 221116,China
Wj999lx@163.com

Against the disadvantage of slow convergence and falling into local optimal illustrated by the traditional fault diagnosis based on the BP model, the paper proposes a new diagnosis method by combining the variable precision rough set with the PSO-BP neural network. The proposed model covers the new design of heuristic β simplification algorithm based on the classification quality, the idea of which could dissolve the input redundancy of the network so as to get the simplify the network and decrease the time consumption of training; on the other hand, the improved PSO -BP model could better conquer the problem of local optimal for BP neural network by taking good advantage of the global data search ability of PSO algorithm and local quick search ability of BP algorithm. In the experiment section, the paper conducts the fault diagnosis model on the antifriction bearing to illustrate the efficiency of the proposed model.

1. Introduction

The Recently, various kinds of methods for fault diagnosis have been proposed such as the famous experts system(Yang et al.(2012) reported)、Neural network-based model(Liu et al.(2015) reported)、Support vector machine-based model(Ao et al.(2015) reported)、Rough set-based model(Sun et al.(2013) reported) etc. The overall characteristics of the above models are that each of them performs better than the others in some specific field and there doesn't exist the one suitable for all the possible problems. In conclusion, the establishment of new model with all the advantages of the above mentioned models equipped is in great need and will be the most promising direction of the fault diagnosis subject. The paper proposes a new model by combining the variable precision rough set with the neural network, which takes good advantage of the data simplifying ability and the high-precise prediction ability to get the reasonable fault diagnosis model.

2. Diagram of the proposed model

The flowing chart of the proposed diagnosis model is shown as Figure 1, and the key design is setting the variable precision rough set as the former data processing part of the PSO-BP neural network. The specific steps to establish the proposed model is expressed as follows:
(1) Conducting the sampling operation for N times at the level of and convert the sampling signals into the original input data with different state of the equipment;
(2) Extracting the feature vector after the basic pre-process on the input signal to get the original problems vector set with two dimensional spread;
(3) Conducting the discrete processing on the continuous problem data, and turning the attribute of each section into Figurer expression by the value judgment such as "1"、 "2"、 "3" etc., and getting the attribute expression set including the Condition set(C) and Decision set (D), and the definition of the sample is set as U to get the final decision table;
(4) Set the fiducially threshold as β, and conducting the β-format simplification on the condition attribute set based on the variable precision rough set theory to find out the minimum-correlation attribute sub-set which maintains the original attribute of the decision table;
(5) Constructing the neural network model. Grouping the attribute set from step (4) and the corresponding data collection as the training set and put them into the neural network model to train and self-learning to get the best-performance model;

Please cite this article as: Niu X.L., Wang J., Ren Z.H., 2015, A fault diagnosis based on combination model of vprs and pso-bp neural network, Chemical Engineering Transactions, 46, 319-324 DOI:10.3303/CET1546054

(6) Put the trained model into the practical fault diagnosis. (Foot) notes

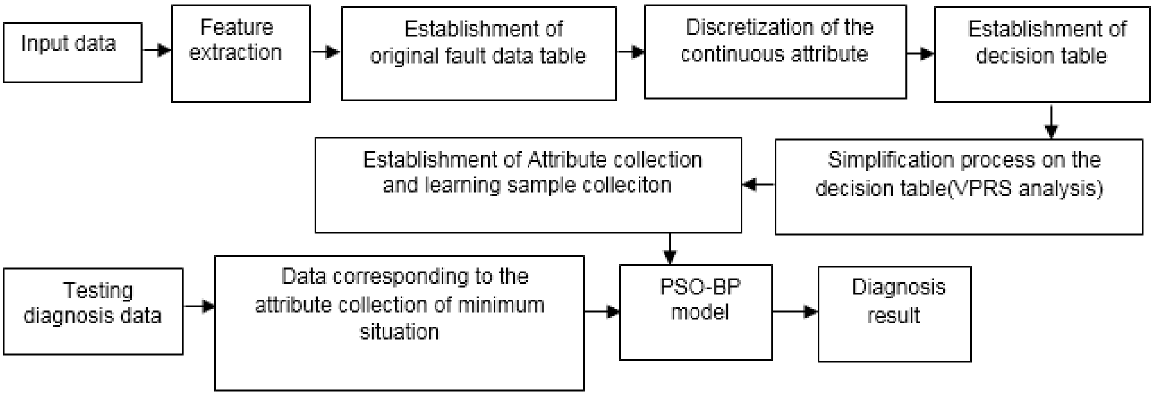


Figure. 1 Diagram of the diagnosis system

3. Theory of the variable precision rough set

The definition of variable precision rough set was proposed on the basis of the RS model (Pawlak (1982) reported), and the key is adding the deviation accuracy to the RS model to make it the certain fault tolerance, and on the condition of , it turns out the original RS model. Later the is redefined as correct classification rate, the domain of which is set as (An et al. (1996) reported), the proposed model of the paper is set according to this definition.

Definition 1: The decision table is set as $S = \{U, C \cup D, V, f\}$, and for any attribute sub-set $\Phi \neq B \subseteq (C \cup D)$, and the expression $I(B)$ is set for illustrating the non-classifying relationship for B on U, so $I(B) = \{(x, y) \in U \times U | \forall a \in B, f(x, a) = f(y, a)\}$.The equal one for U from $I(B)$ could be expressed as $U/I(B)$ or U/B

Definition 2: As for the decision table $S = \{U, C \cup D, V, f\}$, attribute collection $\Phi \neq B \subseteq (C \cup D)$ and the undefined subject $X \subseteq U, x \in U$, as for the predefined threshold $\beta \in (0.5,1]$, it could be concluded that:

The up β –approximate set for collection X on B is defined: $\underline{apr}_B^\beta(X) = \cup\left\{Y \in U/B | \frac{|Y \cap X|}{|Y|} \geq \beta\right\}$

The low β –approximate set for collection X on B is defined: $\overline{apr}_B^\beta(X) = \cup\left\{Y \in U/B | \frac{|Y \cap X|}{|Y|} \geq 1 - \beta\right\}$

Definition 3: For some given decision table $S = \{U, C \cup D, V, f\}$, supposing the decision categories covers $U/D = \{D_1, D_2, \cdots, D_n\}$, and for the condition attribute sub-set $\Phi \neq B \subseteq C$, the approximate set for collection B on D is set as $\gamma_B^{\beta}(D) = \frac{\sum_{i=1}^{n} |\underline{apr}_B^{\beta}(Di)|}{|U|}$ under the condition $\beta \in (0.5,1]$.

Definition 4: For some given decision table $S = \{U, C \cup D, V, f\}$, the definition of β –approximate $red^\beta(C, D)$ for condition attribute C on decision attribute D is defined under the following two conditions:

$$\gamma^\beta(C,D) = \gamma^\beta(red^\beta(C,D),D);$$

For the given condition of cutting any attribute sub-set form $red^\beta(C, D)$, the expression of (1) would be unreasonable.

4. Selection of the threshold for the VPRS and the β-simplification algorithm

4.1 Method for Figuring out the range of β

The definition of β, which is set as right classification accuracy, is an important parameter for the VPRS. The specific method for selection of β is on the basis of the theory (Beynon (2001) reported): If the condition kinds X is distinguishable during the threshold $\beta \in (0.5,1]$, it could be concluded that X is also distinguishable during any range set $\beta_1 \in (0.5, \beta]$; on the contrary, if it is undistinguishable, any other data of X is also distinguishable during the range set $\beta_2 \in (\beta, 1]$.

The theorem described above demonstrates that every dataset is distinguishable when equal to or less than the upper set of β, so we design the algorithm for the calculation of β_{max} under the condition of maintaining the best similar classification quality:

Input dataset: Decision information system$S = (U, A = C \cup D, V, f)$, set C as the non-empty condition attributes set and D as the non-empty decision attribute set.

Output: the range ofβ.

(1) Figureuring out the condition kinds of the decision system $U/C = \{X_1, X_2, \cdots, X_n\}$ and the decision kinds $U/D = \{Y_1, Y_2, \cdots, Y_m\}$;

(2) Figureuring out the inclusion degree for all the decision condition at each conditionX_i, and the result is $P(Y_j \mid X_i)) = \frac{|X_i \cap Y_j|}{|X_i|}$;

(3) Set$m_1 = 1 - \max\{P(Y_j|X_i)|\forall i \leq n, P(Y_j|X_i) < 0.5\}$;

$m_2 = \min\{P(Y_j|X_i)|\forall i \leq n, P(Y_j|X_i) > 0.5\}$; So the distinguishable threshold for decision set Y_i could be expressed as: $\beta_i = \min(m_1, m_2)$;And the max distinguishable threshold under the estimation of decision table requirements is set as:

$$\beta_{\max} = \min_{j \leq m}(\beta_j)$$

(4) The output range for β is set as$\beta \in (0.5, \beta_{\max}]$.

4.2 β -simplification algorithm based on VPRS

Attribute simplification refers to the procedure of cutting the unnecessary or unimportant attribute under the condition of maintaining the classification ability for condition attribute over decision attribute. According to the definition in third section, the paper designs a new set of decision table attribute simplification algorithm by seeing the classification quality as inspiration knowledge:

Input: The discretization decision table $S = \langle U, C, D \rangle$

Output: The relatively simplified edition for the decision table $S = \langle U, RED, D \rangle$

Step1: Figureuring out the condition class $\frac{U}{C} = \{X_1, X_2, \cdots, X_n\}$ and the decision class $\frac{U}{D} = \{Y_1, Y_2, \cdots, Y_n\}$

Step2: Figureuring out the value range of β according to the theory in 4.1

Step3: For the givenβ, calculating the approximate classification quality $\gamma^{\beta}(C, D)$of condition attribute C over decision attribute D

Step4: Calculating all the approximate classification quality $\gamma^{\beta}(C, D)$ for each attribute C_i of the table and then set the calculated value in descending order in size to form the attribute set C', taking the combination process if there has been two or more attribute classification quality equaling to get the kinds of attribute set C'

Step5: Import the fault feature subset $RED = \{\}$ for each attribute set C'

Step6: Find out the attribute C_i with the highest classification quality score fromC', and define $RED = RED \cup C_i$ then Figureure out$\gamma^{\beta}(RED, D)$, if there exists the condition $\gamma^{\beta}(RED, D) = \gamma^{\beta}(C, D)$ then getting the result, or else going back to step6 after defining $C' = C' - C_j$

Step7: Getting the fault feature subset RED

5. Improved PSO-BP algorithm

Particle swarm optimization (Kennedy et al. (1995) reported) is a global optimization algorithm, the BP algorithm is a local search optimization method, By combining PSO algorithm and BP algorithm, it can solve the problem that the BP network(Rumelhart et al. (1986) reported) can easily fall into local extreme value,The idea of traditional PSO-BP model is constructing the projection from the weight and threshold of network to the PSO particle, and realize the training of the network by the way of continuous iteration on the speed and location information of particle to optimize the weight and threshold information(Wang and Jiang (2012) reported). The analysed shortcoming of the model is the high dimensionality of the particles, the performance of which is that the dimensionality of each particle is $I \times H + H \times O + H + O$ as a result of the topological three-layer network architecture I-H-O, and it results in the low efficiency. The paper proposes the heuristic model from the PSO search to BP search, and the specific idea is illustrated as follows:

(1) In the weight update equations of BP algorithm, the learning speed η is set two valuesη_1, η_2, and the corresponding dimensionality of PSO is 2, which includes η_1, η_2 respectively. The idea of PSO is set to find the optimal learning factor, the weight and threshold of BP is refreshed by the BP algorithm. The paper adopts the mean square error distribution (MSE) of the BP neural network as the fitness function

(2) To get the better optimal result, the model additionally set the mutagenic factor in the PSO algorithm, which means the speed and location of the particle is refreshed by the following rules:

$$v_i(k+1) = \omega v_i(k) + c_1 r_1 (pbest_i(k) - x_i(k)) + c_2 r_2 (gbest(k) - x_i(k))$$

$$x_i(k+1) = x_i(k) + pv_i(k+1)$$

(3) The optimal particle and the worst particle were treated. The optimal particle continues to fly from the new position when the fitness meliorates, or else returns to the original position to research. The worst particle will be replaced by the global optimal particle in each step

6. Fault diagnosis by the proposed model

6.1 Collection and preprocess of the data set

The experiment data of rolling bearing for fault diagnosis covers 3 different situations with a total of 32 samples, 11 of which are normal, 9 of which is outer ring fault (ORF), 12 of which is support platform fault (SPF). The testing signals is the vibration acceleration information and each situation is packed with N=2048 sample points spread. The original signal is the approximate entropy in the 8 frequency bands after adopting the 3-layers wavelet packet decomposition. The specific experiment is conducted with method of "training-testing" with grouping the 32 groups of data into training one and testing one (4:1), adopting the Particle swarm-based continuous attribute discretization method (Wang (2013) reported) to process the data, and then the decision table shown in table.1 would be obtained. The condition attribute set is $C = \{c_{30}, c_{31}, c_{32}, c_{33}, c_{34}, c_{35}, c_{36}, c_{37}\}$, the decision attribute set is $D = \{FN\}$. $FN = 1$ represents the normal one, $FN = 2$ represents the ORF one, and $FN = 3$ represents the SPF one.

Table 1. Training dataset after the discretization process

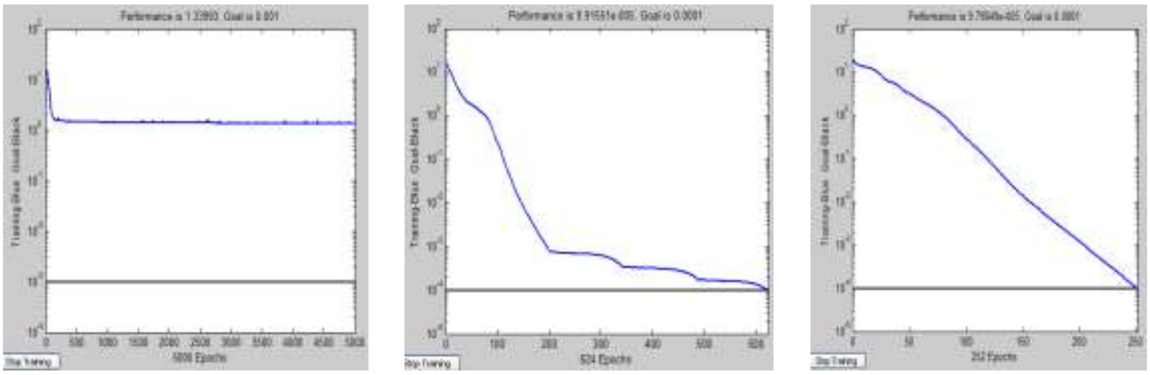
U	C30	C31	C32	C33	C34	C35	C36	C37	FN
X_1	2	3	3	3	1	1	3	1	1
…	…	…	…	…	…	…	…	…	…
X_{11}	2	3	1	2	1	2	1	2	1
X_{12}	3	1	3	3	3	1	3	3	2
…	…	…	…	…	…	…	…	…	…
X_{20}	3	2	2	2	1	2	3	2	2
X_{21}	2	1	2	3	3	2	1	1	3
…	…	…	…	…	…	…	…	…	…
X_{32}	1	1	1	1	2	2	3	2	3

6.2 β -simplification on the attribute dataset

It can be inferred from the theoretical part that the range of β is set as $\beta \in (0.5,1]$, and the paper set $\beta = 0.85$. Conducting the β-simplification on the dataset in table. 1 based on the variable precision rough set theory and the finale result is $\{c_{30}, c_{31}, c_{35}\}$.

6.3 Neural network model training

To illustrate the result overall, the paper try the three mentioned model to conduct the fault diagnosis experiment with the various dataset. Model 1 is BP network system, Network structure is 8-19-3, the target error is set as 0.001, the maximum training steps is 5000, the learning speed is 0.01; Model 2 is PSO-BP network system, the topological architecture of which is 8-19-3, and the other related parameters is set as follows: the group scale is set as ,the dimensionality in the solution space is set as ,the accelerate parameters are set as , the boundary value of w is set as =0.9 and =0.4, the upper boundary value is set as and the maximum iteration times is set as 500; Model 3 is VPRS-PSO-BP network system, The input feature vectors is ,Network structure is 3-7-3, The parameter set of PSO is the same as PSO-BP model. The error curve of each model during the processing is shown as Figure. 2.



(a) BP model (b) PSO-BP model (c) VPRS-PSO-BP model

Figure. 2The error curve of model

6.4 Result and discussion

Table .2 Result of fault diagnosis by three models

Groups	BP model			PSO -BP model			VPRS -PSO -BP model			Expectation
1	0.0001	0.0082	0.9970	0.0000	0.0029	0.9877	0.0000	0.0000	1.0000	(001)
2	0.5232	0.0001	0.4908	0.0359	0.0000	0.9961	0.0789	0.0000	0.9998	(001)
3	0.0050	0.9970	0.0000	0.0000	0.9965	0.0007	0.0000	0.9998	0.0000	(010)
4	0.0004	0.9982	0.0000	0.0000	0.9961	0.0022	0.0000	0.9994	0.0000	(010)
5	0.8817	0.0000	0.1268	0.9872	0.0001	0.0349	0.9990	0.0008	0.0002	(100)
6	0.8720	0.0000	0.1159	0.9995	0.0000	0.0005	0.9998	0.0000	0.0006	(100)

Table .3 Result comparison of fault diagnosis on three models

Models	Network structure	Attribute set	Training error	Training steps	Norm
BP model	8-19-3	$\{C_{30}\text{-}C_{37}\}$	1.33993	5000	0.9458
PSO-BP model	8-19-3	$\{C_{30}\text{-}C_{37}\}$	0.000099	624	0.0090
VPRS-PSO-BP model	3-7-3	$\{C_{30},C_{31},C_{35}\}$	0.000098	252	0.0008

Taking the diagnosis experiment on the 6 test samples in table.1 based on the 3 classical model above, and the practical result is illustrated in table.2. To obtain the further analyzing result of accuracy, the paper proposes the norm calculation of absolute error to better describe the severity of the fault, the criterion of which is the higher the value, the more severity of the machine. The detailed training and result information of the three models is listed in table.3.

In conclusion, the proposed model could overcome the disadvantage of trapping into the local optimal point of BP model, and the speed of convergence is more quickly and time consumption is less than the other two classical models. On the other hand, the structure design of the proposed model could simplify the distribution demand on the sample data and reduce the nodes number of the hidden layer, the design of which results in the optimization neural network model.

7. Conclusion

The paper proposes a new fault diagnosis model by the theoretical analysis on the shortcomings of the existed model, which combines the VPRS model and the classical PSO-BP model to get a better performance model. What's more, the paper put the proposed model into the practical rolling bearing fault diagnosis experiment, and it can be concluded from the result that the importation of VPRS into the existing neural network model, which is set as the tool for data preprocessing, results in the redundancy-removal, simplification of network structure and the increase in training and testing speed. In a word, the result from the simulation suggests that the diagnosis accuracy by the proposed model is better than that without the data preprocessing by VPRS.

Acknowledgements

The research described in this paper has been supported by three main projects: The Fundamental Research Funds for the Central Universities of China (Grant No: 2011QNB21); The Innovative engineering developing on graduates of Jiangsu Province in China (Grant No: CXZZ11-0293); The Chinese Coal Industry Association of science and technology guiding plan project (Grant No: MTKJ2011-306).

Reference

An A.J., Shan N., Chan C., 1996, Discovering rules for water demand prediction: an enhanced rough-set approach,Engineering Applications in Artificial Intelligence, 9(6),pp:645-653, DOI: 10.1016/S0952-1976(96)00059-0

Ao H., Cheng J. , Zheng, J., Truong TK, 2015, Roller Bearing Fault Diagnosis Method Based on Chemical Reaction Optimization and Support Vector Machine, Journal of Computing in Civil Engineering , 29 (5), 04014077, DOI:10.1061/(ASCE)CP.1943-5487.0000394

Beynon M., 2001, Reducts within the variable precision rough sets model: a further investigation, European Journal of Operational Research, 134(3), pp: 592-605, DOI: 10.1016/S0377-2217(00)00280-0

Kennedy J., Eberhart R., 1995, Particle swarm optimization, proceedings of IEEE International conference on Neural Networks, Volume 4, pp. 1942-1948, DOI: 10.1109/ICNN.1995.488968

Liu M.L., Wang K.Q., Sun L.J., 2015, Fault Diagnosis Method of HV Circuit Breaker Based on Wavelets Neural Network, Open Automation & Control Systems Journal , 7(1),pp.126-134 , DOI:10.2174/1874444301507010126

Pawlak Z., 1982, Rough sets, International Journal of Computer and Information Sciences, Volume 11, Issue 5,pp:341-356, DOI:10.1007/BF01001956

Rumelhart D.E., Hinton G.E., Williams R.J., 1986, Learning representations by back-propagating errors,Nature, Vol. 323(6088),pp.533-536, DOI: 10.1038/323533a0

Sun Q.Y., Wang C.L., Wang Z.L., Liu X.R., 2013, A fault diagnosis method of Smart Grid based on rough sets combined with genetic algorithm and tabu search, Neural Computing & Applications , 23(7),pp:2023-2029, DOI:10.1007/s00521-012-1116-x

Wang A.P., Jiang L., 2012, BP Neural Network Learning Algorithm Based on Particle Swarm Optimization. Computer Engineering, 38(21):193-196, DOI: 10.3969/j.issn.1000-3428.2012.21.052

Wang L., 2013, Algorithm of continuous attribute discretization based on improved particle swarm. Computer Engineering and Applications, 49(21):29-32, DOI: 10.3778/j.issn.1002-8331.1305-0492

Yang Z.L., Wang B., Dong X.H., Liu H., 2012, Expert System of Fault Diagnosis for Gear Box in Wind Turbine, Systems, Volume,pp.189-195, DOI:10.1016/j.sepro.2011.11.065