Research on State Evaluation and Prediction System of Hydraulic Turbine Based on Pressure Pulsation Parameters

(ID 60377)

Xiaoping Jiang¹, Xiang Gao¹,* , Jiawei Hu¹, Ziting Wang¹, Lele Wang¹

¹ School of Mechanical Electrical and Information Engineering, China University of Mining and Technology, Beijing No.Ding-11, College Road of Haidian district, Beijing, 100083, China
* Corresponding author. Tel: +8618534410860, 18534410860@163.com

Abstract: The pressure pulsation parameter of hydraulic turbine is one of the important indexes reflecting hydraulic stability of hydraulic turbine. Based on the pressure pulsation of the turbine, the health condition of the turbine unit is evaluated and predicted in this paper. Based on the measured data of pressure pulsation at four characteristic position of the hydraulic turbine, clustering was conducted by k-means method, and the health status of the hydraulic turbine unit was evaluated by finding the optimal k value for health status classification. The time series relationship of pressure pulsation was determined, and the evaluation result was taken as the prediction feature to establish the trend prediction model of pressure pulsation of hydraulic turbine set based on the time series change of long and short term memory network (LSTM). To realize real-time health assessment and prediction of unit running state and achieve quantitative assessment of equipment health state can provide technical basis for guiding unit maintenance.

Key word: Hydraulic turbine, Pressure pulsation, Clustering, health condition, Time series, Trend prediction;

1. Introduction

The pressure pulsation of the hydraulic turbine [1-5] is one of the important indicators that reflect the hydraulic stability of the turbine, and reducing the pressure pulsation level is also one of the basic requirements of hydraulic turbine unit design. The pressure pulsation of the hydraulic turbine will not only cause the vibration of the unit and the swing of the output, but also cause cracks in the blades and tear of the tail pipe wall. In severe cases, it can even cause resonance of hydraulic units and hydraulic structures, which directly threatens the safe and stable operation of the power station. Therefore, the evaluation and prediction of pressure fluctuations are of great significance.

The literature [6] proposed a method based on neural network prediction and support vector machine prediction to predict the vibration trend of hydraulic turbines to aim at the study of pressure
pulsation; Literature [7] uses time series analysis to identify the dynamic parameters of the hydraulic turbine blades; Literature [8] proposed the use of data mining and time series prediction methods to predict abnormal faults of hydraulic turbines.

In this paper, the neural network prediction method in deep learning is applied to the prediction of the health status of hydropower units. Based on the values of historical pressure pulsation parameters obtained in practical engineering applications, a neural network method prediction model is established to stress the operating state of the hydraulic turbine. The pulsation parameters are predicted, and then applied to the actual prediction and judgment of the health status of the hydropower unit, which plays an important role in monitoring the abnormal state of equipment and judging the potential abnormality of the unit.

2. Classification of Turbine Health Based on Pressure Pulse Harm

2.1. Analysis of pressure pulsation data set

Through the analysis of research status at home and abroad, it can be seen that the pressure pulsation parameters are of great significance to the health assessment of hydraulic turbines. Therefore, this paper selects the pressure pulsation at four key positions of the turbine's draft tube vortex, Karman vortex, labyrinth valve, and blade as the characteristics. It can be seen from Figure 1 that the pressure pulsation of the turbine is relatively stable, and all the data are healthy sample data under similar working conditions.

Based on the measured data of pressure pulsation at each characteristic location, the health status samples of the turbine unit are classified based on k-means, and the pressure pulsations at the four characteristic locations are classified according to the clustering results, which are divided into A, B, C, and D four grades, according to which the influence degree of the pressure pulsation on the hydraulic turbines unit can be judged.

Since the hydraulic turbine samples obtained in this paper are all healthy samples, the processing of fault samples requires healthy samples to identify abnormal samples, so a health model of characteristic quantities is established. The model is based on k-means. For a newly input sample point, calculate the distance from each cluster center to determine whether it is an abnormal point. In addition, because the sample data is the data under the same working condition, the data of the new sample point should also be the data under this working condition, otherwise it cannot be judged whether it is abnormal. For healthy samples under different working conditions, it needs to be modeled and processed separately.

A preliminary analysis of the data is carried out. In the time series, the pressure pulsation parameters of the four different characteristic positions of the hydraulic turbine under full power are measured, and the instantaneous pressure curve is obtained using 512Hz as the sampling frequency. The relationship between pressure pulsation data and time at four different characteristic locations is shown in Figure 1.
2.2. Analysis of the classification model of the health status of hydraulic turbine units

According to the pressure pulsation data collected at four different locations on the hydropower station turbine, the data is analyzed and a quantitative classification model based on data analysis is established. The health status classification flowchart is shown in Figure 2. This paper will use K-means clustering method to quantify and classify it, and use the elbow diagram method and the contour coefficient method to determine the optimal K value.

The core index of the elbow method is the sum of the squared errors (SSE).
\[ SSE = \sum_{i=1}^{k} \sum_{p \in C_i} |p - m_i|^2 \]  

(1)

Among them, \( C_i \) is the \( i \)th cluster, \( p \) is the sample point in \( C_i \), \( m_i \) is the centroid of \( C_i \) (mean of all samples in \( C_i \)), and SSE is the clustering error of all samples, which represents the quality of the clustering.

Using the elbow method to select the optimal cluster number \( k \) for the structured pressure pulsation data. The specific method is to let \( k \) take the value from 1 until the upper limit you think is appropriate (in general, the upper limit is not too large, here we choose the upper limit to 8), cluster each \( k \) value and note the SSE, then draw the relationship between \( k \) and SSE, as shown in Figure 3, and finally select \( k \) corresponding to the elbow as our best cluster number.

![Elbow diagram of pressure pulsation data](image)

The core index of the contour coefficient method is the Silhouette Coefficient. The contour coefficient of a sample point \( X_i \) is defined as follows:

\[ S = \frac{b - a}{\max(a, b)} \]  

(2)

Among them, \( a \) is the average distance between \( X_i \) and other samples in the same cluster, called the degree of aggregation, and \( b \) is the average distance between \( X_i \) and all samples in the nearest cluster, called the degree of separation. And the nearest cluster is defined as

\[ C_j = \arg \min_{C_i} \frac{1}{n} \sum_{p \in C_i} |p - X_i|^2 \]  

(3)

Where \( p \) is a sample in a certain cluster \( C_i \). In fact, the average distance from all samples of a cluster to the cluster is used to measure the distance from the point to the cluster, and then the cluster closest to \( X_i \) is selected as the closest cluster.

On the other hand, use a structured pressure pulsation data set, consider the case where \( k \) is equal to 1 to 8, perform clustering for each \( k \) value and find the corresponding contour coefficient, and then make a relationship between \( k \) and contour coefficient, as shown in Figure 4. As shown in the figure, \( k \) with the largest contour coefficient is selected as our best clustering coefficient.
It can be seen from the elbow diagram that when the k value corresponding to the elbow is between 3 and 5, the diagram shows an obvious turning point, that is, the corresponding elbow corresponding point in the elbow diagram. Therefore, it can be judged from the elbow diagram method that the optimal k value at this time is between 3 and 5; from the contour coefficient diagram, it can be seen that when k is 2, the maximum clustering effect of contour coefficient is the best, but when k is 2 there are only two types of results: healthy state and unhealthy state, which are not consistent with the healthy state classification model, and the corresponding SSE coefficient is relatively large, that is, clustering is meaningless when k is 2. Since k is 2 meaningless, the optimal k value is chosen between 3 and 5. According to the above clustering results, the compactness index (CP), spacing index (SP), Davis Paulding index (DB), Dunn index (DVI) at k=3, 4, and 5 are calculated respectively, as shown in Table 1 below.

| k | CP     | SP     | DB      | DVI     |
|---|--------|--------|---------|---------|
| 3 | 0.0959 | 0.0398 | 2.7029  | 0.0426  |
| 4 | 0.0606 | 0.1098 | 2.5074  | 0.5498  |
| 5 | 0.1493 | 0.0837 | 7.6902  | 0.1674  |

The clustering results of k=3, k=4, and k=5 are evaluated and compared by the compactness index (CP), interval index (SP), Davis Paulding index (DB), and Dunn index (DVI). When k is 4, the compactness index (CP) is lower than k=3, indicating that the intra-class distance is closer when k is 4; the spacing index (SP) is higher than k for 3 and 5, indicating that k is 4 When k=4, the Dunn index (DVI) is significantly higher than the DVI with k=3 and 5. Therefore, the optimal number of clusters is 4.

Decompose the pressure pulsation clustering results of the four-dimensional features, as shown in Figure 5 below. The diagram decomposes the four-dimensional pressure pulsation clustering results into 12 two-dimensional graphs. Red, blue, yellow and green each represent one class. In the figure, the abscissa is the pressure pulsation parameter of the four characteristic positions, and the corresponding ordinate is the pressure pulsation parameter of the other three positions. Finally, the optimal cluster number is determined to be 4, that is, the health status of the turbine unit is divided into four categories: A, B, C, and D.
According to the clustering results, combined with the pressure pulsation data, draw the corresponding categories of the pressure pulsation data. To facilitate graphical observation, take the first 40 sets of data for analysis, and convert the four categories A, B, C, and D into 1, 2, 3, 4 categories, in order to facilitate the drawing, take the data range of each position and scale the value from 1 to 4. The schematic diagram is shown in Figure 6.

The classification of this section of sample data is mainly determined by the Karman vortex in figure 2, and the pressure pulsation of the rest data is relatively stable. Regarding the pressure pulsation data at the Karman vortex, it can be clearly seen that when the sample data fluctuates slightly at 0-10 and 25-40, the pressure pulsation amplitude is relatively large, and the categories are C and B, 10-15, 30 respectively. The pressure pulsation amplitude at ~35 is small, and the category is A at this time, and the pressure pulsation amplitude is large at 15-30, and the category is D at this time. Therefore, it can be judged that the category is from A to D, and the greater the amplitude of the pressure pulsation, the greater the impact on the running state of the turbine.

Since the pressure pulsation data measured in this article are healthy samples under similar working conditions, the characteristic quantity health model is established to better identify the "abnormal" samples under the working condition. The Euclidean distance is used to calculate the distance between the data. By calculating whether the Euclidean distance between the new sample and the cluster center
is greater than the maximum distance of its category, it is judged whether the unit operation is abnormal.

![Schematic diagram of "abnormal" sample points](image)

**Fig. 7 Schematic diagram of "abnormal" sample points**

As shown in Figure 7, when the Euclidean distance between the new input sample point and the cluster center is greater than the maximum distance within the class, it can be determined as an "abnormal" point, which can be converted into a fault sample point with artificial assistance and saved in the fault sample library.

3. **Forecasting the health status of the turbine**

3.1. **Prediction of health status based on neural network**

The BP (Back Propagation) network was proposed by a team of scientists led by Rumelhart and McCelland in 1985. It is a multi-layer feedforward network trained by the error back propagation algorithm and is currently one of the most widely used neural network models. The network consists of input layer, hidden layer and output layer nodes. The hidden layer can be one layer or multiple layers. Neurons are represented by nodes, and the nodes of the front layer to the nodes of the rear layer are connected by weights. The neurons between the hidden layers are not connected.

Since the weights and biases of the hidden layer of the BP neural network, and the weights and biases of the output layer are randomly given, the results of each training are slightly different. Through continuous multiple training, adjust the number of layers, hidden layer nodes, learning rate and other hyperparameters. Until the trained loss function converges to the expected error, then save the model. Put the test set into the model for prediction and get the prediction result. In order to facilitate the observation and analysis of the prediction results of the BP neural network on the health status of the turbine unit, the first 40 test samples are taken as shown in Figure 8. The blue "×" represents the true value, the red "○" represents the value predicted by the BP neural network model, and the health status levels 1 to 4 represent the clustering results A, B, C, and D respectively. The table corresponding to Figure 7 is shown in Table 2.
As shown in Table 3 below, the prediction effect of the BP neural network model is evaluated according to the root mean square relative error, root mean square error, average absolute error and average relative error.

| Sample number | True value | BP Predicted value | Sample number | True value | BP Predicted value |
|---------------|------------|---------------------|---------------|------------|---------------------|
| 1             | 3          | 3.0821              | 21            | 2          | 2.0104              |
| 2             | 3          | 3.1232              | 22            | 2          | 2.0043              |
| 3             | 3          | 3.1375              | 23            | 2          | 2.0038              |
| 4             | 3          | 3.2284              | 24            | 2          | 1.9830              |
| 5             | 3          | 3.1866              | 25            | 2          | 1.9600              |
| 6             | 3          | 3.2261              | 26            | 2          | 1.9504              |
| 7             | 3          | 3.2422              | 27            | 2          | 2.0103              |
| 8             | 3          | 3.2216              | 28            | 2          | 2.1070              |
| 9             | 3          | 3.0801              | 29            | 3          | 2.4217              |
| 10            | 3          | 2.7459              | 30            | 3          | 2.8219              |
| 11            | 3          | 2.5887              | 31            | 3          | 3.0060              |
| 12            | 2          | 2.3437              | 32            | 3          | 2.9592              |
| 13            | 2          | 2.1073              | 33            | 3          | 2.9841              |
| 14            | 2          | 2.0682              | 34            | 1          | 2.5507              |
| 15            | 2          | 2.0328              | 35            | 1          | 1.6338              |
| 16            | 2          | 2.0140              | 36            | 1          | 1.0987              |
| 17            | 2          | 2.0194              | 37            | 1          | 0.9967              |
| 18            | 2          | 2.0375              | 38            | 1          | 0.9885              |
| 19            | 2          | 2.0319              | 39            | 1          | 0.9928              |
| 20            | 2          | 2.0342              | 40            | 2          | 0.9858              |

As shown in Table 3 below, the prediction effect of the BP neural network model is evaluated according to the root mean square relative error, root mean square error, average absolute error and average relative error.

| Prediction object | Data Classification | MSE  | RMSE | MAE  | MAPE |
|-------------------|---------------------|------|------|------|------|
| Health status of turbine | Training phase     | 0.8241 | 0.7241 | 0.7241 | 1.2542 |
|                     | Prediction stage    | 0.8487 | 0.5721 | 0.5911 | 1.2461 |

It can be seen from Table 3 that the values of root mean square error (RMSE), mean absolute error (MAE), and root mean square relative error (MSE) are relatively small in the training phase and the prediction phase, and the model has certain prediction effect, but due to the large average relative error.
error (MAPE), it can be analyzed from Figure 8 that the accuracy of the local position model prediction is high, but there are large deviations in other places. The main reason may be caused by overfitting of BP neural network model training.

3.2. Health state prediction based on long and short-term memory network

LSTM (long short term memory) is a special recurrent neural network with a chain form of repetitive neural network modules. Remembering long-cycle useful information is its basic function. Therefore, this paper uses LSTM time series algorithm to solve the problem of time series prediction in this paper. Similarly, a four-dimensional characteristic pressure pulsation data set is used as a training set for a time series prediction model, and the result of the above clustering is used as a target for time prediction.

LSTM is an improved RNN. Each LSTM unit has a cell, and its state $h_t$ at time $t$ is called the cell state. The top level of the LSTM model defines a conveyor belt, which is used to describe the cell state and memorize the current network information. The conveyor belt selects and memorizes the information through three control gates, which are input gate, output gate, and forget gate. The switching state of the door is affected by the input layer information and the hidden layer information at the previous moment.

![Fig.9 Schematic diagram of LSTM structure](image)

After the current input and the output at the previous moment pass through the sigmoid function, the output $h_{t-1}$ is converted to a number between 0 and 1, and then multiplied by $C_{t-1}$.

$$f_t = \sigma(o_f \cdot [h_{t-1}, x_t] + b_f)$$

In the above formula, $o_f$, $b_f$, $\sigma$ represents the weight of the forget gate, the offset of the forget gate, and the sigmoid function.

The model of the input door save information is as follows:

$$i_t = \sigma(o_i \cdot [h_{t-1}, x_t] + b_i)$$

In the above formula, $o_i$, $b_i$ respectively represents the weight of the input gate and the offset of the input gate.

For the output gate, it controls the effect of long-term memory on the current output $o_t$,

$$o_t = \sigma(o_o \cdot [h_{t-1}, x_t] + b_o)$$

The final output of LSTM is determined by the output gate and unit state:

$$h_t = o_t \ast \tanh(c_t)$$
Fig. 10 Instantaneous pressure pulsation history data

Fig. 11 Clustering results corresponding to pressure pulsation parameters
Tab.4 LSTM model training parameters table

| Parameter               | Value   |
|-------------------------|---------|
| Time series length      | 30      |
| Batchsize               | 50      |
| LSTM layers             | 1       |
| dropout                 | 0.7     |
| Learning rate           | 0.001   |
| Learning rate decay     | 0.95    |
| Epoch                   | 10000   |
| Enter the number of features | 4    |
| Expected error          | 1e-3    |

The whole model is trained 50,000 times, and a plot is taken every 200 times. The following loss function graph is obtained through training. From the figure, it can be seen that the loss function loss shows a decreasing trend, and the cross entropy is at least 1.03, and the training effect is obvious.

![LSTM training loss function loss](image1)

Save the trained model, extract the last 50 data of the pressure pulsation dataset as the test dataset, and get the following verification result map, where "x" is the health status value predicted by the trained model; "." is the real status value.
Since the weights and biases of the forget gate, input gate and output gate of LSTM are given randomly, the training results of each time are slightly different. Through continuous multiple training, adjust the number of LSTM layers, learning rate, and learning Rate attenuation, dropout. Until the training results converge to the expected error, then save the model. Put the test set into the model for prediction and get the prediction result. In order to facilitate the observation and analysis of the prediction results of the LSTM network on the health status of the turbine unit, the first 40 test samples are taken as shown in Figure 15 below. The blue "×" represents the true value, and the red "○" represents the value predicted by the LSTM model.
Tab.5 Prediction performance evaluation of neural network model

| Sample number | True value | BP Predicted value | LSTM Predicted value | Sample number | True value | BP Predicted value | LSTM Predicted value |
|---------------|------------|---------------------|----------------------|---------------|------------|---------------------|----------------------|
| 1             | 3          | 3.0821              | 2.9809               | 21            | 2          | 2.0104              | 1.9820               |
| 2             | 3          | 3.1232              | 2.9810               | 22            | 2          | 2.0043              | 1.9876               |
| 3             | 3          | 3.1375              | 2.9810               | 23            | 2          | 2.0038              | 1.9876               |
| 4             | 3          | 3.2284              | 2.9810               | 24            | 2          | 1.9830              | 1.9876               |
| 5             | 3          | 3.1866              | 2.9726               | 25            | 2          | 1.9600              | 1.9781               |
| 6             | 3          | 3.2261              | 2.9821               | 26            | 2          | 1.9504              | 1.9832               |
| 7             | 3          | 3.2422              | 2.8162               | 27            | 2          | 2.0103              | 1.9875               |
| 8             | 3          | 3.2216              | 2.9861               | 28            | 2          | 2.1070              | 1.9875               |
| 9             | 3          | 3.0801              | 3.0160               | 29            | 3          | 2.4217              | 2.9671               |
| 10            | 3          | 2.7459              | 3.0161               | 30            | 3          | 2.8219              | 2.9620               |
| 11            | 3          | 2.5887              | 2.0172               | 31            | 3          | 3.0060              | 2.9763               |
| 12            | 2          | 2.3437              | 2.0172               | 32            | 3          | 2.9592              | 2.9763               |
| 13            | 2          | 2.1073              | 2.0172               | 33            | 3          | 2.9841              | 2.9451               |
| 14            | 2          | 2.0682              | 2.0171               | 34            | 1          | 2.5507              | 1.0723               |
| 15            | 2          | 2.0328              | 2.0140               | 35            | 1          | 1.6338              | 1.0723               |
| 16            | 2          | 2.0140              | 2.0615               | 36            | 1          | 1.0987              | 1.0522               |
| 17            | 2          | 2.0194              | 2.0172               | 37            | 1          | 0.9967              | 1.0732               |
| 18            | 2          | 2.0375              | 2.0172               | 38            | 1          | 0.9885              | 1.0822               |
| 19            | 2          | 2.0319              | 1.9820               | 39            | 1          | 0.9928              | 1.0137               |
| 20            | 2          | 2.0342              | 1.9820               | 40            | 2          | 0.9858              | 1.9925               |

4. Conclusion analysis

As shown in Table 6 below, the prediction effect of the LSTM network model is evaluated according to the root mean square relative error, root mean square, average absolute and average relative error, and the prediction effects of the LSTM network model and the BP neural network model are compared.

Tab.6 Error index of LSTM network model training and prediction stage

| Prediction object | Prediction model | Data Classification | MSE  | RMSE  | MAE  | MAPE  |
|-------------------|-----------------|---------------------|------|-------|------|-------|
| Health status of turbine | BP neural network | Training phase | 0.8241 | 0.7241 | 0.7241 | 1.2542 |
| | | Prediction stage | 0.8487 | 0.5721 | 0.5911 | 1.2461 |
| | LSTM network | Training phase | 0.3612 | 0.5102 | 0.4871 | 1.0262 |
| | | Prediction stage | 0.3493 | 0.4986 | 0.4872 | 1.0187 |
It can be seen from Figure 15 that the prediction effect of the LSTM network model is significantly better than the prediction effect of the BP neural network model, and its predicted value is closer to the real value. Compare the error values of the LSTM model and the BP neural network model from Table 6. Found that the various error values of the LSTM model are also lower than the BP neural network model, indicating that the prediction accuracy of the LSTM model is higher, which solves the problem of BP neural network overfitting.

5. Conclusion
This paper mainly studies the health evaluation and prediction system of hydraulic turbine units based on pressure pulsation. According to the four groups of characteristic pressure pulsation parameters, based on finding the optimal K value and k-means clustering results, the characteristics of various types were evaluated and analyzed, and the clustering results were taken as the health state of the water turbine set, and training samples were provided for the prediction model. The deep learning model LSTM was used to establish a time series prediction model based on four-dimensional characteristic pressure pulsation, and the correlation between the time law of turbine measurement points and the measurement points was mined. Based on a large amount of training data, the network model was optimized and adjusted, and the equipment operation rule model under normal operation state was established to monitor the abnormal state of equipment. Finally, the feasibility and validity of the model are verified by the test.

References
[1] Zhou Yikui. Study on the pressure pulsation in hydraulic turbine condition of a pump turbine[D]. Harbin Institute of Technology, 2018
[2] Qian Zhongdong, Lu Jie, Guo Zhiwei, Zhang Jianjun. Characteristics of pressure fluctuation in pump-turbine under turbine mode, 2016, 34(08):672-678.
[3] Zhu Wenlong, Zhou Jianzhong, Xia Xin, et al. Diagnosis strategy of hydraulic turbine pressure pulsation based on operating conditions of hydropower units [J]. Vibration and Shock, 2015, 34(8): 26-30.
[4] Pan Luoping. Research on Hydropower Unit Fault Diagnosis System Based on Health Assessment and Deterioration Trend Forecast[D]. China Institute of Water Resources and Hydropower Research, 2013
[5] Gui Zhonghua, Le Zhenchun, Dong Yangwei, Tang Yongjun, Xu Hongquan, Chen Rui. Research on the Evaluation Method of Pressure Fluctuation of Francis Turbine [J]. Hydropower and Pumped Storage, 2018, 4(04):38-41.
[6] Zhan Pei. Prediction of vibration trend of hydropower units based on artificial intelligence [D]. 2016.
[7] Wu Peiqiang, Liu Yibing, ZHANGPei-qiang, et al. Dynamic parameter identification of turbine blades based on time series analysis [J]. Mining and Metallurgy, 2007, 16 (3): 91-94.
[8] Cui Weiyu. Research on abnormal warning method of hydraulic turbine based on data mining [D]. 2018.
[9] Liu Ling, Guan Xiujuan, Lv Yanguang. Study on the Method of Detecting Pressure Fluctuation in Draft Tube[J]. Large Electric Machine and Hydraulic Turbine, 2002(3):39-42.
[10] Han Xiaohong, Hu Yu. Research on K-means clustering algorithm [J]. Journal of Taiyuan University of Technology, 2009(03):33-36.0
[11] Kanungo, T, Mount, D.M, Netanyahu, N.S, etc. An efficient k-means clustering algorithm: analysis and implementation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(7):0-892.
[12] Ding C H Q , He X . Cluster Structure of K-means Clustering via Principal Component Analysis[C]/ Advances in Knowledge Discovery and Data Mining, 8th Pacific-Asia Conference, PAKDD 2004, Sydney, Australia, May 26-28, 2004, 61 Proceedings. DBLP, 2004.
[13] Hamerly G , Elkan C . Learning the K in K-Means[J]. Advances in Neural Information Processing Systems, 2003, 16(2):861-868.
[14] Aristidis Likas, Nikos Vlassis, Jakob J. Verbeek. The global k-means clustering algorithm[J]. Pattern Recognition the Journal of the Pattern Recognition Society, 2003, 36(2):451-461.
[15] Zhang Hao, Wang Qijie, Zhu Jianjun, et al. INFLUENCE OF SAMPLE DATA PREPROCESSING ON BP NEURAL NETWORK-BASED GPS ELEVATION FITTING[J]. Journal of Geodesy and Geodynamics, 2011(02):129-132
[16] Xia Shixiong, Li Wenchao, Zhou Yong, et al. An improved k-means clustering algorithm (English)Journal of Southeast University(English Edition), 2007(03):123-126.
[17] Wu Guangjian, Zhang Jianlin, Yuan Ding. Automatically Obtaining K Value Based on K-means Elbow Method[J]. Computer engineering & Software, 2019, 040(005):167-170.
[18] Yi Fan. Neural Network Forecast Research[D]. Southwest Jiaotong University, 2005
[19] Zhou Aiwu, Yu Yafei. The Research about Clustering Algorithm of K-Means[J]. Computer Technology and Development, 2011, 021(002):62-65.
[20] Wang Xin, Wu Ji, Liu Chao, et al. Exploring LSTM based recurrent neural network for failure time series prediction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018(4):772-784.
[21] Chen Chang, Li Xiaolei, Cui Weiyu. Hydraulic turbine operation status detection based on LSTM network prediction[J]. Journal of Shandong University(Engineering Science), 2019, 49(3).