Prediction Model for Rotary Kiln Coal Feed Based on Hybrid SVM

Xiaogang Zhang*, Jingjing Zhao

"College of Electrical and Information Engineering, Hunan University, Changsha 410082, China"

Abstract

A SVM (Support Vector Machine) classification for rotary kiln based on time series trend characteristics was presented in this paper. The amount of coal feed depended on its trends which were predicted by the dynamic rules of rotary kiln data. The original sample was cleaned by data preprocessing, and the piecewise linear representation based on key points was used to extract the trend characteristics of rotary kiln time series data. The parameters of SVM were decided by Particle Swarm Optimization method. It was proved that the proposed model with high prediction level which enhanced the robustness of the control of rotary kiln could be applied on the actual situation.

Keywords: rotary kiln; Support Vector Machine; Particle Swarm Optimization; time series

1. Introduction

As a large thermal equipment, rotary kiln is widely used in building materials, metallurgy, chemical industry, environmental protection and other manufacturing industries. With high energy consumption, production instability, and low automation level, its production process is still largely dependent on human experience. In order to make the operation conditions stable, increase production and reduce consumption, the key to rotary kiln system is how to enhance the level of automation.

Rotary kiln have a lot of complex characteristics, such as large delay, strong coupling, nonlinear, time-varying and so on, so it is difficult to extract expertise, and the system modeling is also complex [Da-ping H et al., 2005]. At present, the study on rotary kiln is mainly focused on the analysis of flame image. By establishing model of sintering temperature to control the temperature, the stable condition is achieved.

As the temperature detection method based on image processing is prone to interfere by dust and smoke, an earlier idea of rotary kiln control using flame image feedback was proposed in [Collins S, 1993, Shimoda M et al. 1990]. In view of it, the fuzzy expert control model based on multi-sensor data integration and a RBF network predictive control model were established in [Xiao-gang Z, et al. 2002, 2007]. In order to improve the accuracy of image and recognition level of working conditions, the flame image feature and key rotary kiln data were used to recognize sintering condition in [Peng S, et al. 2008].

There are some progresses on measurement and control of sintering temperature have been achieved, but the over-reliance on sintering temperature leads to the narrow operation scope of controller, and it is difficult to achieve long time control. In addition, these training methods using directly field signals

* Corresponding author Xiaogang Zhang. Tel.: 13507403156; E-mail address: Zhangxiaogang@126.com
depend on prediction of single sample without further extraction of trend characteristics. So the dynamic trend of thermal signals is easy to be overlooked, the accuracy and generalization ability of the model decreased correspondingly.

Therefore, this paper proposes a rotary kiln model in a new way. Compared to the modeling methods directly using field data, the prediction model based on Support Vector Machine extracts multi-sample coal feeding trend in certain period of time, other than single sample prediction. The dynamic characteristics of multiple thermal data are used to forecast the coal feeding trend, and manual work would be under the guidance of command. So, single signal rotary kiln control which only depends on sintering temperature will be avoided, and the robustness of kiln control is also improved. Firstly, rotary kiln thermal data are preprocessed, and the key change points of coal feeding is extracted, then time series is conducted on the basis of piecewise linear representation. The trend characteristics are extracted to form the training samples, and the parameters of SVM are optimized by Particle Swarm Optimization. It was proved that the proposed model with high prediction level which enhanced the robustness of the control of rotary kiln could be applied on the actual situation.

2. Basic scheme

As shown in Fig. 1, rotary kiln coal feeding prediction model includes three parts: data preprocessing, piecewise linear time series based on key points and SVM classifier.

In this paper, thermal data is from an alumina rotary kiln. The sample point interval is 3 minutes, each sample point contains 8 kinds of key indicators (called attributes of rotary kiln): the amount of coal feeding, sintering temperature, kiln head temperature, kiln inlet temperature, main motor current, cooler current, slurry flow rate and material temperature.

When preprocessing, the original data would be denoised and normalized. The kiln thermal data is considered as the time series sample in piecewise linear representation, and the piecewise trend characteristic sample is formed by extracting kiln key points. Next, the prediction sample which is classified by SVM classifiers is used to predict the trend of coal feeding.

3. Design and Implementation for SVM model

3.1. Data preprocessing for rotary kiln

The original thermal data sample obtained from the kiln site is defined as \( RD = \{x_{i,j}\} \), where \( i = 1,2,\cdots,l \) is the total number of the sample, \( j = 1,2,\cdots,8 \), \( j \) is the number of data attributes. The original sample may not be entirely satisfactory, so there may be unreasonable noise points (In such
cases, when the feed speed limit is achieved, the amount of coal feeding is still increased). Further more, the attributes dimensions of kiln indicators are different, the relationship among these attributes would not be reflected properly. In this paper, the preprocessing method is represented as follows:

Step1: Computing the each sequence average of $RD$, save as $ave(j)$;
Step2: Setting the permissible fluctuation range of each column under manual operation experience, save as $bd(j)$;
Step3: If $|x_{i,j} - ave(j)| > bd(j)$, $x_{i,j} = x_{i-1,j}$, otherwise, $x_{i,j}$ is unchanged;
Step4: Seeking the maximum $mt(j)$ and minimum $mn(j)$ of each column, a non-dimensional sample normalized interval $(0,100)$ is generated based on normalizing $x_{i,j}$ as: $x_{i,j} = 100(x_{i,j} - mn(j))/(mt(j) - mn(j))$.

3.2. The characteristic extraction of time series trend based on key points

Nowadays, there are extensive researches on time series analysis [Prat KB, E. Fink, 2002]. Rotary kiln thermal data can be viewed as a large time series sample. In this paper, key points to increase or decrease the amount of coal feeding are found to process kiln time series based on the method of piecewise linear representation. Extracting the characteristics of sub-trend (where 1 is to increase the coal feed, -1 is to decrease the coal feed), then the training sample of coal feeding trend is formed.

In this paper, the algorithm is designed to find kiln key points as follows, where $\eta$ is the change threshold:

Step1: Initializing, setting the value of trend characteristic $z_i = 0$, $\eta > 0$;

When $i = 1$, $z_i = \begin{cases} -1 & x_{2,1} - x_{1,1} < -\eta \\ 1 & x_{2,1} - x_{1,1} > \eta \end{cases}$ ;When $i > 1$, $z_i = \begin{cases} -1 & x_{i+1,1} - x_{i-1,1} < -\eta \text{ and } x_{i,1} - x_{i-1,1} \geq -\eta \\ 1 & x_{i+1,1} - x_{i-1,1} > \eta \text{ and } x_{i,1} - x_{i-1,1} \leq \eta \end{cases}$ ;

Kiln key points is the samples which satisfy $z_i \neq 0$.

Step2: The sample point interval of rotary kiln thermal data is set as 3 minutes. Segmenting the time series with same distance by key points, each section contains $t/3 + 1$ sample points, therefore, when $i \geq (t/3) + 1$;

Step3: Setting the total number of key points as $n$, the sample points which meet $z_i \neq 0$ will be stored in the new collection $\{y_m\}_{m=1}^{n}$, named the set of key points characteristic;

Step4: Fitting a straight line between the interval $[i-(t/3),i-(t/3)+1,\ldots,i]$ for each column of key point $x_{i,j}$. The fitting function is $g(x) = k_{m,j}x + e$;

Step5: Let $x_m = [k_{m,2}, k_{m,3}, \ldots, k_{m,8}, x_{m,2}, x_{m,3}, \ldots, x_{m,8}]$, where $x_{m,j}$ is sample value of the key point,$KD = \{y_m \mid y_m \}_{m=1}^{n}$ is the trend sample based on piecewise linear representation.

Step6: Taking account of the complexity of rotary kiln system, a single value is insufficient to reflect the change of kiln coal feeding trends accurately, the value $x_{m,j}$ is added into the trend sample to improve the reliability of the sample and the performance of classifier(as shown in the third step).

3.3. SVM classifier of rotary kiln

SVM[Platt J, 1998, Cristianini N, Shawe-Taylor J, 2004, Chih-Chung C, Chih-Jen L, 2001] is built on the basis of statistical theory, the principle of SVM is structural risk minimization which seeks a compromise between experience risk and confidence intervals. The basic idea of SVM is to find a hyperplane in a certain hypothesis space. The hyperplane can divide the data with different characteristics
at large, and its original model is maximum margin classifier. Therefore, the problem becomes a convex quadratic programming problem which finds the largest geometric intervals of classification. For the trend sample \( KD = \{ y_m, x_m \} \), the optimization objective function is:

\[
\min_{\alpha \in \mathbb{R}^n} W(\alpha) = \frac{1}{2} \sum_{m=1}^{n} y_m \alpha_m \sum_{q=1}^{n} y_q \alpha_q K(x_m, x_q) - \sum_{m=1}^{n} \alpha_m y_m \quad \text{subject to} \quad \sum_{m=1}^{n} \alpha_m = 0 \quad 0 \leq \alpha_m \leq C
\]  

(1)

Where \( \alpha_m \) is the lagrange multiplier, \( C (>0) \) is the penalty parameter, which controls the punishment degree of misclassification and gets a compromise between the maximum interval and the total number of kiln sample. \( K(x_m, x_q) \) is the kernel function which meets the conditions of Mercer. Non-linear separable data in low-dimensional space is mapped to linear separable higher dimensional space by the kernel function, then the classifier could find the optimal hyperplane in higher dimensional space. In this paper, Gaussian kernel function is adopted for SVM, then the classification decision function could be got from

\[
f(x) = \text{sgn}\left( \sum_{m=1}^{n} \alpha_m^* y_m K(x, x_m) + b^* \right)
\]  

(2)

Where \( b^* \) is the corresponding offset, \( \alpha_m^* \) is the nonzero lagrange multiplier which determines the separating hyperplane.

PSO is a swarm intelligence optimization algorithm which is inspired to solves the problem from population characteristics of birds preying[Kennedy J, Ebethart RC, 1995].

Firstly, initializing a group of particles in the solution space, each particle represents a potential optimal solution of the extremal optimization problem which indicators are the value of position, speed and fitness. The fitness value which represents the quality of the particles is determined by fitness function.

Particles move in the solution space. \( P_{best} \) is defined as the best fitness position in individuals experiencing extreme values, \( G_{best} \) is the optimal fitness location of all particles. The fitness would be calculated when each particle is updated, \( P_{best} \) and \( G_{best} \) could be updated by the comparison of fitness of new particles, fitness of individual extreme values and groups extreme value.

If penalty parameter \( C \) is too high, overfitting will happen, and the generalization ability will also be brought down. In other words, the classification accuracy of train set is high, but test set is very low. Another key parameter \( \sigma \) of the kernel function has great impact on the classification, so optimization for classifier design process is also difficult. In this paper, PSO is used to optimize the parameters \( C \) and \( \sigma \). The group with lower \( C \) in all pairs which meets the highest classification accuracy is chosen as the parameters of SVM classifier. The algorithm steps are as follows:

Step1: Setting PSO parameters, including speed range, location, range, population size, number of iterations, and so on;
Step2: Choosing the classification accuracy as fitness function;
Step3: Initializing the particles and speed, and calculating the initial fitness;
Step4: Identifying \( P_{best} \) and \( G_{best} \) of parameters \( C \) and \( \sigma \) respectively;
Step5: Updating speed and population, mutating particles based on self-adaption;
Step6: Calculating the fitness value; If the fitness not only meets the algorithm requirements but also completes the iterations, the optimal \( (C, \sigma) \) will be obtained, otherwise, turn to the fourth step.

According to the analysis above, classification prediction process of coal feeding trend based on SVM is designed as follows:

Step1: Preprocessing the original thermal data \( RD \).
Step 2: Finding the key points and conducting the time series based on piecewise linear representation, then the coal feed sample is formed.

Step 3: Getting the training set and test set of KD.

Step 4: Obtaining optimal C and σ using PSO.

Step 5: On the basis of optimal parameters from the fourth step, training SVM classifier by train set.

Step 6: Classification predicting by using test set. Results and accuracy of classification are obtained.

4. Conclusion of experimental analysis

LIBSVM [Chih-Chung C, Chih-Jen L, 2001] is used in this paper, and the experimental data is from on-site thermal data of an alumina rotary kiln. Gaussian kernel $K(x, x') = \exp(-\|x - x'\|^2 / \sigma^2)$, is selected for SVM kernel function, where $\sigma > 0$ is the width parameter of kernel function. The remaining SVM parameters adopt the default values of LIBSVM. The sub-segment length $t$ is identified as 30 minutes, and the data sample is established on key points which have been found.

As shown in Figure 2, the key points of the alumina rotary kiln coal feeding time series from 17:36 on August 5 to 00:35 on August 6 in 2009 are extracted by the presented method in part B in $\Box$, longitudinal coordinate shows the speed of twin screw of coal feeding. The marked points indicate coal feeding changes of the next moment. The amount of coal feeding would be increased in the next moment when the triangle points upwards, and vice versa. Obviously, the key points in Fig. 2(a) are consistent with the actual curve. That is to say, the key points extracting method presented in this paper is effective.

![Key points of rotary kiln coal feed in 7 hours](image1)

![Prediction of coal feed in 7 hours based on traditional method](image2)

In the traditional model of rotary kiln, the field thermal signals are used for training directly, and the training sample is formed by single sample data. Constructing SVM classifier and predicting the situation in 7 hours coal feeding curve shown in Fig.2(a), the results are illustrated in Fig.2(b).

![Prediction of coal feed in 7 hours with random parameters](image3)

![Optimization of parameters by PSO](image4)
There are 12 error prediction points marked in Fig.2(b), and the arrow indicates the correct coal feeding trend. The remaining 12 points match with the actual. That is to say, the accuracy rate of prediction is only 50%, and it can’t meet the requirements in field.

SVM classifier is established with random parameters C and \( \sigma \) by using the method proposed in this paper, the prediction results in 7 hours coal feeding curve are shown in Fig.3(a). The accuracy rate is raised to 58% with 10 error points and 14 correct points.

So, it is proved that the prediction method based on trend characteristics has higher reliability than traditional method based on single signals. However, the current prediction accuracy still can’t meet the practical requirements of industrial application. Therefore, the algorithm which could obtain the optimize parameters \( C \) and \( \sigma \) is designed to improve prediction accuracy of SVM classifier.

Setting the SVM classification accuracy rate as the fitness value, and the maximum iteration is 200. The traditional Genetic Algorithm[Cheng-Lung H, Chieh-Jen W,2006] and the PSO proposed in part b of C in [9] are respectively used to optimize the parameters \( C \) and \( \sigma \). The training results are shown in Table 1.

<table>
<thead>
<tr>
<th>Population size</th>
<th>GA</th>
<th>PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Best C</td>
<td>Best ( \sigma )</td>
</tr>
<tr>
<td>20</td>
<td>40.5</td>
<td>2.9</td>
</tr>
<tr>
<td>40</td>
<td>56.7</td>
<td>14</td>
</tr>
<tr>
<td>60</td>
<td>48.62</td>
<td>3.25</td>
</tr>
<tr>
<td>80</td>
<td>35.54</td>
<td>7.8</td>
</tr>
<tr>
<td>100</td>
<td>12.43</td>
<td>1.5</td>
</tr>
</tbody>
</table>

As is shown in Table 1, PSO algorithm has more advantages not only in optimal accuracy but also in terms of running time than traditional GA algorithm. The best accuracy rate of PSO algorithm is 66.67 when the population size is 60. The optimization process for the parameters is shown in Fig.3(b).

According to the Table 1, with the penalty parameter \( C = 30.05 \) and kernel width parameter \( \sigma = 5.14 \), the optimal SVM model is set up to predict coal feeding in 7 hours shown in Figure 2. The prediction result is illustrated in Fig.4. The error predicted state reduced to 8 points, and the prediction accuracy of coal feeding trend characteristic reaches 70%. The results can meet the level of actual application. what’s more, the practical feasibility of the proposed method in this paper can be proved.
5. Conclusion

A SVM classification for rotary kiln based on time series trend characteristics was presented in this paper. It is proved that the proposed model has more reliability than the traditional methods based on single sample points. Next, optimizing the parameters of SVM by the traditional GA and PSO respectively, and adopting the rules of dynamic rotary kiln thermal data to predict the coal feeding trend directly. The speed of feeding coal is adjusted to make the operation of rotary kiln maintain at a stable state. Then the control condition which only depends on sintering temperature is avoided, it improves the robustness of the control of rotary kiln, and have a constructive practical significance. The comparative analysis based on experiments shows that the proposed classification with a high prediction accuracy can be applied in the actual rotary kiln production to predict coal feeding trend. But rotary is a complex control object, the subject for further study is to explore the prediction method of kiln coal feeding in view of various factors.

Acknowledgements

This work was supported by National Natural Science Foundation of China(60874096). The authors wish to thank the mill of ZhongZhou Branch China Aluminum CO.LTD and the lab mates of the Hunan University for their valuable support and assistance during this project.

References