

Application of LS-SVM in fault Diagnosis for Diesel Generator Set of Marine Power Station

CHEN Liwei, LIU Huan

College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

chenliwei@hrbeu.edu.cn, liuhuan@hrbeu.edu.cn

Abstract - In this paper support vector machine (SVM) algorithm being discussed, then discussed least squares support vector machine (LS-SVM) algorithm, at the same time, the applications of SVM in the fault diagnosis of turbocharger system for diesel generator set of marine power station being discussed, the least squares support vector machine algorithm being used in the research of fault diagnosis, being compared with BP neural network, experiments result show the operation speed of the least squares support vector machine algorithm is fast, its generalization ability is stronger, SVM can solve small sample learning problems as well as no-linear, high dimension and local minimization problems in the fault diagnosis of turbocharger system for diesel generator set of marine power station.

Index Terms - LS-SVM, fault diagnosis, marine power station.

I. Introduction

Marine power station is the most important component of the marine power station system; it is the core of the marine power system. It has great significance to ensure safe navigation for marines, it is important to achieve the reliability of marine power station operation. With the continuous improvement of automation and complexity of marine power station, it means the more faults that having impact on its normal work ability, it is not enough by using technical design and equipment monitoring to ensure safe and reliable operation of the marine power station.

When marine power station has faults and the system is fail, the experience and expertise of marine motor staff can play a decisive role for rapid diagnosis and troubleshooting. But as the marine power station update constantly and become complexity, when the marine power station has faults, it is fail to quickly identify and resolve as soon as possible only by members of the marine, which will cause power outages, downtime and even suspended the marine, and then bring a great deal of damage for the marine sailing security and economic benefits.

So an appropriate fault diagnosis method must be used to reduce the maintenance complexity of marine power station and cut the number of crew members. From the development trend of fault diagnosis, the knowledge-based intelligent fault diagnosis method is an important breakthrough in diagnostic technology in the future direction. Vapnik's statistical learning theory (SLT) and support vector machine (SVM) open up a new way to solve the small sample case of troubleshooting. Currently.

II. Support Vector Machines

A. Basic theory of Support Vector Machines

Support vector machines (SVM) is developed from the linearly separable optimal hyperplane, the basic idea is in case of two dimensional and two classes linear separable. The SVM algorithm derives from solving the problem of Binary classification problems, and for non-linear classifier, a kernel function is employed to map the data into a higher dimensional input space to achieve. Now support vector machine has been successfully applied to many practical problems such as handwriting recognition, face recognition, pattern classification and function estimation, regression analysis.

Support H is a classification line of two classes without error separate, H_1 , H_2 are lines parallel to classification lines and they are nearest to classification lines, the distance between them is called class interval, showed in figure 1.

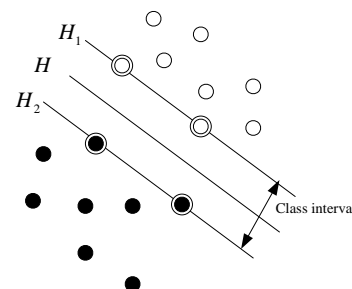


Fig 1 Optimal hyperplane

The so-called maximum classification interval line is that it is not only able to class correctly but also able to gain the largest class interval. The former is to ensure the empirical risk be minimizing (it is zero), the later can make the confidence range be smallest, so as to make the real risk minimization. Extended to higher dimensional space, the optimal classification line becomes optimal classification hyperplane^[1].

For the given training sets $\{x_k, y_k\}$, it has N samples, $x_k \in \mathbb{R}^D$ are the input, $y_k \in \{-1, 1\}$ are the output. The general form of linear discriminate function in D dimensional linear space is

$$g(x) = \omega \cdot x + b \quad (1)$$

The function of classification surface is $\omega \cdot x + b = 0$,

normalized the discriminate function, make the two class samples to meet $|g(x)| \geq 1$, that is the $|g(x)|$ is equal to 1 for all samples that nearest to the optimal classification hyperplane.

So the maximum interval is equivalent to make $\|\omega\|$ (or $\|\omega\|^2$) minimum, if

$$\gamma_i[(\omega \cdot x_i) + b] - 1 \geq 0, i = 1, 2, \dots, n \quad (2)$$

the optimal classification hyperplane is correct to all samples.

In condition of nonlinear and inseparable, introduce ξ_i , the optimal question of standard support vector machines change to

$$\min_{\omega, b, \xi} J(\omega, b, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_i \quad (3)$$

$$\text{s.t. } y_i[\omega \phi(x_i) + b] \geq 1 - \xi_i, i = 1, 2, \dots, n$$

$$\xi_i \geq 0, i = 1, 2, \dots, n$$

C is error penalty coefficient, $\phi()$ is the function of space changing.

B. Least Squares Support Vector Machine

LS-SVM introduces least squares linear system into SVM, which is put forward by Suykens in 1999, LS-SVM is solving the following convex optimization problem^[2,3].

$$\min J_{LS}(w, b, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^l e_i^2 \quad (4)$$

$$\text{s.t. } y_i(w^T \phi(x_i) + b) = 1 - e_i, i = 1, 2, \dots, l \quad (5)$$

Definite Lagrange function:

$$L(w, b, e; \alpha) = J_{LS} - \sum_{i=1}^l \alpha_i \{y_i[w^T \phi(x_i) + b] - 1 + e_i\} \quad (6)$$

α_i is Lagrange multiplier, it can be positive or negative.

According to Karush-Kuhn-Tucker (KKT) condition:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^l \alpha_i y_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \Rightarrow \alpha_i = \gamma e_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \Rightarrow y_i(w^T \phi(x_i) + b) - 1 + e_i = 0 \end{cases} \quad (7)$$

The matrix equation can be obtained easily as follows:

$$\begin{bmatrix} I & 0 & 0 & -Z^T \\ 0 & 0 & 0 & -Y^T \\ 0 & 0 & \gamma I & -I \\ Z & Y & I & 0 \end{bmatrix} \begin{bmatrix} w \\ b \\ e \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \bar{1} \end{bmatrix} \quad (8)$$

In which

$$Z = [\phi(x_1)^T y_1, \dots, \phi(x_l)^T y_l]$$

$$Y = [y_1, \dots, y_l]$$

$$e = [e_1, \dots, e_l]$$

$$\alpha = [\alpha_1, \dots, \alpha_l]$$

$$\bar{1} = [1, \dots, 1]$$

Ignore w, e can obtain

$$\begin{bmatrix} 0 & Y^T \\ Y & Z \cdot Z^T + \frac{1}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{1} \end{bmatrix} \quad (9)$$

According to Mercer condition,

$$\begin{bmatrix} 0 & y_1 & \dots & y_l \\ y_1 & y_1 y_1 K(x_1, x_1) + \frac{1}{\gamma} & \dots & y_1 y_l K(x_1, x_l) \\ \vdots & \vdots & \dots & \vdots \\ y_l & y_l y_1 K(x_l, x_1) & \dots & y_l y_l K(x_l, x_l) + \frac{1}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_l \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad (10)$$

Analyze equation (10), if the training sample set contains l samples, then equation (10) contains $l+1$ unknown numbers, $l+1$ equations and $(l+1)^2$ multipliers. We can see, the number of sample points that determine the coefficient matrix of equation (10).

Thus the classification problem can be solved by equation (10), at the same time avoiding quadratic programming problem. When the coefficient matrix is full rank, KKT system is a square system, and it has unique solution.

Eventually obtain least square support vector machine classifier:

$$y(x) = \text{sign}[\sum \alpha_i y_i K(x, x_i) + b], \quad i = 1, 2, \dots, l \quad (11)$$

α, b are the solution of equation (10), $K(\bullet)$ is kernel function.

III. Fault mechanism analyses for diesel generator set of marine power station

The diesel generator set of marine power station is a multi-system, multi-level complex nonlinear system; it is composed of the original motivation, generators and ancillary equipment. The original motivation is generally marine diesel generator; the diesel generator system can be divided into turbocharger system, lubrication system, the cylinder piston, combustion system, fuel system, cooling system, transmission and steering systems and other subsystems. The fault of subsystem will cause performance degradation of the marine power station, fault diagnosis hierarchy shown in Figure 2^[4-6].

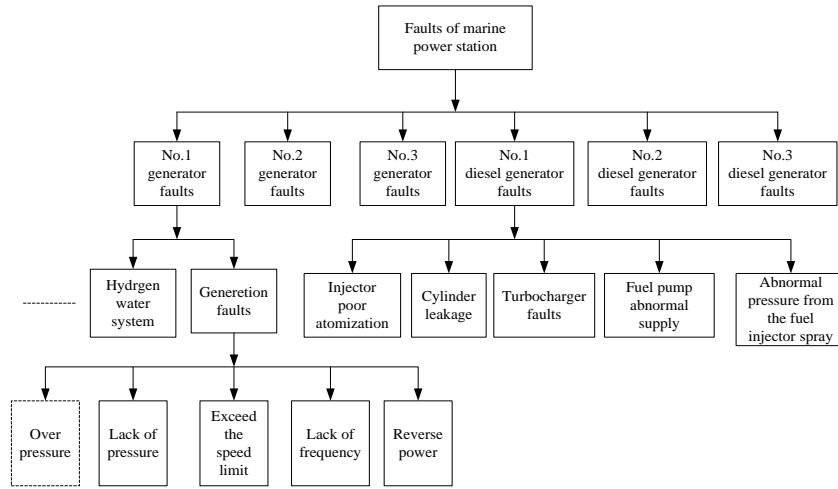


Fig 2 Fault diagnosis hierarchy

IV. The fault diagnosis for turbocharger system of diesel generator

According to the working process of theoretical analysis and practical operating experience for low speed diesel generator, the cause and location of turbocharger system can be determined, they can be regarded as fault variables or output variables, and determine the sign variables as the SVM input variables^[7].

A. Output variables

Output variables are as follows: normal conditions (no fault) F_1 , the efficiency of turbocharger reduce (compressor, turbine or the mechanical efficiency decrease) F_2 , air-cooled heat transfer deterioration F_3 , turbine protection grille block F_4 , turbine flow section plug F_5 , air filter block F_6 , air side flow resistance increase F_7 and the flow resistance of exhaust heat boiler increase F_8 , the total variables are eight. The change range of variables F_1, F_2, \dots, F_8 is from 0 to 1, 0 means without such fault, 1 means the fault is seriously (except normal conditions). The output variables of SVM are $F_1, F_2, \dots, F_5, F_6, F_7, F_8$ can be diagnosed directly from the component output.

B. Input variables

The input variables are: the exhaust manifold temperature T_r , sweep gas tank pressure P_s , the largest average burning pressure of every cylinder P_{max} , turbocharger speed n_{tc} , pressure loss coefficient of sweeping airway, compressor export temperature T_c , sweep gas trunk temperature T_s , pressure loss coefficient of filter net, pressure loss coefficient of air cooler, pressure loss coefficient of exhaust gas boiler, the total variables are ten. Plus a parameter

that reflecting the diesel generator loads, the total variables are eleven. Some of these parameters are either directly measured or derived from the measured parameters. The parameters that being measured by sensors or by the real-time electronic dynamometer need to obtain the deviation of the corresponding reference value (calculated by the simulation), divided by the absolute value of the maximum deviation to obtain the input variables, that is normalized to $[-1, 1]$, "+" means increase, "-" means decreased.

On the basis of the principles of the internal combustion engine and guidance requirements of marine diesel engines, if the upper and lower deviation of the operating parameters of turbocharger diesel generator excess the following data, the fault is exist:

(1) Cylinder exhausts temperature: 30K (absolute temperature); (2) sweep gas tank pressure: 0.06MPa; (3) maximum explosion pressure: 3.0MPa; (4) turbocharger rotational speed: 1500r/min; (5) filter net pressure loss coefficient: 0.10; (6) air-cooled pressure loss coefficient: 0.10; (7) waste heat boiler pressure loss coefficient: 0.10; (8) sweep exhaust pressure loss coefficient: 0.06; (9) compressor output temperature: 30K; (10) sweep gas tank temperature: 40K.

One fault of turbocharger system corresponds to one sample, for further diagnose the severity of a fault, and take two samples from each fault variable; the target value is 0.5 and 1.

In order to reflect the correspondence of fault and the symptoms of generating unit operating load range, give samples of four kinds of conditions: the rated load (100% MCR), part load (90% MCR, 75% MCR) and half-load (50% MCR).

Taking into account the range of ocean marines are unrestricted navigation area, so the air temperature is divided into three sections, namely 283K-294K, 294K-306K, 306K-318K. the sample center is 288K (15 °C), 300K (27 °C), 312K (39 °C), through a large number of simulation calculations, can derive the corresponding samples (total sample size of 108), these samples can be used to train SVM

diagnostic model, so as to realize the fault diagnosis of marine diesel generator.

V. Fault diagnosis experiments

The main work of training stage is to select appropriate category parameters based on the sample set, where the parameters include the kernel function parameters and the penalty factor C . First, choose Gaussian function $K(x, x') = \exp(-\frac{\|x - x'\|^2}{2\sigma^2})$ as mapping kernel function, when testing, the training samples being regarded as testing samples, the accuracy rate can obtain more than 95%. The categories are not wrong, then classify the test samples, the accuracy rate can obtain more than 85%. Using the data acquisition and recording in front, after preprocessing, regard them as normal operating condition's sample data of the marine diesel generator. According to the regression theory, select $\sigma^2 = \frac{1}{2}, C = 200, \varepsilon = 0.001$.

Take cylinder exhausts temperature T_r , sweep gas tank pressure P_s , maximum explosion pressure P_{\max} , turbocharger rotational speed n_{tc} , sweep gas tank temperature and compressor output temperature T_c as the input variables of SVM, take filter net pressure loss coefficient, air-cooled pressure loss coefficient, waste heat boiler pressure loss coefficient as component feature parameters, diagnose directly.

The severity of the fault can be divided into three grade, namely, fault grade I, fault grade II and no-fault (normal conditions). According to the output vector of SVM, the results of treatment are as follows:

- if $0.75 < F < 1.50$, it is grade I (major fault)
- if $0.25 < F < 0.75$, it is grade II (moderate fault)
- if $F < 0.25$ or $F > 1.5$, it is normal (no fault)

Thus can get diagnosis results.

For data comparison, the SVM and BP neural network simulation results are compared, current ly BP neural network has more applications in fault diagnosis^[5]. The experiment results are shown in Table 1.

Table 1 simulation and prediction results comparison of SVM and BP neural network

error		SVM		BP neural network	
		fitting data	prediction data	fitting data	prediction data
absolute error	maximum	0.860	0.235	0.782	0.513
	minimum	0.001	0.004	0.015	0.012
	average	0.264	0.108	0.183	0.210
relative error	maximum	1.65	0.43	1.5	0.7
	minimum	0	0.01	0	0.1
	average	0.56	0.21	0.42	0.38

The results of SVM and BP neural network are analyzed and compared. SVM method can not only make the error as small as possible but also make the regression function as smooth as possible, thus improve its generalization (forecasting) capabilities. At the same time the structure of artificial neural network is optimal structure that being get by a large number of test, its training time is very long, it takes about 38 minutes to get predictable results; while SVM method requires only small samples to calculate, it takes about 4s to get an acceptable predictive value. So in terms of computing speed or generalization ability, SVM methods have shown a good performance.

VI. Conclusion

SVM being used in the turbocharger system modeling of marine diesel generator, simulation results show the learning efficiency of SVM and its ability in solving over-fitting are perfect. SVM is based on small sample statistical theory, and it has rigorous theoretical system, so there will certainly be very broad application prospects in the field of non-linear modeling.

Acknowledgement

This paper is funded by the Fundamental Research Funds for the Central Universities ((HEUCF130811)) and and the National 863 program (P081512005)

References

- [1] WU Hong xing, PENG yu, PENG Xi yuan. SVM Multi-classifier Design for Analogous Circuits Fault Diagnosis. Journal of electronic measurement and instrument. 2007, 21(4):27-31
- [2] DAI Li xiong, MENG Xian rao. New method of fuzzy fault diagnosis of marine diesel engine. MARINE ENGINEERING. 2009,31(1):56-59
- [3] ZHAO Chun hui, QIAO Lei. Classiifcation of hyperspectral remote sensing image using improved LS-SVM. APPLIED SCIENCE AND TECHNOLOGY. 2008, 35(1):44-47.
- [4] WAN Hui, WEI yan. An improved least squares support vector machine algorithm. JOURNAL OF CHONG QING NORMAL UNIVERSITY. 2010,27(4):69-73.
- [5] Wu Zhiliang Guo Chen. Study on Reliability and Maintainability for Marine Power Station Based on Fault Tree. JOURNAL OF WU HAN UNIVERSITY OF TECHNOLOGY. 2010, 34(1):60-63.
- [6] NIU Hong yu. Fault Diagnosis of the Marine Diesel Generator Set Based on Artificial Neural Network. Lanzhou: Lanzhou University of Technology, 2007.
- [7] ZHAN Yulong, ZHAI Hailong. Feasible Research on Fault Diagnosis of Marine Diesel Engine Based on Support Vector Machine. Navigation of China.2007,2(71):89-92.