The Research on Fault Diagnosis for Gas Recovery of Single Coal Bed Methane well Based on Improved Particle Swarm Optimizing Support Vector Machine

Yu Miao[1, a], Jianhua Yang[2, b] and Wei Lu[3, c]

[1]School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China;

[2]School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China;

[3]School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China.

[a]18241114050@163.com, [b]jianhuay@dlut.edu.cn, [c]luwei@dlut.edu.cn

Keywords: coal-bed methane; support vector machine; fault diagnosis; Particle Swarm Optimization

Abstract. As a new type of energy, coal-bed gas plays an important role in the national resource structure. This paper introduce the principle and process of gas recovery of single coal-bed methane well, according to the analysis of faults occurred in the system of gas recovery of single coal-bed methane well, by analyzing the characteristics parameters of gas recovery of single coal bed methane well system, combining with the advantages of support vector machine theory can solve the problems of nonlinear and high dimension. Because of the parameters selection of support vector machine has great influence on fault diagnosis, this article use particle swarm algorithm to optimize the parameters of support vector machine. In order to improve the shortcoming of Particle Swarm Optimization (PSO) algorithm which is easy to fall into local optimal, this article proposed that utilize improved particle swarm optimization support vector machine model for gas recovery of single coal-bed methane well system. The simulation results show that the new diagnosis model has a good fault diagnosis practicality and can be applied to fault diagnosis of single well.

Introduction

Coal-Bed gas exists in the coal strata by physical adsorption state, in production, through the continuous drainage of the water pump, the pressure of the coal-bed methane[1] drops below adsorption pressure, methane gas can be collected. Through the analysis on the historical data of the coal-bed methane well and the status of coal-bed methane, we find that there are three kinds of typical fault in the single-well field: pumping unit stuck failure, suppress the well failure and nozzle clogging failure. The typical parameters which can reflect the working status of the single well and also can be measured in the field are: fluid level, pipe pressure and gas flow rate.

Researches on fault diagnosis system lasted for a long time，fault diagnosis methods become more sophisticated, the current methods of artificial intelligence include gray theory[2], rough set[3] and support vector machine, artificial intelligence are widely applied to industrial fault diagnosis, also achieved better diagnostic results. The goal of support vector machine[5] is to minimize the structure, support vector machine can solve the small sample problems, which has good generalization ability, but the model parameters has a great influence on the model of support vector machine.

Support Vector Machine (SVM) Model

Set $T=\{(x_1,y_1),...,(x_n,y_n)\}\in(X\times Y)^n$ as the training data, n as the number of sample, x_i as the feature vector, $y_i\in\{-1,1\}$ as the classification of labels.

The linear distinguish function is: $f(x)=(\omega\cdot x_i)+b$, The hyper plane equation is: $(w\cdot x_i)+b=0$, The sample should meet: $|f(x)|\geq 1$, The constraint condition is: $y_i[(w\cdot x_i)+b]-1\geq 0,\ \ i=1,2,...n$。

For linear separable problem, the process of solving the optimal hyper plane can be transformed into two planning problem:

$$\begin{cases}\min\limits_{\omega,b,\xi_i}\left[\dfrac{\|w\|^2}{2}\right]\\ s.t.\ \ y_i[\left(w\cdot x_i\right)+b]\geq 1,\ \ i=1,2,...,n\end{cases}\tag{1}$$

In the above formula, the greater of the value of C, the more severe of the punishment, the higher complex of the model, the optimization problem can be transformed into the minimum value problem of Lagrange's function:

$$L(w,b,\alpha)=\frac{\|w\|^2}{2}-\sum_{i=1}^{n}\alpha_i\left\{y_i\left[(w\cdot x)+b\right]-1+\xi_i\right\}\tag{2}$$

Seeking partial differential and make it equal to 0, resulting a Convex Quadratic Programming dual problem:

$$\begin{cases}\max(\sum\limits_{i=1}^{n}\alpha_i-\dfrac{1}{2}\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}\alpha_i\alpha_j y_i y_j(x_i\cdot x))\\ s.t.\quad 0\leq\alpha_i\leq C, i=1,...,n\\ \sum\limits_{i=1}^{n}\alpha_i y_i=0\end{cases}\tag{3}$$

α_i^* is the only optimal solution, if the value of α_i^* is not equal to 0, α_i^* is the support vector, $w^*=\sum_{i=1}^{n}\alpha_i^* y_i x_i$, the value of b^* can be achieved by any of support vector substituted into equation $y_i[\left(w^*\cdot x_i\right)+b^*]-1=0$, finally the linear support vector machine classification decision function is:

$$f(x)=\mathrm{sgn}\left\{(w\cdot x_i)+b\right\}=\mathrm{sgn}\left\{\sum_{i=1}^{n}\alpha_i^* y_i(x_i\cdot x)+b^*\right\}\tag{4}$$

For nonlinear problems, in order to avoid "dimension disaster" problem, using kernel function $K(x_i,x_j)$ instead of the high-dimensional space of the dot product operation, the nonlinear support vector machine classification decision function is:

$$f(x)=\mathrm{sgn}\left\{(w\cdot x_i)+b\right\}=\mathrm{sgn}\left\{\sum_{i=1}^{n}\alpha_i^* y_i K(x_i\cdot x)+b^*\right\}\tag{5}$$

According to the experiences, Gauss radial basis function $K\left(x_i,x\right)=exp\left[-\dfrac{\|x_i-x\|^2}{2\sigma^2}\right]$ is chosen as the kernel function, σ^2 reflects the distribution characteristics of the training samples.

Improved Particle Swarm Optimization Algorithm

The existence of traditional PSO algorithm[6] is easy to fall into local optima, slow convergence and low precision. This paper propose improved PSO algorithm support vector machine, improved PSO[7] as follows:

1) Initialization parameters, the acceleration constants are c_1 and c_2, the initial population is $X\left(n\right)$, the speed of the particle in the population is $\left\{v_1,v_2,...,v_m\right\}$, the maximum evolution algebra is *maxgen*.
2) Calculate the fitness of each individual particle of population $X\left(t\right)$.
3) Leaving the individual particles with high fitness, eliminate low fitness.

4) Select $2k$ ($k \le n/2$) particles randomly from the population, speed and location are X and Y, cross calculate its velocity and position:

$$X_1 = \gamma X + (1-\gamma) Y \tag{6}$$

$$Y_1 = (1-\gamma) X + \gamma Y \tag{7}$$

γ is a random number between (0,1), X and Y are the parent of the particle velocity and position of a particle, X_1 and Y_1 are the child of the particle velocity and position.

5) Calculate the fitness of individual particles after crossing, if higher than the parent, crossover of particles to replace the parent particle and update the individual extreme p_{best} and global extreme g_{best}, the update equation is:

$$\begin{aligned} v_{ij}(t+1) &= \omega_1 \left[v_{ij}(t) + c_1 r_1 \left(p_{best}(t) - x_{ij}(t) \right) \right] + \\ & \omega_2 c_2 \left(g_{best} - x_{ij}(t) \right) \\ x_{ij}(t+1) &= x_{ij}(t) + v_{ij}(t+1) \end{aligned} \tag{8}$$

$\omega_1 = 1 - 0.6 * t / max\,gen$； $\omega_2 = 0.6 * t / max\,gen$， t is the evolutionary generation，*maxgen* is the maximum evolution generation, at early time of the algorithm *t* is small, the value of ω_1 is big, a good play to their ability to search the individual particles, with the increasing number of iterations, the value of ω_2 becomes larger, and the exchange of information between particles increases, speed up the convergence of the global optimum.

6) Update the velocity and position of each particle obtain the population $X(t+1)$.

7) Determine whether the model meets the optimal fitness or the maximum number of iteration, if satisfied then quit, select the best particle output. Otherwise set *t=t+1* and skip 3).

Coal Bed Methane

The typical parameters which can reflect the working status of the single well and also can be measured in the field are: fluid level, pipe pressure and gas flow rate. The relation between typical fault and parameters trend is shown below:

Table.1 The relation between fault and the parameters

fault status	parameters' trend
pumping unit stuck failure	fluid level rise up quickly pipe pressure decrease gas flow rate decrease
suppress the well failure	fluid level rise up quickly pipe pressure decrease gas flow rate decrease
nozzle clogging failure	fluid level decrease normal pipe pressure rise gas flow rate decrease

Experimental data and results

In this paper, the coal-bed methane data from the projects at JinCheng Shanxi province, 180 samples of raw data were collected from the historical data. Which include 60 sample data in a normal condition, 40 sample data in the pumping unit stuck failure status, 40 sample data in the suppress the well failure status, 40 sample data in the nozzle clogging failure status. Each sample is in 4-dimensional vectors, they are: fluid level, pipe pressure, gas flow rate and an output label, the output of the normal state is 1, the output of the nozzle clogging failure state is 2, the output of the pumping unit stuck failure state is 3, the output of the suppress the well failure state is 4, select 40 sample data in a normal condition, 20 sample data in the pumping unit stuck failure status, 20 sample data in the suppress the well failure status, 20 sample data in the nozzle clogging failure status as training set to training SVM fault diagnosis model, the remaining 80 samples as test samples to verify the diagnosis SVM fault diagnosis model.

In this paper, the improved particle swarm optimization support vector machine fault diagnosis model's specific process is as follows:

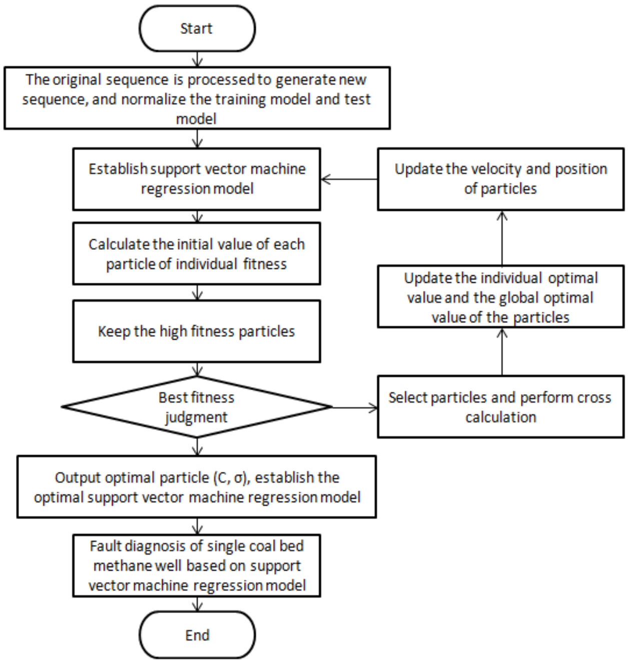

Fig.2 Improved particle swarm optimization support vector machine

The parameters of SVM model have great influence on the model, in order to obtain a better fault diagnosis model, optimization method is needed on SVM, this paper use legacy algorithm, PSO and improved PSO SVM model to optimized the model's parameters, the simulation results as below:

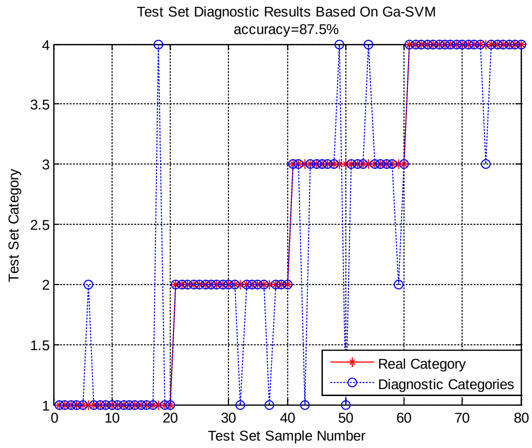

Fig.3 Test Set Diagnostic Results Based on Ga-SVM

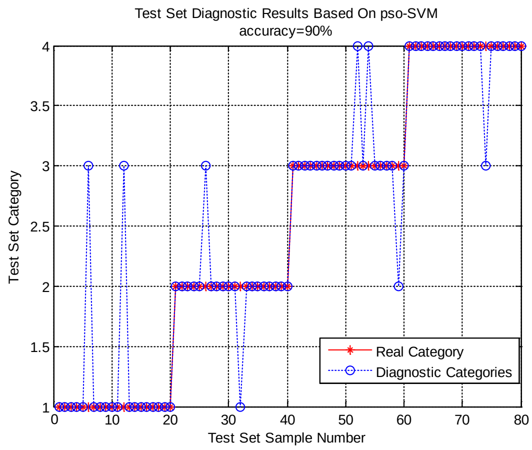

Fig.4 Test Set Diagnostic Results Based on pso-SVM

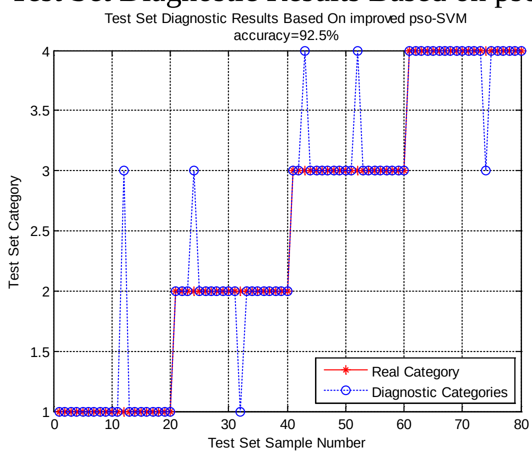

Fig.5 Test Set Diagnostic Results Based on Improved pso-SVM

Diagnosis for 80 test sample data based on genetic algorithms support vector machine model, particle swarm optimization support vector machine model and improved particle swarm optimization support vector machine model, the rate of the results show in table 2:

Table 2 The results of three methods

optimization model	C	σ	rate of Fault diagnosis
GA	0.0933	51.2896	87.5%
PSO	0.0751	43.6258	90.0%
Improved PSO	0.1005	40.6527	92.5%

As we can see from the table, improved particle swarm optimization support vector machine model has a good learning ability and generalization ability.

Summary

This paper described the principle and process of gas recovery of single coal bed methane well, introduced the key parameters that affect coal-bed methane drainage process, fault diagnosis for single well for coal-bed methane based on improved particle swarm optimization support vector machine model. Experimental results show that fault diagnosis for coal-bed methane based on improved particle swarm optimization support vector machine model not only has a good learning ability and training effect, but also has good generalization ability, it can be effectively applied to practical problems.

Acknowledgements

This work was financially supported by Major National Science and Technology Programs of china in the "Twelfth Five-Year" Plan period (2011ZX05039).

References

[1] Sun Z, Zhang W, Hu B, etal. Geothermal field and its relation with coalbed methane distribution of the Qinshui Basin [J]. Chinese Science Bulletin, 2005, 50(1).

[2] Tsai M, Hsiao S, Liang W. Using Grey theory to develop a model for forecasting the demand for telecommunications[J]. Journal of Information and Optimization Sciences, 2005(26):535-547.

[3] Pawlak Z. Rough sets [J]. International Journal of Information and Computer Science, 1982, 11(5): 314-356.

[4] Simon Haykin. YE S, SHI Z translated. The Principle of Neural Network[M]. Beijing: Press of Mechanical Engineering. 2004: 60-80.

[5] CRISTIANINI N, TAYLOR J S. An introduction to Support Vector Machine[M] . LI Z, WANG M, CENG H translated. Beijing: Electronic Industry Press, 2004.

[6] XUE T. The research and improvement of the particle swarm optimization algorithm[M]. Dalian: Dalian maritime University, 2008.

[7] WANG D, LUO Q, QI J. Modified particle swarm optimization[J]. Computer Engineering and Applications, 2008, 44(9): 80-81.