State Recognition of Mill Load Based on Improved K-means Clustering Algorithm

Yunpeng Gao1,a, Cong Wu2, Jizong Peng3, Yuchen Lu4

1Hunan University College of Electrical and Information Engineering Changsha, Hunan Province, China
2Hunan University College of Electrical and Information Engineering Changsha, Hunan Province, China
3Ecole de technologie superieure, Tuninsight AI, Montreal, Canada
4Universite de Montrea JNNRVFK1QC1 / Mila, Tuninsight AI, Montreal, Canada

Corresponding author’s e-mail: gaoyp@hnu.edu.cn

Abstract—In this work, a new identification method for the load identification of ball mill based on improved K-means clustering is proposed. Since the external device response signals of ball mill are closely related to its load during the operation, the appropriate parameters are selected. The number of clusters and initial clustering centers are meliorated based on their density and genetic algorithm. The dimensionality of input features are reduced with kernel principal component analysis (KPCA), and the final clustering centers are obtained by iterating in the space with reduced dimensionalities. Tests on datasets have demonstrated the higher accuracy and stability of our improved K-means method compared with other methods. The case-analysis results show that the mill load status can be identified clearly according to the displacement variations of cluster centers calculated by this method. Our research provides a new reference for mill load identification in the viewpoint of multi-source information fusion.

1. INTRODUCTION

The ball mill is the key equipment in the concentration plant for ore beneficiation. Statistics show that approximately 5% of the power generated in China annually is consumed by ore grinding, and more than 50% of steel ball is consumed by the mineral concentration industry, and the energy consumption of ball mills accounts for 50–60% of that in the entire beneficiation process [1, 2]. Ball mill load directly affects the production efficiency and energy consumption of beneficiation equipment and systems. Therefore, identifying the load status during operation and maintaining the ball mill under ideal load through timely operation and control can greatly facilitate the efficiency, safe production, and energy conservation.

It is complicated and difficult to directly identify the ball mill load, which can only be indirectly characterized or presented by some key variables. Because ball mill operation involves many external devices, such as the machine base, the motor, the classification mechanism, and the transmission mechanism, the external response signals of ball mill can only be used to indirectly characterize the load. Based on a parallel RBF kernel function or a neural network, reference [3] proposed to measure the ball mill load by identifying a set of parameters including coal supply, hot air flow, recirculated air
flow, coal mill outlet temperature, coal mill inlet and outlet pressure difference, and coal mill current. References [4, 5] established a soft measurement model with a multi-sensor information fusion algorithm to predict the internal load parameters of the ball mill based on the vibration, grinding sound and power signals of ball mills. Reference [6] measured the inlet and outlet pressure difference, hot air flow, inlet air negative pressure, outlet pulp temperature, recirculated air flow, and feed volume of double-in and double-out ball mill and integrated the vibration, vibration sound and useful power information to construct the BP neural network prediction model, which can identify the ball mill load to a certain extent. However, this method failed to analyze the correlation of the external features of the mill. In addition, it requires a large number of data samples, a long training time, and is difficult to apply it to practical situations. Reference [7] used external feature data such as inlet and outlet negative pressure, inlet pressure difference, mill power, vibration, and signal data including ore supply, hot air flow, and circulating air flow. It divided the parameters according to their credibility, used the Dempster-Shafer evidence theory to measure the load. This method solved many difficult problems, but the Dempster-Shafer evidence theory requires a large number of experimental samples, especially the prior data, and the time variability of the ball mill load makes it difficult to implement and has a low accuracy.

To seek more effective and applicable ball mill identification method, this paper analyzes “one-stage” grinding process and selects 7 related parameters, which are analyzed after KPCA dimension reduction through the improved K-means algorithm based on density and genetic algorithm. Then, the paper calculated the weight of parameters by the entropy weight method and used it as the abscissa, and proposed to identify the change of the ball mill load status according to the cluster centers displacement in the data space, so as to provide foundation for online monitoring of ball mill load.

2. BALL MILL LOAD AND RELATED PARAMETERS

2.1. Ball mill load classification

Ball mill load is the instantaneous total load of the grinding medium, new ore supply, new water supply, and slurry, which are pre-filled inside the cylinder [8]. Due to the long-term operation of the ball mill in a closed state and the dynamic change of the internal environment, specific load parameters such as material-to-cylinder ratio and slurry concentration cannot be accurately and quantitatively described. During the production process, the amount of ore and water is usually according to the experience of professional operators to ensure the operation safety and stability. Therefore, based on the operator experience, this paper divided the ball mill load into three states: underload, ideal load and overload.

When the mill is under-loaded, the water, steel balls, and minerals inside can fill about 20% to 30% of the ball mill, in which its production capacity is not fully reached, causing a waste of resources; whereas under the status of ideal load, about 30% to 40% of the ball mill is filled, in which it has an optimal working load. The capacity of the ball mill table is increased, and energy waste is reduced. When the ball mill is over loaded, the materials inside accounts for about 40% or above, which is prone to blockage, which not only causes the low production, but also may affect the safe operation of the ball mill and lead to economic losses.

2.2. Grinding process

The conventional “one-stage” grinding process [9] is shown in Figure 1, which includes equipment such as ball mills, cyclones, sand pump ponds and slurry pumps.

As shown in Fig.1, the crushed ore is transferred from the belt to the inlet of the ball mill and a certain amount of water is added. Then, the mixture of minerals and water enters the cylinder, and is lifted and then dropped with the pre-added steel balls as driven by the high-power motor. The ore is ground by the impact and friction, and the fully ground slurry flows from the ball mill outlet into the sand pump pool, and supplemental water is also added to adjust the concentration. The slurry in the sand pump pool goes into the cyclone slurry inlet driven by the slurry pump and then classified by
centrifugal force and gravity. The slurry that meets the fineness requirement enters the flotation process from the overflow outlet. Otherwise, the slurry is returned to the ball mill inlet for secondary grinding.

2.3. Ball mill load related parameters
The difference in the load status of the ball mill leads to the grinding efficiency variations, which causes the changes of the parameters measured by the instruments in the grinding process. Therefore, these parameters are able to reflect the load status to a certain extent. This paper used the FCS historical data in the industry as the original dataset, from which 7 load-related parameters were selected. The specific introduction is shown in Table 1.

3. BALL MILL LOAD IDENTIFICATION METHOD BASED ON IMPROVED K-MEANS
Seven parameters of an operating MQYΦ5.5×8.5m ball mill on the industrial site were collected as shown in Table 1. We exhibits in Fig.2 the evolution of these parameters in a single day (September 2018).

Fig.2 presents that due to the complex time-varying and coupling between the parameters during the grinding process, it is hard to directly find clues of the underlying rules of their changes. To fully unleash the potentially useful information, this paper established a clustering method to process the parameters when the ball mill reaches dynamic equilibrium, and distinguished the load status based on the clustering results.

3.1. Kernel Principal Component Analysis (KPCA)
The 7 load-related parameters have different units and, some are non-linear, and some are overlapped or even redundant. In order to effectively characterize the load status while reducing the calculation, this paper applies KPCA to map the data to high-dimensional space, and then use PCA to reduce the dimension of the dataset [10].

The load-related parameter dataset was set to be $X=[X_1, X_2, \ldots, X_7]^T \in \mathbb{R}^{7 \times n}$, $X_i=[x_{i1}, x_{i2}, \ldots, x_{in}]^T \in \mathbb{R}^{1 \times n}$, which is a curve matrix composed of $n$ sampled data within 10 minutes of the $i$-th parameter of the ball mill under dynamic equilibrium. The mapping of the data $x_i$ in the high-dimensional feature space $F$ was set to be $\phi(x_i)$ and its projection on the feature vector $v_2$ in $F$ was calculated, then the principal element vector was expressed as
### TABLE 1  THE RELEVANT PARAMETERS OF MILL LOAD

| Parameter                  | Unit     | Testing instrument                        | description                                                                                           |
|----------------------------|----------|-------------------------------------------|-------------------------------------------------------------------------------------------------------|
| Ore supply, water supply   | t/h      | Electronic belt scale, electromagnetic flow meter | Inlet parameters are controlled by the operators. If the amount of ore is increased and the water is decreased, the ball mill is under-loaded; if the ore is reduced and the water is increased, the ball mill is over-loaded. |
| Water supplement           | t/h      | electromagnetic flow meter                | It is to adjust the concentration of the pump pool and controlled by the operators. If the supplementary water increases, the pulp concentration at the mill outlet is relatively large, and the ball mill is over-loaded; otherwise, it is under-loaded. |
| Liquid level in the pump pool | m        | Ultrasonic liquid level gauge             | It shows the discharge flow of ball mill outlet. If the liquid level decreases, the flow is reduced, and the ball mill is under-loaded; otherwise, it is over-loaded. |
| Overflow volume            | t/h      | electromagnetic flow meter                | It shows the the pulp flow at the cyclone outlet. If the overflow is large, the grinding efficiency is high and the ball mill is under-loaded; otherwise it is over-loaded. |
| -200 mesh percentage       | %        | Particle sizer                            | It shows the concentration of the slurry at the cyclone outlet. If the -200 mesh percentage increases, the grinding efficiency is high and the ball mill is under-loaded; otherwise it is over-loaded. |
| Motor power                | kW       | Power meter                               | It shows the energy consumption of the ball mill. In a certain range, the motor power will increase with the rise of the load, and when the load increases to a critical point, the load rise will lead to the decline of motor power. |

Figure 2 The changes of relevant parameters in one day
\[ L_k = \sum_{i=1}^{n} \alpha_i \left[ \phi(x_i) \cdot \phi(x) \right] \]

(1)

Where \( \alpha \) is the feature vector; \( K \) is the kernel function; the value of \( k \) is determined by the principal component contribution ratio.

A \( 7 \times k \)-order matrix \( L = [L_1, L_2, \ldots, L_7]^T \in \mathbb{R}^{7 \times k} \) could be obtained using KPC, which retained useful information of the original data while eliminating redundant information.

3.2. Improved K-means algorithm

The K-means algorithm seeks to split the observation data into several clusters. Data within the same cluster have higher similarity, while those in different clusters are less similar. The method first randomly selects \( K \) data objects as the initial clustering centers, calculates the distances between other objects and the clustering center and then assigns them into the nearest cluster. For a new iteration, it recalculates the average value of each cluster based on the previous assignment and uses it as the clustering center for iteration cluster until the criterion function converges or the algorithm reaches a pre-defined iteration limits. The disadvantages of K-means algorithm include sensitivity to noise and initial clustering center selection, and that the cluster number needs to be determined in advance [11, 12]. In this paper, the algorithm was improved based on density and genetic algorithms to obtain stable and reliable clustering results.

The specific steps of the improved K-means algorithm are as follows:

1) Apply the optimal division of hill-climbing algorithm [13] to divide the data space into \( s \) regions;

2) Calculate the average density \( Di \) and density coefficient \( \delta_i \) of each area separately, determine the high-density area, and use its number as the K value, namely,

\[ D = n / s \]

\[ \delta_i = D_i / D \]

(2)

(3)

Where \( n \) is the total number of data objects; \( D_i \) is the actual density of the \( i \)-th area, i.e., the number of data objects in the area.

3) Calculate the distance between the local density \( \rho_i \) and high local density points \( h_i \) of each object separately. If the object \( \rho_i \) is small but \( h_i \) is very large, it is judged to be a noise point, and classified into a dataset named \( S_1 \), and the remaining objects into another dataset named \( S_2 \), hence,

\[ \rho_i = \sum_j \chi(d_j - d_i) \]

\[ h_i = \min_{j \neq i, \rho_j} (d_j) \]

(4)

(5)

4) Set parameters such as the selection strategy of genetic algorithm, population size \( L \), mutation probability \( P_m \), cross probability \( P_c \), genetic algebra \( MaxGen \) to solve the minimum circumscribed circle center \( c_i \) of the \( i \)-th high-density region as the initial clustering center point and define the fitness Function as

\[ f_i = f(O_i) = 1 / r(a_i, b_i) \]

(6)

Where \( O_i = (a_i, b_i) \) is the genetic chromosome, gene \( a_i \), \( b_i \) is the coordinate of the circumscribed circle center; and \( r(a_i, b_i) \) is the objective function that satisfies the minimum circumscribed circle condition, so,

\[ r(a_i, b_i) = \max_{i \in \text{size}} \sqrt{(x_i - a_i)^2 + (y_i - b_i)^2} \]

(7)

5) Calculate the Euclidean distance between each point in \( S_2 \) and the current clustering center, and classify them into the cluster with the smallest distance;
\[
\text{dist}(x_i, c_j) = \sqrt{\sum_{d=1}^{d} (x_{i,d} - c_{j,d})^2}
\]  
(7)

where \(d\) is the data space dimension.

6) Calculate the mean of the current \(i^{th}\) cluster

\[
c_{i,m} = \frac{\sum_{x_i \in C_i} x_i}{|C_i|}
\]

(8)

7) Calculate the sum of squared errors of the current clusters

\[
J_T = \sum_{i=1}^{K} \sum_{x_i \in C_i} \text{dist}(x_i, c_i)^2
\]

(9)

8) Repeat steps 5) to 8) until the clustering center no longer changes or \(J_T\) no longer converges; otherwise, use \(c_{i,m}\) as iteration clustering center and go to step 5);

9) Calculate the Euclidean distance between each object and the clustering center point in \(S1\), and classify them into the nearest cluster;

10) Output the final clustering center point \(c_i\);

The improved algorithm, through density judgement, overcomes the shortcomings of traditional K-means clustering algorithm, namely, dependent on artificial experience and difficult to select \(K\) value. At the same time, it uses genetic algorithm to calculate the minimum circumscribed circle center of high-density area, which tackles the problems of excessive dependence on selecting initial clustering center point and hence the algorithm instability. The improved method is especially suitable for planar data and is easier to find the convex characteristics of the data and improve the clustering effect.

3.3. Load identification based on improved K-means

The paper proposed to identify the mill load status based on KPCA and improved K-means clustering, integrated effectively related parameters of different load. The specific steps are as follows:

1) Obtain 7 parameters in Table 1 as the original data set when the ball mill is in dynamic equilibrium through the DCS system of the concentration plant.

2) Apply KPCA to reduce the dimension of the 7 parameters to obtain their principal component values;

3) Apply the entropy weight method to calculate the weights of the parameters as follows:

\[
w_j = \frac{1 - E_j}{\sum_{j=1}^{d}(1 - E_j)}
\]

(10)

Where \(E_j\) is the information entropy value of parameters.

4) Establish the data space with the principal component value as the ordinate and the weight as the abscissa, and input the data to the improved K-means clustering to obtain the coordinates of the clustering center;

5) Analyze the displacement of the clustering center in the data space to identify the load status of the ball mill.

4. CASE ANALYSIS

4.1. Dataset Testing

The numerical dataset Iris, most commonly used for clustering algorithm performance testing in the UCI dataset, was first adopted to compare the traditional K-means algorithm, the K-means algorithm based on the maximum density [14], and the improved K-means algorithm of this paper in order to verify the effectiveness of the proposed algorithm.
The Iris dataset contained 150 pieces of numeric-type data, each of which had a dimension attribute of 4, which was divided into 3 categories. Data 1 to data 50 were Iris-setosa, 51 to 100 were Iris-versicolor, and 101 to 150 were Iris-virginica.

The Iris dataset was reduced to 2 dimensions using KPCA for clustering, and the dimension-reduced data were used as test samples. The clustering accuracy (CA), clustering similarity (CS), and iteration number \( I \) were used as indicators. The three algorithms were executed 10 times each. The average clustering results are shown in Table II after measuring the algorithm performance. Among them, CA was the percentage of correct clustering for each sample; CS was the similarity after using the same clustering algorithm for the same data set. The higher the similarity, the more stable the algorithm. This paper used the Pearson correlation coefficient to obtain the coordinates of the clustering center of each clustering result as the similarity index. Number \( I \) was the number of iterations to obtain the final result after determining the initial clustering center points, which was used for algorithm efficiency measurement.

| Algorithm          | Correct identification number | CA    | CS    | Iteration number \( I \) |
|--------------------|--------------------------------|-------|-------|--------------------------|
| K-means            | 70                             | 140   | 123   | 0.820                    | 0.7163       | 11               |
| Reference [14]     | 129                            | 138   | 134   | 0.893                    | 0.8527       | 2                |
| Improved K-means   | 132                            | 143   | 138   | 0.917                    | 0.9089       | 2                |

Table 2 shows that the improved K-means algorithm had higher accuracy in clustering, and could accurately determine the location of the initial clustering center point. It avoided the problems of unstable accuracy and efficiency caused by random selection of the initial clustering center of the traditional K-means algorithm and improves the similarity. Meanwhile, the improved algorithm appropriately utilized the relationship between the clustering center point and the data sample density, which greatly reduced the iteration number and algorithm operation time, and improved the algorithm efficiency. In addition, the improved algorithm eliminated the problem of determining \( K \) value based on artificial experience. It utilized the density and optimal partition to determine the \( K \) value completely based on the spatial distribution characteristics of the data samples and the \( K \) value was the same as the real results. This algorithm was proved to be valid, since it solved the shortcomings of lacking scientific basis and inaccurate results when the \( K \) value was determined by human experience if the prior conditions were unknown.

4.2. Measured data analysis

Seven parameters of historical data, including ore supply, water supply, water supplement volume, liquid level of pump pool, overflow volume, -200 mesh percentage and motor power from September 22 to October 19, 2018 for a total of 28 days were acquired. 60 sets of data with a duration of 10 mins under different load status were selected, and the proposed algorithm was used for analysis.

Since the DCS system of the concentration plant updated the data every 10s, the original dimension of each set of data was 60×7. The KPCA principal component contribution rate was set to be 0.99, and RBF kernel function was adopted, whose parameter satisfied \( 2\sigma^2 = 10000 \). KPCA was used to reduce the dimension of the original data. Part of the dimension reduction results after KPCA are shown in Table 3. The parameter weights obtained by the entropy weight method are shown in Table 4.

With the weight as the abscissa and the principal component value as the ordinate, the distribution of 60 sets of data in this two-dimensional space is shown in Fig.3. The improved K-means algorithm was used to obtain the clustering center of each data set. The results are shown in Fig.4.

Fig.4 shows that the clustering center of the parameters under different load conditions changed. Among the 60 sets of data, (0.1525, -0.0113) was the lowest the clustering center, which corresponded to the 53rd set of data out of the 61 sets of data and to 19:10-19:20 on October 10 when the mill was under-loaded; (0.1525, -0.0095) and (0.1525, -0.0094) had the highest proportion among the 61 sets of
data, which was in line with the reality that the ball mill was under ideal load when controlled by the operators. The second set of data with a clustering center of (0.1525, -0.0095) corresponded to 15:00-15:10 on September 22 when the ball mill had a normal load; (0.1525, -0.0088) was the highest clustering center, which corresponded to the 55th data and to 22:00-22:10 on October 10 when the ball mill was over-loaded.

Many complex random factors exist in load data, making it difficult to interpret and accurately classify the load status. This paper proposed to use an improved K-means clustering algorithm to cluster 7 load-related parameters. As for each clustering center, the displacement variation of the cluster centers can be obtained by comparing the load status of each set of data for effective identification of the load status changes.

| No. | PC 1     | PC 2     | PC 3     | PC 4     | PC 5     | PC 6     | PC 7     |
|-----|----------|----------|----------|----------|----------|----------|----------|
| 1   | -0.0084  | -0.0137  | -1.0107  | -1.0054  | -0.0018  | -0.0143  | 0.644    |
| 2   | -0.008   | -0.0136  | -0.0105  | -0.0154  | -0.0026  | -0.0143  | 0.645    |
| 3   | -0.0094  | -0.0136  | -0.0102  | -0.0151  | -0.0023  | -0.0139  | 0.644    |
| 4   | -0.0097  | -0.0135  | -0.0106  | -0.015   | -0.0017  | -0.0139  | 0.644    |
| 5   | -0.0077  | -0.0136  | -0.0109  | -0.0153  | -0.0028  | -0.0141  | 0.0645   |
| 6   | -0.0075  | -0.014   | -0.0114  | -0.0159  | -0.0003  | -0.0149  | 0.0641   |
| 7   | -0.0078  | -0.0138  | -0.0114  | -0.0157  | -0.0007  | -0.0147  | 0.0642   |
| 8   | -0.0087  | -0.014   | -0.0113  | -0.0155  | -0.0004  | -0.0145  | 0.0642   |
| 9   | -0.01    | -0.014   | -0.0107  | -0.0153  | 0.0002   | -0.0143  | 0.0641   |
| 10  | -0.0099  | -0.0137  | -0.0101  | -0.015   | -0.002   | -0.0138  | 0.0645   |

| parameter | Ore supply | Water supply | Water supplement | Liquid level | Overflow volume | -200 mesh percentage | power |
|-----------|------------|--------------|------------------|-------------|-----------------|----------------------|-------|
| Entropy value $E$ | 0.9873 | 0.9803 | 0.9177 | 0.9217 | 0.9762 | 0.9306 | 0.9747 |
| Weight $w$ | 0.0408 | 0.0632 | 0.2642 | 0.2513 | 0.0765 | 0.2227 | 0.0812 |

Figure 3 The spatial distribution of industrial data
5. Load Status Identification System Based on LabVIEW

According to the algorithm proposed in this paper, a load status identification system for ball mills based on LabVIEW was established. The overall block diagram is shown in Fig.5.

Inlet load parameters, such as ore supply and water supply, and outlet load parameters including overflow volume and particle size were collected by sensors such as electronic belt scales, electromagnetic flow meters, and particle sizers in the DCS system. The field analog-to-digital conversion used 4 Siemens IO modules, namely, DI, DO, AI and AO. Among them, the AI and AO modules were connected to the field through an isolation barrier, and the DI and DO modules were connected to the field through an intermediate relay, which greatly enhanced the stability of the transmission signal and the anti-interference ability. The parameters converted from analog to digital were collected and transmitted to the host computer system by myDAQ of American NI company. The dimensions of each load parameter were reduced through KPCA in the host computer system, and the entropy weight method is used to obtain their weights, and finally the corresponding load status.

The system can identify the ball mill load status of industrial DCS data and display the status change. In addition, it has the functions of data storage, query and alarm and a good human-machine interface. The status identification interface is shown in Fig.6.
6. CONCLUSION

This paper proposed to identify the load status of the ball mill based on the displacement of the clustering centers of the load-related parameters. It improved the K-means algorithm to make the clustering results stable and reliable. The case analysis results showed that the improved K-means algorithm greatly reduced the number of iterations and improved the stability. The proposed method selected the $K$ value based on the density and the genetic algorithm to determine the initial clustering center. It was adopted in the clustering of 7 load-related parameters such as ore supply, water supply, water supplement, liquid level of the pump pool, overflow volume, -200 mesh percentage, and motor power. The results effectively characterized the changes of load status. The results could be further analyzed and applied to industry for accurate monitoring and identification of the ball mill load status.

ACKNOWLEDGMENT

Thanks to Zhouhan Lin at MILA (Montreal Institute for Learning Algorithms) for helpful discussions with us about this work. This paper is Supported by the National Natural Science Foundation of China (Grant No. 51777061), the State Key Laboratory Foundation of China (BGRIMM-KZSKL-2017-02) and the Key Research and Development Project of Changsha, China(kq1901029).

REFERENCES

[1] Varas, Richard A. Cartes, and M. Aníbal Valenzuela. "Empirical determination of the effect of lifter wear in mill power for dry grinding," IEEE Transactions on Industry Applications 53.3 (2017): 2621-2627.

[2] Zhou, Xinjian, Nengwen Zuo, and Huicai Su. "Study on intelligent control for ore grinding and grading process," 2010 IEEE International Conference on Progress in Informatics and Computing. Vol. 1. IEEE, 2010.

[3] Cui, Bao-xia, Xing-yu Qu, and Yong Duan. "Study of BBD Ball Mill Material Measure Based on Rough Sets and RBF Neural Network Data Fusion," 2009 International Conference on Intelligent Human-Machine Systems and Cybernetics. Vol. 1. IEEE, 2009.

[4] Zhou, Ping, Tianyou Chai, and Hong Wang. "Intelligent optimal-setting control for grinding circuits of mineral processing process," IEEE Transactions on Automation Science and Engineering 6.4 (2009): 730-743.

[5] Tang J, Zheng X, Chai T. "Multi-source data fusion soft-sensor method for ball mill operating parameters based on transformation of radial basis function," 2010 8th World Congress on Intelligent Control and Automation. IEEE, 2010.

[6] Cui B, Li R, and Duan Y. "Study of BBD ball mill load measure method based on rough set and NN information fusion," 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application. Vol. 1. IEEE, 2008.
[7] L Tian, D Zeng, and X Liu. "Data Merging-based Determination of the Optimal-load Operating Point of a Ball Mill," Journal of Engineering for Thermal Energy and Power, 19.2(2004): 198-201.

[8] Bhaumik, Arup, Suman Banerjee, and Jaya Sil. "Designing of intelligent expert control system using PetriNet for grinding mill operation," WSEAS Transactions on Information Science and Applications 2.4 (2005): 360-365.

[9] Chai, Tianyou, S. Joe Qin, and Hong Wang. "Optimal operational control for complex industrial processes," Annual Reviews in Control 38.1 (2014): 81-92.

[10] Hu, Q, Qin, A, Zhang, Q, He, J, and Sun, G. "Fault diagnosis based on weighted extreme learning machine with wavelet packet decomposition and KPCA," IEEE Sensors Journal 18.20 (2018): 8472-8483.

[11] Kanungo, T, Mount, D. M, et al. "An efficient k-means clustering algorithm: Analysis and implementation," IEEE transactions on pattern analysis and machine intelligence 24.7 (2002): 881-892.

[12] Peng, Kai, Victor CM Leung, and Qingjia Huang. "Clustering approach based on mini batch kmeans for intrusion detection system over big data," IEEE Access 6 (2018): 11897-11906.

[13] Zhang, J. P, Yang, Y, Yang, J, and Zhang, Z. B. "Algorithm for initialization of K-means clustering center based on optimized-division," Journal of system simulation 9 (2009): 033.

[14] J Chen, H He. "Research on Density-based Clustering Algorithm for Mixed Data with Determine Cluster Centers Automat," Acta Automatica Sinica, 2015, 41(10): 1798-1813.