

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

APSO-RVM for Fault Detection of Liquid Rocket Engines Test-bed

Fengyu Zhu, Qi Wang and Zhengguang Shen
Harbin Institute of Technology, Harbin, Heilongjiang 150001, China

Abstract: Selection of Relevance Vector Machine (RVM) kernel function parameter is one among ineffectively resolved issues which is first resolved in the literature by Adaptive Particle Swarm Optimization (APSO). A novel APSO-RVM method is proposed to optimize and select the RVM kernel parameter, thus forming, taking the advantage of APSO dramatically convergence. Furthermore, the method is applied to the fault detection of liquid rocket engines test-bed. In order to verify the validity of dramatically effectiveness in fault detection, this paper demonstrates the proposed APSO-RVM approach by performing both simulations and experiments using Oxygen Valve Outlet Pressure (Pejy) data. Results show that APSO-RVM can rapidly detect faults effectively and has a high practical value.

Key words: Relevance vector machine, particle swarm optimization, fault detection, liquid rocket engines

INTRODUCTION

Liquid rocket engines experiment is a extremely sophisticated integrated technology and huge systematic project. Compared with other large scientific experiments, it has particularly high comprehensiveness encompassing large numbers of subjects such as rocket engine expertise, monitoring and control technology, propellant chemistry technology, cryogenic technology, vacuum technology, high-altitude environment simulation technology, environmental monitoring technology, environmental technology and organization and management science (Figuroa and Schnalzel, 2006). Because of complex structure, high precision and hefty expense, experiment failure can lead to huge losses and even equipment damage. According to the fault records form Aerospace Testing Technology Institute, the failure rate due to experiment system fault was 20%. The early detection of faults can help avoid failure from spreading, reduce system shutdown and prevent accident involving human fatality and material damage (Wu, 2005). Therefore, it has great economic and security significance for early fault detection and timely preventive maintenance of the liquid rocket engine test-bed.

Our liquid rocket engines test-bed still adopts off line monitoring method in the present in China which merely forces on several main parameters, while may result in fallacious prediction even erroneous failure detection relating to the sensors fault (Wu, 2005). Many widely employed methods based on statistics are simple and reliable, however, whose accuracy rely heavily on statistical results of test-bed. And it is difficult to establish accurate mathematical model by mathematical

analysis on account of nonlinear and instability of the liquid rocket engine test-bed. In recent years, as computer technology and artificial intelligence constantly developed, many new theories and methods in the field of fault detection and diagnosis are applied to propulsion system. Fuzzy hyper-sphere neural network based real-time monitoring failure system is presented (Huang *et al.*, 1999) which is verified that it outperforms the BP neural network in the article. Neural network based real-time fault detection algorithm in ground test process is implemented with MATLAB (Huang *et al.*, 2007) which performs effectiveness in engine failure detection through many offline evaluation and real-time online test. A Support Vector Machines (SVM) based multi-fault classifier is established for data mining of liquid rocket engine steady state test (Han and Hu, 2007) which is verified by experiment that the algorithm has excellent classification and anti-jamming ability and requires few training data.

To deal with the defects of difficultly determined neural network structure and easily immersing in partial minimum frequently, a novel adaptive particle swarm optimization relevance vector machine (APSO-RVM) is proposed which is based on liquid rocket engines test-bed fault detection method. Through, mapping the raw data to high dimension space with kernel function, it solves nonlinear in the high dimension space to avoid the lower linear inseparable problem. It not only resolves over learning problems in virtue of small samples and nonlinear but also generates better generalization ability by solving sparse model.

The main contributions in this study are the development of the following:

- A novel RVM kernel parameter selection method is proposed, reflecting faster computing speeds and locating the global optimum
- A novel liquid rocket engines test-bed fault detection method was suggested, reducing the alarm time and promoting the system reliability

THE RELEVANCE VECTOR MACHINE

RVM, like SVM, has a sparse probabilistic model (Tipping, 2001). However, RVM appears to be advantageous against SVM due to Bayesian treatment which does not suffer from the basic limitations of SVM. The RVM trains in the Bayesian framework, obtains the priori probability of weight by a set of hyper-parameters and find the optimal value via iterative algorithm. RVM applied Bayesian inference based on Gaussian process methods to SVM, derives probability distribution and makes kernel function free from Mercer. It has good performance (Tipping and Lawrence, 2005) on function regression and classification.

For regression, given a set of input vector $\{x_n\}_{n=1}^N$ and relevant output vector $\{y_n\}_{n=1}^N$, supervised learning aims at designing a model with those training data and prior knowledge (Tipping and Faul, 2003). Based on new input vector x_n , the model can forecasted relevant output vector $y(x_n)$. t_n are observed output which can be taken as a unknown function $y(x, w)$ contained the Gaussian noise with σ^2 variance:

$$t = y(x, w) + \epsilon \tag{1}$$

where, ϵ is independently distributed noise; w is adjustable parameters weight. So, obtain:

$$y(x, w) = \sum_{i=1}^M w_i \Phi_i(x) = w^T \Phi(x) \tag{2}$$

where, $\Phi_i(x) = K(x, x_i)$. The selection of kernel function will be free from Mercer. Thus, we can choose the most popular kernels such as Gaussian kernel, polynomial kernel, Radial Basis Function (RBF) kernel, etc. The likelihood function corresponding training set is :

$$p(t | w, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\{-\frac{1}{2\sigma^2} \|t - \Phi w\|^2\} \tag{3}$$

Where:

$$t = (t_1, \dots, t_N)^T, w = (w_0, \dots, w_N)^T$$

and:

$$\Phi = [\phi(x_1), \phi(x_2), \dots, \phi(x_N)]^T$$

that:

$$\phi(x_n) = [1, K(x_n, x_1), K(x_n, x_2), \dots, K(x_n, x_N)]^T$$

The way got the optimal weight w of Empirical Risk Minimization (ERM) can result in over learned. To avoid this situation, given a priori conditional probability distribution using sparse Bayesian method for the weight w :

$$p(w | \alpha) = \prod_{i=0}^N N(w_i | 0, \alpha_i^{-1}) \tag{4}$$

Meanwhile, for hyper-parameters α and noise variance $\beta = \sigma^2$ definite hyper-prior distribution as Γ distribution:

$$p(\alpha) = \prod_{i=0}^N \text{Gamma}(\alpha_i | a, b), \quad p(\beta) = \text{Gamma}(\beta | c, d) \tag{5}$$

Where:

$$\text{Gamma}(\alpha | a, b) = \Gamma(a)^{-1} b^a \alpha^{a-1}, \quad \Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$$

In normal circumstances, parameter a, b, c, d is very tiny, it can provide $a = b = c = d = 0$. So, it will get consistent hyper-prior.

Then, according to the Bayes rule, it can find posteriori formula for all of unknown parameters as following:

$$p(w, \alpha, \sigma^2 | t) = \frac{p(t | w, \alpha, \sigma^2) p(w, \alpha, \sigma^2)}{p(t)} \tag{6}$$

Given to new observed point x_* , the distribution of corresponding target forecasted value t_* is:

$$p(t_* | t) = \int p(t_* | w, \alpha, \sigma^2) p(w, \alpha, \sigma^2 | t) dw d\alpha d\sigma^2 \tag{7}$$

Considering:

$$p(w, \alpha, \sigma^2) = p(w | t, \alpha, \sigma^2) p(\alpha, \sigma^2 | t) \tag{8}$$

Thus:

$$p(w | t, \alpha, \sigma^2) = \frac{p(t | w, \alpha, \sigma^2) p(w | \alpha)}{p(t | \alpha, \sigma^2)} \tag{9}$$

$$= (2\pi)^{-(N+1)/2} |\Sigma|^{-1/2} \exp\{-\frac{1}{2} (w - \mu)^T \Sigma^{-1} (w - \mu)\}$$

where, covariance matrix is $\Sigma = (\sigma^{-2} \Phi^T \Phi + A)^{-1}$, $A = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_N)$, mean vector is $\mu = \sigma^{-2} \Sigma \Phi^T t$.

This study use the peak of delta function to approximate the hyper-prior $p(\alpha, \sigma^2 | t)$ in the above

formula. Aim at forecasting, we do not care $p(\alpha, \sigma^2) \approx \delta(\alpha_{MP}, \sigma_{MP}^2)$, where α, σ^2 possible values are $\alpha_{MP}, \sigma_{MP}^2$. Focus on:

$$\int p(t|\alpha, \sigma^2) \delta(\alpha_{MP}, \sigma_{MP}^2) d\alpha, d\sigma^2 \approx \int p(t|\alpha, \sigma^2) \int p(\alpha, \sigma^2|t) d\alpha d\sigma^2 \quad (10)$$

Thus, RVM problem is converted into posterior mode for hyper-parameters problem, that is, seeking the maximum of posterior hyper-parameters which is equal to seeking the maximum of α and β . In the condition of consistent hyper-prior, it only needs to take the maximum of $p(t|\alpha, \sigma^2)$. So, obtain:

$$p(t|\alpha, \sigma^2) = \int p(t|w, \sigma^2) p(w|\alpha) dw \quad (11)$$

$$= (2\pi)^{-N/2} |\Sigma^T I + \Phi A^{-1} \Phi^T|^{-1/2} \exp\{-\frac{1}{2} t^T I + \Phi A^{-1} \Phi^T\}^{-1} t\}$$

To obtain $p(t|\alpha, \sigma^2)$, it need to integrate out α and β with following iterative formula:

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i^2} \quad (12)$$

$$(\sigma^2)^{new} = \frac{\|t - \phi\mu\|^2}{N - \sum \gamma_i}$$

where, $\gamma_i = 1 - \alpha_i N_{ii}$, N_{ii} are the i th diagonal elements of posteriori weight covariance matrix.

Ideally, if given an new input x_* , it can predict target through following steps:

$$p(t_*|t, \alpha_{MP}, \sigma_{MP}^2) = \int p(t_*|w, \sigma_{MP}^2) p(w|t, \alpha_{MP}, \sigma_{MP}^2) dw \quad (13)$$

$$p(t_*|t, \alpha_{MP}, \sigma_{MP}^2) = N(t_*|y_*, \sigma_*^2) \quad (14)$$

$$Y_* = \mu^T \phi(x_*) \quad (15)$$

$$\sigma_*^2 = \sigma_{MP}^2 + \phi(x_*)^T \Sigma \phi(x_*) \quad (16)$$

Corresponding to the new observations, the predicted output of RVM is $y(x_*, \mu)$.

SELECTION OF RVM KERNEL PARAMETER

Yet it has never been proposed any analytic instructional method for RVM kernel parameter selection (Li *et al.*, 2010; Tao *et al.*, 2008). At present, empirical method and ergodic searching have been mostly used to study kernel parameter setting. Ergodic searching has certain blindness for parameters setting and demands enormous sacrifices of time.

Characteristics analysis of RVM kernel parameter: In this study, based on standard function sinc, effects of kernel parameters on training results is discussed.

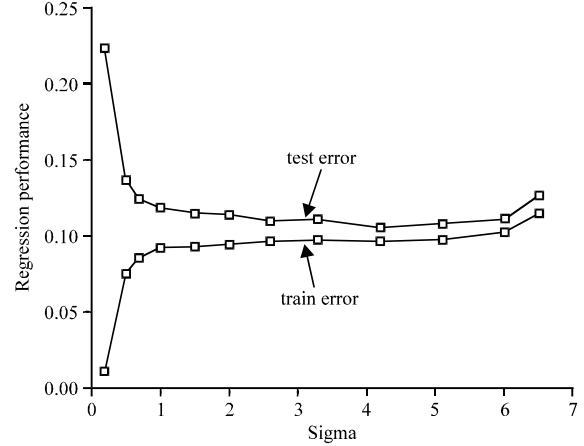


Fig. 1: Regression performance diagram

Table 1: Regression performance of different σ

σ	RV mean	Training error	Test error
0.2	71.2	0.0108	0.2225
0.5	22.0	0.0750	0.1364
0.7	16.9	0.0854	0.1238
1.0	11.5	0.0918	0.1181
1.5	8.9	0.0926	0.1143
2.0	7.9	0.0938	0.1137
2.6	6.9	0.0962	0.1095
3.3	6.8	0.0969	0.1105
4.2	7.0	0.0962	0.1049
5.1	7.0	0.0967	0.1077
6.0	8.2	0.1017	0.1110
6.5	10.2	0.1146	0.1258

The function of sinc is:

$$f(x) = \frac{\sin x}{x} + 0.1 * \text{noise}$$

where, noise is random noise in the range [-1, 1]. Even taken 100 sample points as samples, this paper adopted 5-fold cross-validation to calculate training error and test error. The Gaussian kernel used for the construction of the basis matrices is $K(x, y) = \exp[-\|x-y\|/\sigma^2]$.

From the Table 1, it can conclude that training error approximates zero as σ approaches zero which means all samples can be compact fitted. Whereas, we must notice that corresponding test error approaches infinity, the number of corresponding relevance vector is 100, that is to say, all the training samples are relevance vector, for test samples are 100. At this stage there is over learned and RVM lost learning ability. All above arguments show ERM based on traditional learning approach such as neural network cannot ensure excellent generalization ability. The sizes of relevance vector get smaller with the increase of σ (in a small range) and training error increasing but test error decreasing. That means the generalization ability of RVM is improving. When σ approximates 4.2, it can get minimum test error. Figure 1 illustrates minimum test error is not coincide with minimum

training error which verifies traditional ERM cannot guarantee excellent generalization ability. When σ reach a certain value (nearly 6 in the Table 1), the sizes of relevance vector and corresponding training error and test error are increasing again. That indicates the performance of discrimination and generalizations gets worse. Actually, the generalizations ability of RVM based on Gaussian kernel goes from low to high and to low.

Adaptive particle swarm optimization: After RVM kernel parameter characteristics analysis, a novel method about optimizing RVM kernel parameter with adaptive particle swarm optimization is proposed. Particle Swarm Optimization (PSO) based on swarm intelligence theory is an optimization technique for locating the global optimum (Tu *et al.*, 2011) which produces swarm intelligence to guide optimization with cooperation and competition among particles. This algorithm simulates birds foraging behavior, considers each individual always maintain the optimal distance with adjacent ones during activity. It promotes the co-evolution in the virtue of information sharing. Comparing with other intelligent optimization algorithms, PSO has an excellent performance on parallel searching and meanwhile it has few parameters and shape convergence speed. In PSO, each optimization problem is regarded as searching a flight particle in the space. Flight direction and distance of particle depend on velocity and optimal objective function determines fitness. Particles dynamically adjust their speed based on flying experience and find the optimal solution by iteration.

According to the above problem, defined PSO fitness function as:

$$\text{Fitness} = \sqrt{\frac{1}{n} \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2} \quad (17)$$

where, w_i is relevance vector weight, y_i is truth value, \hat{y}_i is predicted value.

The implementing steps of this optimization process are:

- **Step 1:** $V_{\max} = 1$, $V_{\min} = -1$, $c_1 = c_2 = 1.5$, iteration step is 100, population size is 20. Definite mutation operator as 0.9 on the basic PSO to prevent from immerging in partial minimum on account of iterative efficiency reduced in the post course of the optimization procedure
- **Step 2:** Random generate velocity and position of each particle
- **Step 3:** Calculate Fitness according to fitness function

Table 2: Optimization results of APSO

σ	RV	Fitness
4.7221	7	0.1054
4.7648	7	0.1046
4.5672	7	0.1062
4.3964	7	0.1008
4.8588	7	0.1050
4.5056	7	0.1062
4.8216	7	0.1051
4.4345	7	0.1026
4.8277	7	0.1044
4.4402	7	0.1041

- **Step 4:** Compare the current fitness value and the best fitness value in history for each particle. If the current fitness value is smaller, then take it as the best experienced individual fitness value
- **Step 5:** Calculate the minimum of the best fitness value of all particle experience and record it as the global best fitness value
- **Step 6:** Update particle velocity and position and limit each new velocity and position
- **Step 7:** Random generate mutation factor. If reach the variation conditions, re-initialize the particles
- **Step 8:** If reach the maximum of iteration step, stop searching and return the best fitness value and the position. Else turn to step 3

Following the above steps, 10 optimization results are give in Table 2.

From Table 2, the experimental results show that σ makes a nearly perfect fit with the optimal value at a high probability, meanwhile which prove consistent relevance vector number and close sparse level (7/100). From what has been discussed above, it may safely draw the conclusion that the proposed APSO is efficiently applicable to the problem of RVM kernel parameter optimization.

PROJECT EXAMPLE

Establishing mathematical model is the emphasis and difficulty of complex and nonlinear dynamic systems fault detection (Yu *et al.*, 2009). For a project of liquid rocket engines test-bed, hydrogen supply system is the key part of rocket engine test-bed, whose work status will have a directly impact on the engine test. For security, it have to real-time monitor many significant parameters, such as oxygen valve outlet pressure (Pejy), oxygen tank pressure (Pxy), pre-pump pressure of oxygen (Pohy), etc. which ensure reliability of entire system. Because system conditions rapidly changed before ignition start, the value of parameters mentioned above changed abruptly. It is necessary to use sensors real-time monitoring signal change, in order to prevent unnecessary damage.

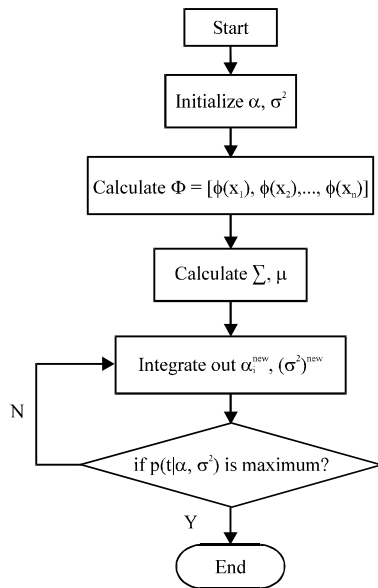


Fig. 2: Modeling process

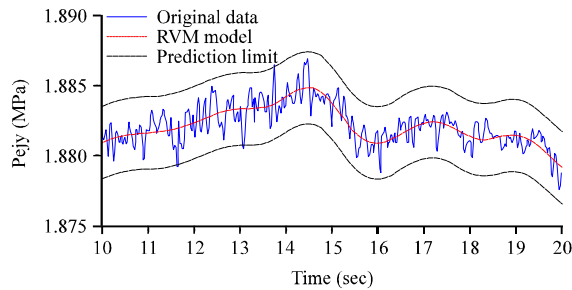


Fig. 3: The results of training Pejy via APSO-RVM

Table 3: Pejy training parameters and results

σ	Fitness	RV	MSE	RMSE
0.103	0.0929	8	7.24×10^{-7}	8.5×10^{-4}

Here, it demonstrates how the proposed APSO-RVM framework is efficiently applicable to the project of liquid rocket engines test-bed fault detection. This paper focus our analysis on the parameter of Pejy and model it using 37 times normal test data which are truncated 10 seconds data in pre-cooling phase before ignition with 50 Hz sampling frequency. The total sample points are 500. The modeling process is as follows:

- Data preparation. First of all, make data preprocessing and normalize to [0, 1]. Besides, select Gaussian kernel and optimize kernel parameter using APSO-RVM. Return optimal kernel parameter σ
- Modeling using the theory of section 2. Simplified process is as Fig. 2
- Calculating prediction error and evaluating upper and lower limit of model
- According to Chebyshev inequality $P(|x-\mu| \geq n\sigma) \leq 1/n^2$, via APSO-RVM model, obtain:

$$P(|\sum_{i=1}^M w_i \Phi_i(x) - \mu| \geq n\sigma) \leq \frac{1}{n^2} \quad (18)$$

In practice, making estimated standard deviation S instead of the standard deviation σ . The threshold range of sensor signal characteristics can be formulate as:

$$C_i = [w^T \phi(x_i) - nS, w^T \phi(x_i) + nS] \quad (19)$$

where, w^T is relevance vector, $\phi(x_i)$ relevance vector kernel matrix, $n = 3$. The probability of falling into this interval is 99.74%. The results of training Pejy model via APSO-RVM are shown in Fig. 3. Training parameters and results are presented in Table 3.

The experimental results indicated that APSO-RVM has excellent modeling ability (fitness is 8/500), while reduces detection time.

Furthermore, for verifying the performance of model presented via APSO-RVM, it employed on the actual data of an oxygen valve leakage. Using the method proposed in this literature, alarm occurred at 17.53 sec which is 0.95 sec ahead of alarm time via red threshold detection method.

CONCLUSION

In this study, a novel APSO-RVM kernel parameter method is proposed for augmenting RVM kernel parameter selection by adaptive PSO. What this paper done make up for amending the defects in the analytic instructional method for RVM kernel parameter selection. Results show that kernel parameter will converge in a tiny region using analytic instructional method proposed instead of experience and ergodic searching. It is efficiently applicable to the project of liquid rocket engines test-bed fault detection and can rapidly detect failure which meet the engineering requirements of real-time and reliability with high practical significance. According to the different observer points, it can flexibly analysis data features and select the appropriate parameters. It will have extensively effect.

As further study, the proposed model remains to integrate consider the pressure and flow direction among different observer points.

REFERENCES

Figueroa, F. and J. Schmalzel, 2006. Rocket Testing and Integrated System Health Management. In: Condition Monitoring and Control for Intelligent Manufacturing, Hafez, M.M. (Ed.). Springer Verlag, London, pp: 373-391.

- Han, Q. and X. Hu, 2007. Application of support vector machine in steady state fault detection and diagnosis of liquid-propellant rocket engine. *Missile Space Veh.*, 4: 54-58.
- Huang, C., Y. Zhang and Q. Chen, 1999. Neural network approach to fault detection of liquid rocket engine-pattern recognition technology. *J. Propul. Technol.*, 20: 1-4.
- Huang, Q., J.J. Wu, H.G. Liu, T.F. Xie, 2007. Implementation of real-time fault detection algorithms based on neural network for liquid propellant rocket engines. *J. Nat. Univ. Defense Technol.*, 29: 10-13.
- Li, G., G. Wang and H. Xue, 2010. GA optimizing method to kernel function parameters of RVM. *Control Eng. China*, 17: 335-337.
- Tao, X., J. Xu, B. Du and Y. Xu, 2008. Bearing fault detection based on RVM using phase space. *J. Vibr. Shock*, 27: 6-10.
- Tipping, M.E., 2001. Sparse Bayesian learning and the relevance vector machine. *J. Machine Learn. Res.*, 1: 211-244.
- Tipping, M.E. and A.C. Faul, 2003. Fast marginal likelihood maximization for sparse Bayesian models. *Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics*, January 3-6, 2003, Florida, USA.
- Tipping, M.E. and N. Lawrence, 2005. Variational inference for Student-t models: Robust Bayesian interpolation and generalised component analysis. *NeuroComputing*, 69: 123-141.
- Tu, J., Y. Zhan and F. Han, 2011. An improved PSO algorithm coupling with prior information for function approximation. *Inform. Technol. J.*, 10: 2226-2231.
- Wu, J., 2005. Liquid-propellant rocket engines health-monitoring-a survey. *ACTA Astronaut.*, 56: 347-356.
- Yu, C., H.X. Han and W. Min, 2009. Missile fault detection based on linear parameter varying fault detection filter. *Inform. Technol. J.*, 8: 340-346.