The Degradation State Recognition of Rolling Bearing Based on GA and SVM

Yonghe Wei
School of Mechanical Engineering
Shenyang Ligong University,SLU
Shenyang, China
e-mail：yonghewe@sina.com

Minghua Wang
School of Mechanical Engineering
Shenyang Ligong University,SLU
Shenyang, China
e-mail：517752960@qq.com

Abstract —In order to accurately recognize the degradation state of rolling bearing, a hybrid method combining Genetic Algorithm（GA）and a Support Vector Machine (SVM) was proposed,and the model for degradation state recognition of rolling bearing was constructed. Firstly the feature vectors of degradation state were extracted through the combination of GA and SVM from statistical characteristic. Then the degradation state probability distribution and historical remn ant life of rolling bearing are calculated to deter mine the optimal number of degradation state, whi ch is employed to construct the SVM model for deg radation state recognition. Finally extracted the characteristic vectors which have been optimized and deleted by GA from the test data of different degradation states, and then using the character ristic vectors as the input of SVM which parame ters has been optimized by GA to identify the degradation state of rolling bearing.The analytical results for full lifetime datasets of a certain bearing demonstrate the validity of the method.

Keywords—genetic algorithm(GA);support vector machine (SVM); rolling bearing;fault; degradation state

I. INTRODUCTION

The core idea of PHM is that acquire all kinds of data information with as little as possible sensors firstly, then with the aid of intelligent reasoning algorithm to assess the health status of system itself, and forecast the fault of machine. Failure prediction includes three important steps: feature extraction, degenerate state recognition and the forecast of residual life or failure probability. Among them, the degradation state recognition is the key technology of fault prediction, it direct relationship to the accuracy of fault prediction and the accuracy of the residual life predi ction[1-2].

Rolling bearing is an important compon ent of rotating machinery, its condition directly affects the run of rotor and even the whole machine[3-4]. The reliable identify of degrada tion state of rolling bearing mainly includes two key problems: one is establish a correct model of degradation state; the other is select a suitable degradation status indicators. To solve above two problems, scholars has carried out a preliminary study on the performance degrada tion recognition of rolling bearing. Hack-Eun [5] presented a method that using SVM to estima tion the state probability of equipment and applied the experimental data verified it. Vachts evanos [6] has studied the fault diagnosis using the method of dynamic wavelet neural network.

The main problem of above method is the number of degradation condition of rolling bearing need artificial setting according to our experience.So,this article presents a method of discretization and identification of degradation state automatically based on GA and SVM. Using the method of health state probability evaluation for degradation state discretization of rolling bearing from the signal of fatigue experiments of whole life cycle firstly, determine the optimal number of degradation state automatically.Then the method of genetic algorithm (GA) is used to optimize parameters of SVM, the model of degradation state recognition of rolling bearing then will be set up to recognition the degradation state.

II. DEGRADATION STATE FEATURE SELECTION BASED ON GA

The statistical characteristics of vibration signal contains abundant information.When the degradation of bearing in a different state,the statistical characteristics of the signal will corresponding changes.

In the classification method of intellig ence, sometimes because of the redund ant or irrelevant features covered the main classifica tion characteristics, then will lead to"dimension disaster". To eliminate these redundant or irr elevant features, feature selection is required,it can reduce the dimens ion of feature space and improve the accuracy of classification.The fitness function of GA as a evaluation standard for the performance of feature subset,will search the effective feature subset. This article introduced the algorithm of [7], construct the fitness function of Gene tic algorithm to feature optim ization accord ing to the recognition rate of degradation state model of SVM in the follow ing section.

The fitness function for feature extract ion is as follows:

$$f(x_i)=\frac{\exp(A_{X_i}-\eta\cdot A)}{\gamma\cdot(\sum_{j=1}^{n}x_j)/n} \tag{1}$$

Where $X_i(x_1,x_2\cdots x_j\cdots x_n)$ reflects a possible feature subsets, then $f(X_i)$ is the corresponding fitness function value of X_i, n is the total number of features, A_{X_i} is the corresponding training accuracy of X_i, A is the accuracy of including all the training characteristics.The two parameters are used to adjust the tolerate of threshold value of classification accuracy and the contribution weight of characteristics reduce to the fitness function.

We can search the feature subset which have the highest fitness by the algorithm, it can improve the characteristic optimization effect of degradation condition greatly.

III. DEGRADATION STATE RECOGNITION MODEL

A. The Model of Regression Classification Based on SVM

SVM is a kind of machine learning algorithm which is usually use to solve the problem of classification and prediction of small sample[8-9]. Given target sample set{(x1, y1), (x2, y2), …, (x_i, y_i) }, $x_i \in R^N$, N is the dimensions of each training sample.Suppose that there are k classes, y_i={1,…,k}, then the model have k(k-1)/2 binary classifiers. For training data from the ith and the jth classes,SVM solves the folling classification problem:

$$\min \ \frac{1}{2}\left(w^{ij}\right)^T\left(w^{ij}\right)+C\sum_{i=1}^{n}\xi^{ij}\left(w^{ij}\right)^T \qquad (2)$$

$$\text{s.t.} \ \left(w^{ij}\right)^T\phi\,(x_t)+b^{ij}\geq 1-\xi_{ij},\ \text{if } y_t=i,$$

$$\left(w^{ij}\right)^T\phi\,(x_t)+b^{ij}\geq -1+\xi_{ij},\ \text{if } y_t=j,$$

$$\xi^{ij}\geq 0$$

Where ξ^{ij} is the relaxation factor, C is the penalty factor, w^{ij} is the coefficient vector, $\phi\,(x_t)$ is the kernel function.

The parameter g control the flexibility of RBF k ernel function, the coefficient c plays a role of balan ce the complexity of the decision function and the n umber of sample error[10],These two parameters of su pport vector machine (SVM) are usually rely on exp erience or artificial test in practical application.

For the selection of the optimization parameters c and g of the SVM, the common method is using K—fold cross validation (K—CV) at present, the pri nciple of it is divide the original data into K groups (usually is divide equally) firstly, each subset will a s a test set respectively, at the same time the rest of the K-1 set of subsets of the data will as a training set.Then we will get K models,and acquire the best parameters c and g which make the test sets have t he highest classification accuracy.But if we want to a cquire the best parameters c and g in a larger range ,it will be very time consuming.Nevertheless, using heuristic algorithm—genetic algorithm (GA) on the b asis of the K—CV don't need to traverse all the par ameters points can also find the global optimal soluti on.

Because the classification accuracy using K—C V reflected the pros and cons of the performance of classifier, so using K—fold cross validation (K—CV) accuracy as the fitness function value of genetic alg orithm (GA) to find the best parameter c and g whi ch make the classification effect of SVM better in th is paper, The Group number K usually will take 5 o r 10. According to the existing research,K=5 will obt ain good classification effect[11].

B. The Discretization of Degra dation State

Traditional state monitoring and fault diagnosis u sually divided the health status of bearing into two k inds—normal and fault simply.In fact，most of the be aring will experience a series of degradation state be fore the final function failure.The appropriate select o f degradation state number has important influence o n the accuracy of degradation state recognition. Thro ugh the residual lifetime error of health state,the met hod of health state probability evaluation can determi ne the optimal number of degradation state, this met hod can effective discretization the degradation state of bearing. The degradation state is divided into 9 ki nds state which are 2 to 10 firstly, using the model in 2.1 respectively to training and forecasting the sa mples, and then calculate the training and predicting accuracy of each case, Finally select the optimal resu lt to determine the final number of degradation state. Specific steps are shown as follows:

(1)The initial sample data is $\vec{x}_t=(x_{t1,}x_{t2,\ldots,}x_{tm})$,m is the data length,t is the time index, y={1,…,k}， k=2,3…, 10, the initial value of k is 2.

(2) Calculating the probability distribution of each state according to the smooth window and instruction function.

$$\text{Prob}(S_t=i\,|\,\vec{x}_t,\ldots,\vec{x}_{t+u-1})=\sum_{j=t}^{t+u-1}I_i(y_j)/u \qquad (3)$$

$$I_i(y)=\begin{cases}0 & y\neq i\\ 1 & y=i\end{cases};$$

Where S_t is the smooth health status, u is the width of the smooth window function.

The sum of probability of each health state is show as follows

$$\sum_{i=1}^{k}\text{Prob}(S_t=i\,|\,\vec{x}_t,\ldots,\vec{x}_{t+u-1})=1 \qquad (4)$$

(3)After the probability of each health state is obtain ed, then we will get the residual life of current state according to the history information of residual life of each state.

$$\mathrm{RUL}(T_t)=\sum_{i=1}^{k}\mathrm{Prob}(S_t=i\mid \vec{x}_t,\ldots,\vec{x}_{t+u-1})\cdot\tau_i \qquad (5)$$

Where S_t is the current probability of each degradation state, τ_i is the residual life of history at state i, m is the number of degradation state.

(4) Calculating the average training and predicting ac curacy according to the following formula.

$$A=(\frac{\sum_{i=1}^{N}|\mu_i'-\mu_i|}{N})\times 100\% \qquad (6)$$

(5)Where N is the number of data. μ_i' is the actual residual life, When μ_i is the assessment residual life of training,then A is the average training accuracy, when μ_i is the assessment residual life of test,then A is the average test accuracy,

(6) If k = 10, steering (5), or k = k + 1, jump to (2). Contrast the status which have researched by drawing, finding the optimal number of degradation state.

IV. THE DEGRADATION STATE RECO GNITION OF ROLL ING BEARING

To verify the effectiveness of the method which have been proposed,we selection the vibration data of rolling bearing from the NSF I/UCR Center for Intelligent Maintenance Systems (IMS).Each data consists of 20,480 points with the sampling rate set at 20 kHz.The model of rolling bearing is Rexnord ZA-2115,and rotation speed was kept constant at 2000 RPM by an AC motor coupled to the shaft via rub belts,all bearings are force lubricated.

Because of the statistical characteristics of vibration signal contains abundant information, so statistical characteristics are used to recognition degraded state of rolling bearing in this paper.Using the reconstruction signal through the combination of GA and SVM to optimize the characteristics of degra dation state.The input feature vectors of SVM contains 16 statistical characteristics, incl uding seven time-domain dimensional char acte ristics, four time-domain dimensionle ss features and five frequency domain characteristics.

TABLE I TATISTICAL CHARACTERISTICS

Time domain features		Frequency domain features
dimensional	***dimensional ess***	
1. Mean 2.Var 3.Root 4.RMS 5.Peak 6.Ske 7.Kur	8.S factor 9.C factor 10.I factor 11.L factor	12. MSF 13. FC 14.VF 15. RMSF 16. RVF

Using the 16 features which has been extracted to reduce the dimension.The parameter setting of GA is as follows: the number of population and maximu m evolution generation are 60 and 300 respectively, The value of P_c is 0.65, P_m is 0.005 ,the paramete rs of fitness function η and γ are 0.91 and 0.6 resp ectively.The result of feature selection is showed by 0 or 1. After the dimensionality reduction,six features which have best classification results are shown in t able 2, they are variance value, Root amplitude, the value of the root mean square, mean square frequenc y, center frequency, and frequency variance. Then the feature vector T which has been acquired is used f or identify the degradation state of rolling bearing.

$$T=[\mathrm{Var},\mathrm{Root},\mathrm{RMS},\mathrm{MSF},\mathrm{FC},\mathrm{VF}] \qquad (7)$$

The accuracy of K-CV is used as the fitness val ue of genetic algorithm to optimizing the parameters c and g of SVM technology. The parameters popula tion and the largest evolution generation is 20 and 2 00 respectively, P_c and P_m also take 0.65 and 0.00 5 respectively. As shown in Fig .6, the final results of SVM parameters optimization are C = 1.3785, g = 11.0259.

According to the features of the support vector machine (SVM) which has already been acquired, an d the theory which has been given in Ⅲ，we can de termine the specific number of the degradation state based on the different status of training and predictio n accuracy which has been analyzed.As shown in Fi g .1, different degradation state represents the process of bearing fault effectively， according to physical c hange of degradation process, despite the state of hig h number have higher training accuracy, but with a l ower prediction accuracy. On the contrary, the state of low number has higher prediction accuracy, but w ith a lower training accuracy.Beyond five health state s the prediction accuracy decreased rapidly and witho ut significant increased in the training accuracy value s.So the number of degradation state is 5.

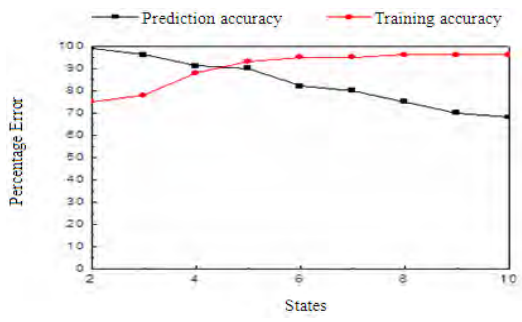

Figure 1. The training and prediction accuracy of different state

Using the feature vector T which has been dim ension reduction to recognition the degradation state of rolling bearing.The data are divided into 5 groups, and each group has 60 samples, including 30 sample s for training the model and other 30 samples for th e recognition test of degradation state.

Selection the test data which has been acquired to identification the degradation state after the SVM model of degradation state has been trained. The division of state categories and recognition accuracy of degradation state are shown in Table 2.

TABLE II THE RECOGNITON ACCURACY OF SVM

class	Training/test samples	Sample points	Accuracy (%)
1	30	20480	96.6667
2	30	20480	93.3333
3	30	20480	96.6667
4	30	20480	96.6667
5	30	20480	100

The recognition result of the test sample is shown in Fig .2.From the figure we can see that the 1st，3th and 4th class status have only one sample data was wrong points, the 2th class status has two sample data were in the wrong points, and the fifth class status of the sample data in identifying when no fault point. The graph shows that the SVMhas a high classify accuracy to the test sample, and the SVM technology which has beenoptimized by genetic algorithm is effective for the degradation state recognition of rolling bearing.

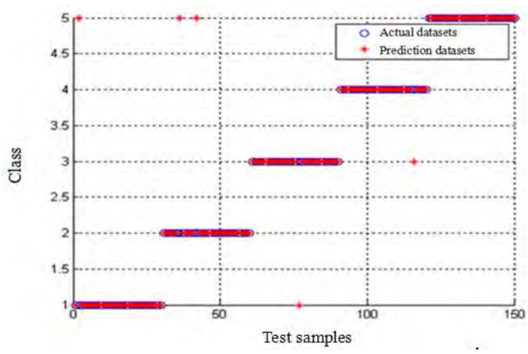

Figure 2. The comparison chart of classification result

To validate the advantages of GA in feature extraction,comparing GA with other traditional commonly method—principal component analysis (PCA) method and particle swarm Optimization (PSO).The contrast analysis of classification results of feature extraction are shown inTable 3.

TABLE III THE RECOGNITION ACCURACY OF THREE FEATURE SELECTION METHODS

Manner	Accuracy (%)	Correct samples/Total samples	Feature number
GA	96.6667	145/150	6
PCA	86.6667	130/150	8
PSO	82.6667	124/150	9

Table 3 shows the result that GA has significantly advantage in terms of feature extraction relative to other two methods,and it has lower dimension and high accuracy.This is because the technology of GA eliminate the initial characteristics which has nothing to do with the recognition of degradation state comparing with other two kinds of methods,making the classification accuracy become better.

V. CONCLUSION

The accurate assessment of degradation state of current equipment is the basic of implementation status maintenance and predictive maintenance. This paper main studies the following conclusions:

(1)GA could extract the sensitive feature which reflect the degradation state of rolling bearing realize the optimal selection of the characteristics,and it can also realize the parameters optimization of SVM.

(2)The optimal number of degradation state can be determined through the combinationof health state probability evaluation method and SVM.

(3)The method of degradation state recognition based on the GA and SVM can identify the degradation state of mechanical system accurately.

Through the actual data validation shows that the effect of proposed technology is remarkable to the degradation state recognition of rolling bearing.

ACKNOWLEDGMENT

This work is supported by National High Technology Research and Development Prog ram Foundation of China under Grant No.20 12AA041303 and Outstanding Young Scholars Growth Plan in University of Liaoning Pro vince under Grant No.LJQ2012019.

REFERENCES

[1]G.Niu and B. S.Yang,“Intelligent conditi on monitoring and prognostics system based on data-fusion strate gy,”Expert Systems with Applications,Vol. 37,Dec. 2010, pp.8831 -8840.

[2]H.T.Pham and B.S.Yang, “Estimation and forecasting of machine health condition using ARMA/GARCH MODEL,”Mechanical Systems and Signal Processs ing, Vol. 24,Feb. 2010, pp.546 -588.

[3]X.L.Zang,X.F.Chen and B.Li, “Review of life prediction for mechanical major equipments,”Journal of Mechanical Engineering, Vol. 47,Nov. 2011, pp.100-116 .

[4]A.Heng,S.Zhang,Tan,C.C.Andy and J.Mathew, “Rotating ma chinery prognos tic cs:state of the art,challenges an d opportu nities,”Mechanical Systems and Signal Proces s ing, Vol. 23,Mar. 2009, pp.724-739.

[5]K.Hack,C.C.Andy,Tan and J.Mathew, “Bearing fault prog nosis based on health state pr obability estimation,”Ex pert Systems with Applications,Vol. 39,Apr. 2012, pp. 5200-5213.

[6]P.Wang and G .Vachtsevanos,“Fault Prognosis using Dynamic Wavelet Neural Networks,” Proc. IEEE Symp. IEEE Systems Readiness Technology Conference, IEEE Press, Aug. 2001, pp.857-870,doi:10.1109/autest. 2001. 949467.

[7]Y.Liu,M.Li and H.Dong,“Research of fault feature opti mization based on genetic algorithm,”Computer Simu lation, Vol. 27, Apr. 2010, pp.92-96.

[8]V.N.Vapnik,“Statistical learning theory,”New York: Wil ey, 1998.

[9]J.F.Gao,W.G.Shi,J.X.Tan,F.J.Zhong,“Support vector mac hines based approach for fault diagnosis of valves in reciprocating pumps,”Proc. IEEE Symp.International Co nference on Electrical &Computer,IEEE Press, 2002,pp. 1622-1627, doi:10.1109/ccece.2002.1012999.

[10]S.S.Zang, “Research of the optimization of SVM param eter based on PFO,”Shenyang:Northestern University, 20 10.

[11]J.L.Yin and Y.L.Zhu,“Parameter optimizati on for support vector machine and its application to fault diagnosis of power transformers,”Electric al Measure ment & Instrumentation, Vol. 49,May 2012, pp.11-16.