Research on Fatigue Life Prediction Method of Ballastless Track Based on Big Data

Ailin Wang(✉)

Railway Engineering College, Wuhan Railway Vocational College
of Technology, Wuhan 430205, China
`gjakg785@yeah.net`

Abstract. In order to improve the precision of fatigue life prediction of ballastless track, a method for predicting fatigue life of ballastless track based on big data is proposed. The big data model is constructed to analyze the fatigue life cycle of ballastless track. Big data mining and feature extraction are used to extract the fatigue life cycle of ballastless track. Combining with the particle swarm optimization method, the feature classification of the failure state trend of ballastless track construction is carried out, and the information fusion is carried out according to the characteristic parameters of the failure state of ballastless track construction. The expert system model for predicting fatigue life of ballastless track construction is established and the fatigue life of ballastless track is predicted by association rule mining method. The simulation results show that the precision of fatigue life prediction of ballastless track is high, and the strength and life cycle of ballastless track are analyzed.

Keywords: Big data · Ballastless track · Fatigue life · Prediction

1 Introduction

With the development of high-speed railway in China, ballastless track is more and more used in new railway. The construction practice of high-speed railway shows that both ballastless track and ballastless track can ensure the safe operation of high-speed train, but because of the difference in technology and economy between the two kinds of track structure, they have their own advantages and disadvantages. Therefore, in order to obtain the best technical and economic benefits, the rail structure selection of our country should be reasonable according to our own national conditions and railway characteristics. The fatigue life of ballastless track is the key to determine the track quality [1]. It has great significance to study the fatigue life prediction method of ballastless track.

In order to ensure high comfort and high safety, high-speed railway requires high ride comfort to track geometry precision. Through the investigation and understanding of the high-speed railway lines that have been opened to traffic and are in the process of fine-tuning the long rail, in the early stage of high-speed rail construction, the tool-track method or rail-row frame method was used for the fine-tuning of the construction [2]. After the track plate construction is completed, the seamless long rail will be laid. In the long track fine adjustment, there are some precision overruns in the constructed track

G. Gui and L. Yun (Eds.): ADHIP 2019, LNICST 302, pp. 140–150, 2019.
https://doi.org/10.1007/978-3-030-36405-2_15

bed plate, such as gauge, elevation, center line, orbit direction, etc. In order to meet the requirements of high-speed railway for rail ride comfort, the accuracy problem in construction must be solved by replacing non-standard fasteners. If the precision control is not in place in the construction, the construction cost will be wasted greatly in the replacement of the fastener in the later period [3]. The short rail with the same specifications as the long rail must be used in the rail row processing, and the machining precision must be the same as that of the track, so as to ensure that the rail row can meet the requirements in the course of its use. The rail-row frame should have enough stiffness and enough stability to prevent the rail-row frame vibration deviation caused by the construction load, which has a negative effect on the track accuracy. Rail support system should be supported on a stable basis to prevent the sinking and dislocation of the supporting member [4].

The fatigue life prediction of ballastless track casting in digital machining mode is carried out. Combined with big data information processing and fatigue life characteristic detection method, the accuracy of fatigue life prediction for ballastless track casting is improved, at present, The fatigue life prediction methods for ballastless track casting in digital machining mode can be divided into two methods: time domain analysis based fatigue life data mining method, frequency domain analysis based fatigue life data mining method, and fatigue life data mining method based on frequency domain analysis [5]. The fatigue life prediction model of ballastless track casting based on statistical characteristic extraction, etc., combined with nonlinear time series analysis and signal detection algorithm, can predict the trend of fatigue life state of ballastless track casting. Based on data mining and feature extraction of fatigue life, this paper presents a fatigue life prediction model for ballastless track casting based on genetic KNN clustering. Firstly, big data association rule mining method is used to collect the fatigue life characteristic information of ballastless track casting, and then the feature classification of fatigue life state trend of ballastless track casting is carried out by combining particle swarm optimization (PSO) evolution method. According to the characteristic parameters of fatigue life of ballastless track casting, the expert system model for fatigue life prediction of ballastless track casting is established, and the fatigue life prediction of ballastless track casting is realized. Finally, simulation experiments are carried out to show the superior performance of this method in improving the accuracy of fatigue life prediction of ballastless track casting.

2 Fatigue Life Data Sampling and Feature Extraction of Ballastless Track Casting

2.1 Big Data Analytical Model for Fatigue Life of Ballastless Track Casting

The first step to predict the fatigue life of ballastless track casting is to construct the fatigue life model of big data. The fatigue life of big data is divided into vibration data according to the means of testing and the principle of diagnosis. The noise data and other fatigue life data of ballastless track casting, considering the input/output parametric model of complex ballastless track parts, the big data distribution channel of ballastless track casting fatigue life is an extended sampling channel. The multi-path channel model is used to collect the fatigue life of big data in ballastless track casting. There are two main characteristics of big data distribution channel of ballastless track casting fatigue life [6]. One is that big data distribution channel of ballastless track casting fatigue life is a channel model with limited bandwidth. Second, big data distribution channel of ballastless track casting fatigue life is constrained by distance, which has multi-path propagation characteristic and is easily disturbed by ballastless track vibration, which leads to fatigue life. In ballastless track casting, the multi-path structure of fatigue life big data distribution channel depends on the array distribution type of data acquisition nodes [7]. Assuming that the data acquisition nodes are composed of S array elements, the radial distance of big data distribution of ballastless track casting is d, and the data receiving model of fatigue life characteristics is as follows:

$$x_m(t) = \sum_{i=1}^{I} s_i(t) e^{j\varphi_{mi}} + n_m(t), -p + 1 \le m \le p \tag{1}$$

Where, $s_i(t)$ is the vibration data sensed by vibration sensor of ballastless track casting equipment, and $x_m(t)$ is the series of thermal sensing data received by element m. Thus, the impulse model of big data distribution channel for fatigue life of ballastless track casting is constructed as:

$$h(t) = \sum_i a_i(t) e^{j\theta_i(t)} \delta(t - iT_S) \tag{2}$$

In the above formula, $\theta_i(t)$ indicates that the data ETL layer provides radial deviation of fatigue life data for ballastless track casting to the data analysis layer, and the width of the time window for sampling vibration sensing information of ballastless track equipment is T, The quantized set of fatigue life data features distributed in the extended channel is represented as:

$$\boldsymbol{x}(t) = [x_{-P+1}(t), x_{-P+2}(t), \cdots, x_P(t)]_{N\times 1}^T \tag{3}$$

$$\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots s_I(t)]_{I\times 1}^T \tag{4}$$

Where, P is the bandwidth of big data's collection of fatigue life of ballastless track casting, and I is the number of array elements. Under the digital machining mode, big data analytical model of fatigue life distribution of ballastless track casting is expressed as follows:

$$c(\tau, t) = \sum_n a_n(t) e^{-j2\pi f_c \tau_n(t)} \delta(t - \tau_n(t)) \tag{5}$$

Where, $a_n(t)$ is the closed-loop management characteristic vector of the nth ballastless track fatigue life diagnosis, $\tau_n(t)$ is the time delay of the nth data channel, the big data analytical model of fatigue life distribution of ballastless track casting under the digital machining mode is constructed. It provides data input basis for fatigue life prediction of ballastless track casting.

2.2 Fatigue Life Feature Extraction

The abnormal state data of ballastless track during fatigue life cycle are decomposed and extracted by Hilbert spectrum extraction method, and the information fusion is carried out according to the extracted value of fatigue life state characteristic parameter of ballastless track casting. Because the cluster control parameter s_0, τ_0 of ballastless track fatigue life data is often unknown, the fatigue life characteristics of ballastless track casting are detected by using the lowest mean square error estimate $\hat{s}_{0ML}$, $\hat{\tau}_{0ML}$:

$$\begin{aligned} l_1(r) &= \int r(t)\sqrt{\hat{s}_{0ML}} f^*(\hat{s}_{0ML}(t-\hat{\tau}_{0ML}))dt \\ &= \max_{s,\tau}\left|\int r(t)\sqrt{s} f^*(s(t-\tau))dt\right| \\ &= \max_{a,b}\left|W_f r(a,b)\right| \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_1 \end{aligned} \tag{6}$$

In the equation, $a = 1/s, b = \tau$, λ_1 is the detection threshold. The statistic of fatigue life detection for ballastless track casting is a sparse array distribution. Because s_0, τ_0 are unknown, the following formulas are selected to extract fatigue life characteristics.

$$\begin{aligned} l_2 &= \max_{b}\left|\int r(t)\frac{1}{\sqrt{a'}} f^*(\frac{t-b}{a'})dt\right| \\ &= \max_{b}\left|W_f r(a',b)\right| \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_2 \end{aligned} \tag{7}$$

The optimal classification plane of fatigue life characteristics of ballastless track casting is calculated. Between the data analysis layer and the data processing layer, the fatigue life characteristic distribution sequence $x_1, x_2, \ldots, x_n, \ldots$, the total number of points is N, Ballastless track vibration data series $\{x_i\}$ is evenly sampled by sampling interval $j\tau$. The output autocorrelation function is expressed as follows:

$$R_{xx}(j\tau) = \frac{1}{N}\sum_{i=0}^{N-1} x_i x_{i+j\tau} \tag{8}$$

The characteristic function $\overrightarrow{X}(l, n_i)$ of fatigue life prediction for ballastless track casting under some kind of fatigue life state is used to calculate the intra-class dispersion matrix $\widehat{S}_w$ of fatigue life distribution. The optimal solution to big data's acquisition problem of fatigue life can be described by χ^*. By using the abnormal operation and maintenance data management method, the adaptive constraint characteristics are obtained as follows:

$$x(t) = \left[1 - \cos(2\pi f_s^{(r)} t)\right] [1 + A\cos(2\pi f_s t + \phi)] \\ \cos[2\pi f_m t + B\sin(2\pi f_s t + \varphi) + \theta] \tag{9}$$

The self-organizing training of big data in ballastless track casting fatigue life was carried out by multi-source information filtering method. The maximum gradient difference of fatigue life prediction of ballastless track casting was obtained.

$$AVG_{\mathrm{X}} = \frac{1}{m \times n} \sum_{x=1}^{n} \sum_{y=1}^{m} |G_{\mathrm{X}}(x, y)| \tag{10}$$

Where, m, n are vector quantized autocorrelation coefficients of fatigue life data digging for ballastless track casting respectively, so that the feature of fatigue life distribution can be extracted [8].

3 Optimization of Fatigue Life Prediction Algorithm

3.1 Big Data Clustering of Fatigue Life Characteristics

Because of the continuous generation of data over time, the data stored in the database increases exponentially, and the relationship between the data becomes more complex. Thus, it is more possible to get some of these relationships through association rules mining algorithm for those data that do not seem to have any connection. It is applied to the fatigue life prediction of ballastless track structure, and the expert system model of ballastless track structure fatigue life prediction is established. The association rule mining method is used to predict the fatigue life of ballastless track structure. Based on big data association rule mining method, the fatigue life characteristic information of ballastless track casting is collected, and the optimal design of fatigue life prediction model of ballastless track casting is carried out. This paper presents a fatigue life prediction model for ballastless track casting based on big data clustering. The feature classification of fatigue life state trend in ballastless track casting is studied by particle swarm optimization method [9]. In this paper, the method of mining correlation dimension feature is used to extract fatigue life category feature. It is assumed that there are n samples in the trend data set of fatigue life state of ballastless track casting, in which the characteristic vector of sample x_i, $i = 1, 2, \ldots, n$ is:

$$\boldsymbol{x}_i = (x_{i1}, x_{i2}, \cdots, x_{is})^T \tag{11}$$

The adaptive optimization of multi-dimensional data is carried out by using particle swarm optimization method, and the data set is divided into 2^n subsets. KNN fuzzy search is carried out in two-dimensional space, and the fuzzy clustering center matrix of

KNN clustering for fatigue life judgement of ballastless track casting is obtained as follows:

$$\boldsymbol{V} = \{v_{ij} | i = 1, 2, \cdots, c, j = 1, 2, \cdots, s\} \tag{12}$$

According to the feature similarity in the clustering process, the KNN fuzzy partition matrix for fatigue life prediction of ballastless track casting is obtained:

$$\boldsymbol{U} = \{\mu_{ik} | i = 1, 2, \cdots, c, k = 1, 2, \cdots, n\} \tag{13}$$

In the K-nearest neighbors of K-means clustering centers, the weight of fatigue life of each kind of ballastless track casting is calculated in turn. The formula is as follows:

$$P_{1J} = \sum_{d_i \in kNN} Sim(x, d_i) y(d_i, C_j) \tag{14}$$

The clustering objective function matrix of big data KNN for fatigue life of ballastless track casting is obtained as (defining clustering objective function):

$$\boldsymbol{J}_m(U, V) = \sum_{k=1}^{n} \sum_{i=1}^{c} \mu_{ik}^m (d_{ik})^2 \tag{15}$$

The recall characteristics of fatigue life characteristics of ballastless track are obtained by using K-valued particle swarm optimization control method and distributed clustering method: (1) the characteristics of fatigue life in ballastless track casting are as follows:

$$W = \frac{\overline{K}}{\gamma} = \frac{1}{\gamma} \sum_{k=1}^{K} \sum_{n=1}^{N} kp_{k,n} \tag{16}$$

The data conversion wait time is:

$$W_q = W - \overline{X} = \frac{1}{\gamma} \sum_{k=1}^{K} \sum_{n=1}^{N} kp_{k,n} - \frac{(N-1)\mu + r}{\mu r} \tag{17}$$

In the k subclass of class 1, the fuzzy mean scheduling method is used to perform adaptive scheduling, and the output of KNN clustering is obtained as follows:

$$U_{util} = \gamma \overline{X} \tag{18}$$

According to the above analysis, according to the evolution of particle swarm optimization and the idea of KNN optimization, the cluster processing of fatigue life characteristic data of ballastless track casting is realized [10].

3.2 Fatigue Life Prediction Output

In the process of fatigue life cycle of ballastless track, the fatigue life big data is clustered by KNN and the convergence is judged [11]. By using the correlation spectrum analysis method, the directivity characteristics of fatigue life categories are obtained as follows:

$$\rho_{XY} = \frac{Cov(\boldsymbol{X}, \boldsymbol{Y})}{\sqrt{D(\boldsymbol{X})}\sqrt{D(\boldsymbol{Y})}} \tag{19}$$

Where, $Cov(X, Y)$ represents the autocorrelation function of the sampled fatigue life data of two groups of ballastless track casting, and $D(X)$ and $D(Y)$ denote the average energy respectively.

The method for decomposing the fatigue life of the ballastless track is carried out by adopting a wavelet scale decomposition method, and the output fatigue life evolution characteristic amount is as follows:

$$s(t) = \underbrace{\sum_{k=1}^{N} p_k \sin(\omega_k n + \Phi_k)}_{u(n)} + \zeta(n) \tag{20}$$

Where, $\zeta(n)$ is the number of fatigue life categories, Φ_k is the phase information of fatigue life distribution, and ω_k is the recursive characteristic of fatigue life data of ballastless track casting. Suppose the balanced scheduling model of ballastless track casting data under the digital machining mode is represented as:

$$x_n = x(t_0 + n\Delta t) = h[z(t_0 + n\Delta t)] + \omega_n \tag{21}$$

In the equation, $h(.)$ is the sample time window function of fatigue life and ω_n is the measurement error. Particle swarm evolution's sample training set is $X = \{x_1, x_2, \ldots, x_n\}$, n is the number of ballastless track casting fatigue life data sets X, The abnormal state data of ballastless track during fatigue life cycle are decomposed by Hilbert spectrum extraction method and the state parameters are extracted. The results of feature extraction with multi-parameter fusion are as follows:

$$y(t) = \frac{1}{\pi} K \int \frac{x(\tau)}{t - \tau} d\tau = x(t) * \frac{1}{\pi t} \tag{22}$$

Where, K is the characteristic matching coefficient of fatigue life, $x(\tau)$ is the discriminant statistic of fatigue life prediction in ballastless track casting, $*$ is convolution, and the judgement value of output is:

$$C_{T'}(f) = \sum_{k=-K}^{K} c_k e^{-j2\pi f k T'} \tag{23}$$

The fatigue life distribution of ballastless track is divided into several IMF components by constructing expert system, and the fatigue life distribution of ballastless track is divided into several IMF components. The fatigue life prediction output of ballastless track casting with $X_k = [x_{k1}, x_{k2}, \ldots, x_{km}, \ldots, x_{kM}]$, corresponding to any training sample is as follows:

$$\boldsymbol{Y}_k = \left[y_{k1}, y_{k2}, \cdots, y_{kj}, \cdots, y_{kJ}\right] \quad (k = 1, 2, \cdots, N) \tag{24}$$

Based on this data, an intelligent expert system is established, and the fatigue life prediction of ballastless track casting is realized with big data analysis method [12].

In summary, the implementation process of the improved model is shown in Fig. 1.

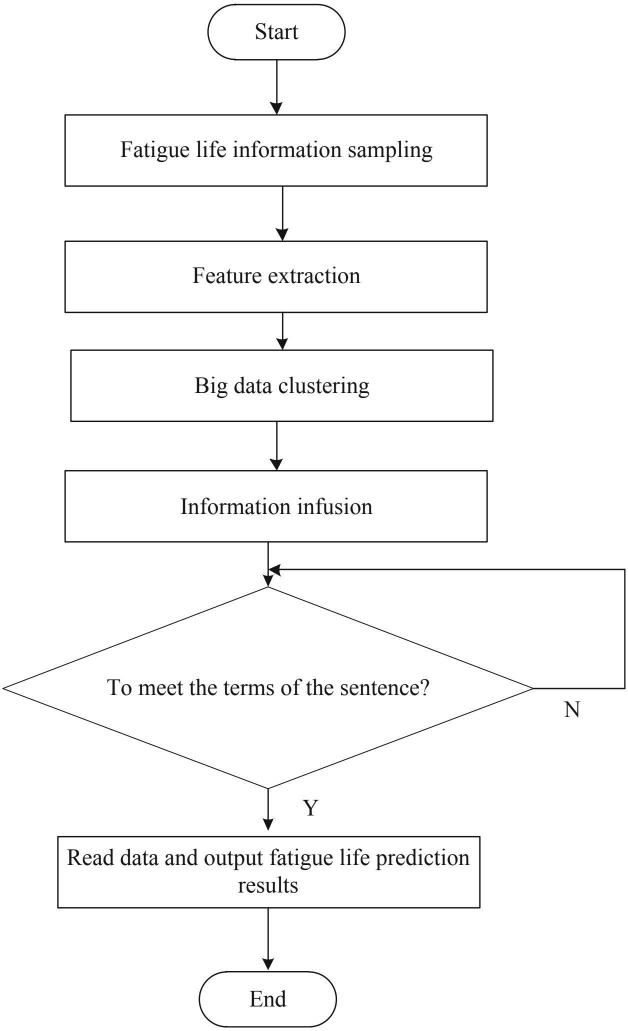

Fig. 1. Implementation flow of algorithm

4 Simulation Experiment and Result Analysis

In order to verify the application performance of this method in the prediction of fatigue life of ballastless track casting, the simulation experiment is carried out. In the experiment, the workpiece of ballastless track casting is selected as a ballastless track hand, and the algorithm design is designed with Matlab 7. The collected fatigue life of ballastless track casting big data is the oscillation amplitude of ballastless track vibration, the time interval of data sampling is 20 s, the length of data is 10244. The number of iterations of particle swarm evolution is 2000, the mutation operator is 0.24, The crossover operator is 0.15. According to the above simulation environment and parameter setting, the fatigue life prediction simulation experiment of ballastless track casting is carried out. First, the original sampling data is given as shown in Fig. 2.

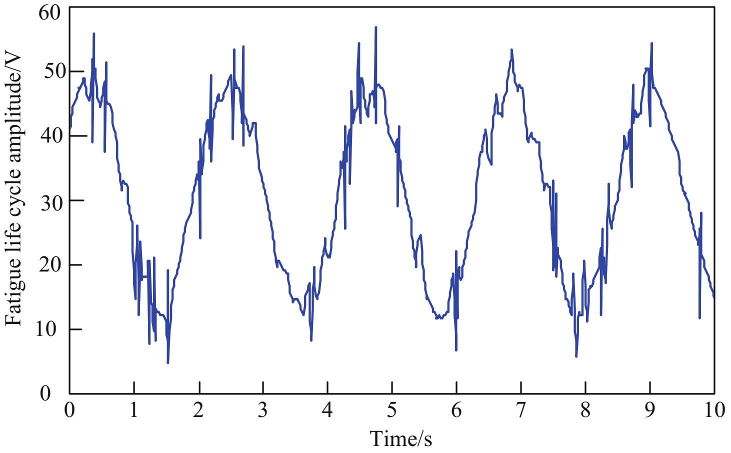

Fig. 2. Oscillation data sampling for ballastless track casting

Taking the data of Fig. 2 as the research sample, a set of relatively stable fatigue life signal analysis models of ballastless track casting are established by means of spectral characteristic analysis method, and the fatigue life data of ballastless track casting is processed by this method. The influence of interference component is effectively suppressed and the prediction ability of fatigue life is improved. On this basis, the cluster analysis of fatigue life sample data is realized, and the clustering output results are shown in Fig. 3.

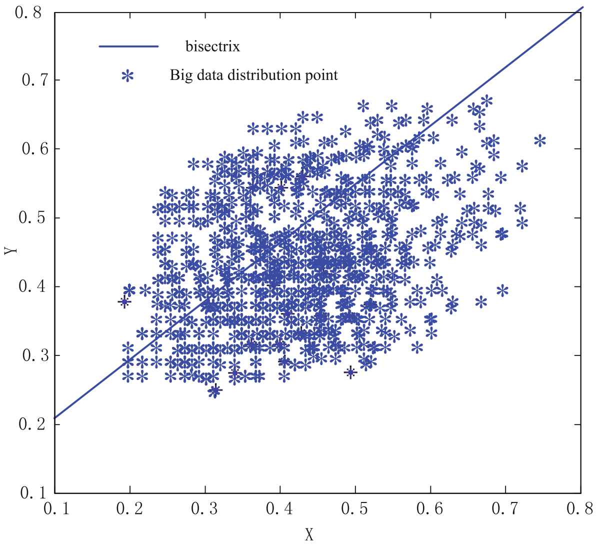

Fig. 3. Analysis of fatigue life of ballastless track casting by big data

According to the result of cluster analysis in Fig. 3, the clustering characteristic quantity is inputted into the expert system model to realize the fatigue life prediction of ballastless track casting. In order to compare the performance, different methods are used to test the accuracy of fatigue life prediction. The comparison results are shown in Fig. 4.

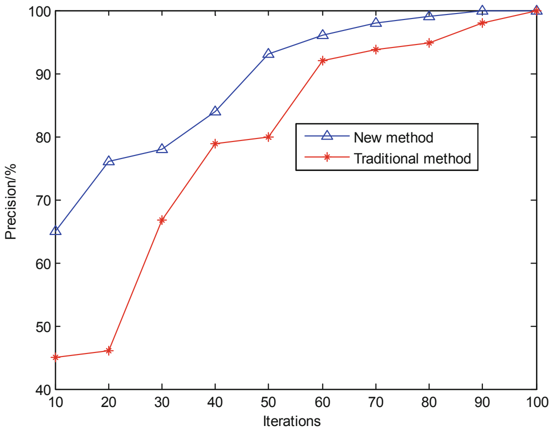

Fig. 4. Comparison of prediction accuracy

The analysis of Fig. 4 shows that the big data clustering method proposed in this paper is more accurate in predicting fatigue life of ballastless track casting, and the feature extraction and clustering of fatigue life state is better than that of non-ballastless track casting.

5 Conclusions

In this paper, the fatigue life prediction model of ballastless track casting is studied to improve the intelligent diagnostic ability of fatigue life of ballastless track casting. The abnormal state data of ballastless track during fatigue life cycle are decomposed and extracted by Hilbert spectrum extraction method. The fatigue life of ballastless track is predicted according to the result of feature extraction, and the fatigue life data of ballastless track casting is processed. The prediction of fatigue life of ballastless track is more accurate and clustering is better, so the prediction ability of fatigue life is improved, and the fatigue life prediction ability of ballastless track casting is higher than that of non-ballastless track casting. It has good application value in fatigue life intelligent diagnosis of ballastless track casting.

References

1. Hu, S., Ding, Z., Ni, Q.: Beamforming optimisation in energy harvesting cooperative full-duplex networks with self-energy recycling protocol. IET Commun. **10**(7), 848–853 (2016)
2. Seo, D.W., Lee, J.H., Lee, H.S.: Optimal coupling to achieve maximum output power in a WPT system. IEEE Trans. Power Electron. **31**(6), 3994–3998 (2016)
3. Dai, H., Huang, Y., Li, C., et al.: Energy-efficient resource allocation for device-to-device communication with WPT. IET Commun. **11**(3), 326–334 (2017)
4. Helmy, A., Hedayat, A., Al-Dhahir, N.: Robust weighted sum-rate maximization for the multi-stream MIMO interference channel with sparse equalization. IEEE Trans. Commun. **60**(10), 3645–3659 (2015)
5. Alfaro, V.M., Vilanovab, R.: Robust tuning of 2DoF five-parameter PID controllers for inverse response controlled processes. J. Process Control **23**(4), 453–462 (2013)
6. Han, D., Chen, X., Lei, Y., et al.: Real-time data analysis system based on Spark Streaming and its application. J. Comput. Appl. **37**(5), 1263–1269 (2017)
7. Sun, D.W., Zhang, G.Y., Zheng, W.M.: Big data stream computing, technologies and instances. J. Softw. **25**(4), 839–862 (2014)
8. Hao, S.G., Zhang, L., Muhammad, G.: A union authentication protocol of cross-domain based on bilinear pairing. J. Softw. **8**(5), 1094–1100 (2013)
9. Zikui, M.A., Chen, W.: Friction torque calculation method of ball bearings based on rolling creepage theory. J. Mech. Eng. **53**(22), 219–224 (2017)
10. Zhou, S.B., Xu, W.X.: A novel clustering algorithm based on relative density and decision graph. Control Decis. **33**(11), 1921–1930 (2018)
11. He, H., Tan, Y.: Automatic pattern recognition of ECG signals using entropy-based adaptive dimensionality reduction and clustering. Appl. Soft Comput. **55**, 238–252 (2017)
12. Zhu, Y., Zhu, X., Wang, J.: Time series motif discovery algorithm based on subsequence full join and maximum clique. J. Comput. Appl. **39**(2), 414–420 (2019)