International Conference on Advanced Computer Science and Electronics Information (ICACSEI 2013)

Research on Modeling Method Based on Least Squares Support Vector Machine

Jun Xu[1], Kang Du[1], Yaohui Zhang[2] and Xiaobing Zhu[1]

[1]Department of Control Engineering, Academy of Armored Force Engineering, Beijing 100072, China

[2]Department of Technical Support Engineering, Academy of Armored Force Engineering, Beijing 100072, China

dkwine@163.com

Abstract - A method of support vector machine based on particle swarm optimization was proposed for the question of parameter selecting difficult of least square support vector machine in modeling of gas sensor. Least square support vector machine is used to build the model of gas sensor. Particle swarm optimization arithmetic was introduced to optimize the parameters of support vector machine. The sensor model is tested with the data measured reality. The results prove the accuracy of the model.

Index Terms - Particle swarm optimization algorithm, Support vector machine, Gas sensor, Mathematical modeling

I. Introduction

As an important device for measuring the gas concentration, the gas sensor has been widely used in the field of energy, electricity and environmental protection. It has a very important significance for the quantitative detection and analysis of the gas mixture in the relevant fields. There are many quantitative analysis methods of gas mixture, such as Artificial Neural Networks (ANNs), Support Vector Machine (SVM) and so on, but some problems also exist in the above methods. The prediction accuracy of ANNs are related with its initial choice of the weights, the output is unpredictable and easy to fall into local minimum. Despite overcoming these shortcomings mentioned above, but the selection of the SVM algorithm parameters directly affect the predictive accuracy of the model. It is difficult to get optimal parameters by searching blindly, and that requires a great deal of time and the number of samples [1] [2].

It was built the LS-SVM model of three components gas mixture, which contains NO_2, SO_2 and NO in this paper. To solve the problem of LS-SVM parameter optimization, the Particle Swarm Optimization (PSO) was introduced to optimize penalty factor and kernel function parameters [3] [4].

II. Least Squares Support Vector Machine

SVM is a method of universal machine learning based on statistical learning theory proposed by Vapnik. Its basic principle is transforming the input space to a high dimensional feature space through a non-linear transformation of the inner product function definition. The non-linear relationship between the input and output variables in the high dimensional space is found by the hypothesis space of the linear function [5].

The LS-SVM is the development of the support vector machine. The different point is that the loss function of the optimization objective is represented by the two-norm error. The inequality constraints in the LS-SVM are instead of equality constraints in SVM, to improve the convergence rate [6]. The structure of the least squares support vector machine shown in Figure 1.

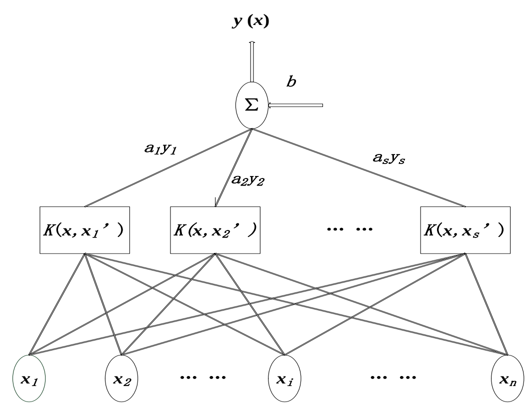

Fig.1 Structure of LS-SVM

As the figure shown, the input space of SVM is constituted by the original observation data, that is mapped via the kernel function to a high dimensional feature space. In this feature space, data is classified or fitted by linear functions in SVM. Finally obtained the relationship between the input x and output y:

$$y(x) = w^T x + b = \sum_{i=1}^{s} a_i K(x, x_i) + b \qquad (1)$$

In this formula, x_i is a support vector, s is the dimension of support vectors, x is the measured input vectors; w is the value coefficient, and its dimension is equal to the dimension of SVM, b is the threshold value or offset, a_i is the corresponding Lagrange multiplier with weight coefficient, $K(x, xi)$ is the kernel function. Kernel functions of SVM come in many forms, including linear kernel, polynomial kernel function and RBF kernel function. Different kernel functions have different output expressions and results [7].

III. Particle Swarm Optimization

PSO was put forward by Dr. Eberhart and Dr. Kennedy from American Purdue University, in 1995 [8]. The basic idea is to find the optimal solution through collaboration and information sharing between groups of individuals. At first, the system initializes a group of random particles (random solution), and then finds the optimal solution by iteration. The particles update themselves by tracking two extreme values *(pbest, gbest)* in each of iteration. After finding the optimal value, the particles update their own speed and position by the following formulas [9] [10].

$$V_{i+1} = \omega \times V_i + c_1 \times rand() \times (pbest_j - x_j) + c_2 \times rand() \times (gbest - x_i) \tag{2}$$

$$x_{i=1} = x_i + V_{i+1} \tag{3}$$

V_i is the velocity of the particles, *pbest* represents local optimal solution, and *gbest* represents the global optimal solution, *rand* is a random number in the range (0,1), X_i is the current position of the particle, c_1 and c_2 are learning factors [11] [12].

IV. Simulation and Results

The amount of data Obtained from gas sensors was very large. 208 groups of data were selected in this simulation and optimization to reduce the amount of calculation and improve the operation speed,

A. Optimization Process

The non-linear RBF kernel was chosen, because the gas sensor has a non-linear characteristic:

$$K(x, x_i) = e^{-\frac{\|x - x_2\|^2}{2\sigma^2}} \tag{4}$$

Where σ is the parameters of the kernel function, and it can improve the predictive accuracy of SVM. The expressions of support vector machine model output:

$$y_p = b + \sum_{i=1}^{n} w_i \exp\left\{-\frac{(x - x_i)^2}{2\sigma^2}\right\} \tag{5}$$

Set number of particle swarm *N = 20*, particle swarm optimization Algebra 100, inertia weight factor ω take the initial value of 0.9, the termination value of 0.3, the learning factor c1 = c2 = 2, particle swarm optimization parameter σ 0 to 20.

The square error between sample forecast results y_{ij} setted for the model test of objective function and the sensor expectation model output y_j created by SVM:

$$f_i = \frac{1}{l} \sum_{j=1}^{l} (y_{ij} - y_j)^2 \tag{6}$$

Where f_i denotes the fitness value of the i-th particle in a value of 1~10, Y_{ij} is the and j-th sample model-predicted value of the i-th particle, y_j is the j-th sample model expectations. The size value (c, σ) of each particle was substituted into the SVM model. According to the calculation results of the sample, the fitness value of corresponding to each particle was obtained by the formula (6). The optimization process is illustrated in Figure 2.

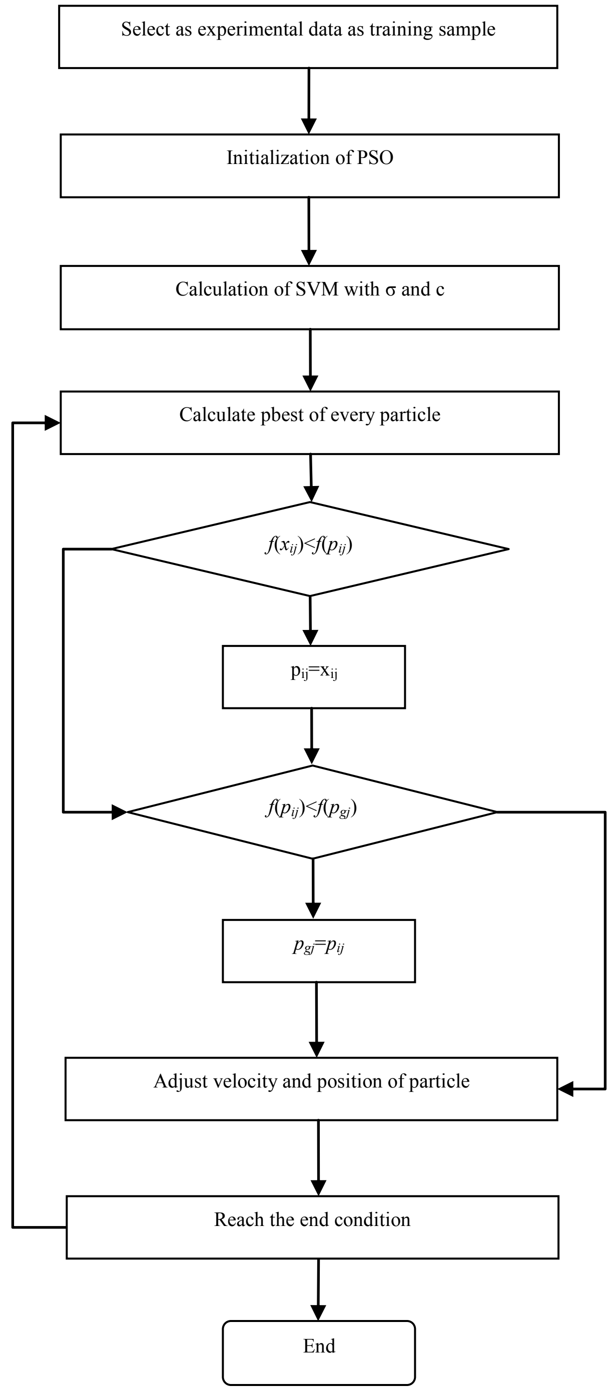

Fig.2 Optimization process of PSO

B. Simulation Results

The optimal parameters had gotten by PSO, and set into the formula (5) to obtain the model of the sensor. The model was compared with the actual measurement of gas concentration to test the accuracy. As shown in Table 1.

The linearity of the three gases NO_2, SO_2 and NO were calculated based on the obtained maximum fitting deviation. As shown in Table 2.

TABLE 1 Part Simulation Results of Measuring Data

Measuring concentration(10^{-6})			Simulation results of model(10^{-6})			Reference errors(%)		
NO_2	SO_2	NO	NO_2	SO_2	NO	NO_2	SO_2	NO
0	281.7	1166.2	0.7	279.4	1162.0	0.098	0.31	0.36
0	423.1	1166.2	1.3	427.2	1174.8	0.18	0.55	0.74
0	741.6	1166.2	3.2	735.0	1164.7	0.45	0.89	0.13
194.3	281.7	0	198.3	281.9	1.2	0.56	0.027	0.10
194.3	423.1	0	191.3	421.7	1.1	0.42	0.19	0.094
194.3	741.6	0	193.2	740.9	1.5	0.15	0.094	0.13
194.3	0	305.4	195.5	0.9	308.1	0.17	0.12	0.23
539.8	0	305.4	538.1	0.4	300.6	0.24	0.054	0.41
539.8	281.7	305.4	539.1	283.7	300.5	0.098	0.27	0.42
539.8	423.1	305.4	538.8	420.0	301.0	0.14	0.42	0.38
539.8	741.6	305.4	534.3	741.0	298.0	0.77	0.080	0.63
712.4	281.7	1166.2	711.0	281.3	1159.8	0.20	0.054	0.55
712.4	423.1	1166.2	708.4	420.1	1160.9	0.56	0.40	0.45
712.4	741.6	1166.2	707.6	738.5	1161.6	0.67	0.42	0.40

TABLE 2 Linearity of the Three Gases

Gas	The maximum reference errors(10^{-6})	Measurement range(10^{-6})	Linearity(%)
NO_2	0.3	712	0.04
SO_2	2.2	740	0.30
NO	6.4	1166	0.55

V. Conclusion

The accuracy of SVM model largely depends on the selection of relevant parameters. This paper has solved the problem of parameter optimization by introducing PSO algorithm. The SVM model has been established by the sample of measurement data and evaluated at the same time. As the experimental results shown, the gas sensor model based on SVM with PSO has high accuracy by comparing with the measured data. is less than 1% and meet the requirements of the accuracy.

References

[1] Y Zhang, X Xu, Z Wang. An overview on support vector machine and time series prediction. Computer Applications and Software, 2010,27 (12):127-157.

[2] N Deng, Y Tian. An New Method of Data Mining— Support Vector Machine. Beijing: Science Publishing Company, 2004.

[3] J Kennedy, R C Eberhart. Particle Swarm optimization// Proceedings of IEEE International Conference on Neural Networks, 1995:1942-1948.

[4] B F Vanden, A P Engelbrecht.A cooperative approach to particle swarm optimization. IEEE Transactions on Evolutionary Computation, 2004,8(3):225-239.

[5] V N Vapnik. An overview of statistical learning theory. IEEE Trans on Neural Network,1999,10 (5): 988-999.

[6] V N Vapnik. The Nature of Statistical Learning Theory . New York: Springer,1995.

[7] J A K Suykens, J Vandewalle. Recurrent least squares support vector machines. IEEE Transactions Circuits and System I, 2000, 47(7): 1109-1114.

[8] W Liu, Q Cai, H Liu. Particle swarm optimization algorithm based on parametric equation method to handle equality constraints. Computer Engineering and Design,2008, 29(3): 697-699.

[9] X Sun, S Zhao, Z Yan. Antisubmarine Search Research on the Optimization of Particle Swarm. Microelectronics & Computer, 2008,25(10):91-93.

[10] [X Wei, Y Zhou. Numerical Integral Method Research Based on PSO. Microelectronics & Computer, 2009,26(7):117-119.

[11] Y Shi, R C Eberhart. Parameter selection in particle swarm optimization. Evolutionary Programming VII, New York, 1998: 591-600.

[12] Y Shi, R C Eberhart. A modified particle swarm optimizer. Proceedings of the IEEE International Conference on Evolutionary Computation, 1998:69-73.