Research on Load Feature Extraction Method of Typical Users Based on Deep Learning

Zhu Lian-huan[1](✉), Wei Wei[1], Zhu Wei-yang[2], Ding Can-song[2], Shen Kai[2], and Fu Guan-hua[1]

[1] State Grid Hangzhou Xiaoshan Power Supply Company, Hangzhou 311200, China
[2] Zhejiang Zhongxin Power Engineering Construction Co., Ltd., Hangzhou 311200, China

Abstract. In order to improve the accuracy of typical user load feature extraction, this paper proposes a typical user load feature extraction method based on deep learning. Using k-means algorithm to cluster user load data, select typical user load sample data from the clustering results, and classify user load categories, and implement the extraction of typical user load features based on the classification results combined with deep learning methods. The experimental test results show that the method in this paper has high accuracy in the extraction of typical user load characteristics, high accuracy in load recognition, and good practical application effects.

Keywords: Deep learning · Typical user · Load feature · Feature extraction

1 Introduction

In order to improve the calculation efficiency and extraction accuracy of typical user load extraction, the research on typical user load extraction method is carried out in order to realize the typical user load extraction quickly and accurately. As one of the key technologies of load monitoring for typical users, the function of load extraction is to extract and analyze various power consumption characteristics of residents, businesses and administration by using the load characteristics extracted from load monitoring data [1]. At present, there are many researches on the implementation methods of load extraction of typical users, which mainly use two kinds of methods: mathematical optimization algorithm and pattern extraction algorithm. When using mathematical optimization algorithm to extract load, most of them use differential evolution algorithm and genetic optimization algorithm. Because the mathematical model of typical user load extraction is more complex, the mathematical optimization algorithm is faced with many problems the low efficiency of solving the problem affects the applicability of mathematical optimization algorithm in the typical user load feature extraction method [2]. The commonly used pattern extraction algorithm is supervised learning algorithm. Its working principle is to select the load historical monitoring data in a certain period of time as the training data set, and use the data set to train the algorithm. Finally, the real-time monitoring data of load is extracted according to the training results.

W. Fu et al. (Eds.): ICMTEL 2021, LNICST 387, pp. 145–156, 2021.
https://doi.org/10.1007/978-3-030-82562-1_13

However, these two methods have the problem of low accuracy of typical user load feature extraction, so this paper proposes a typical user load feature extraction method based on deep learning. This paper uses k-means algorithm to cluster user load data, selects typical user load sample data in the clustering results, and divides user load categories. According to the classification results, combined with deep learning method, the typical user load feature extraction is realized, so as to solve the problems existing in traditional methods and improve the accuracy of typical user load feature extraction. I hope this paper can do some research It can lay a solid foundation for the further development of power data analysis.

2 User Load Feature Extraction Method

2.1 Typical User Load Data Clustering

Before clustering with k-means algorithm, data preprocessing is needed: suppose that there are m sample data in the initial data set for deep learning algorithm, and there are k characteristic data attributes in the initial data set, and the $k > 0$ is an integer. In the whole process of data clustering, m sample data are formed into the sum of squares of cluster center distance with characteristic attributes

$$\Delta F = U(j)\sum_{i=1}^{K}\bar{\omega}^{*}F_i = \sum_{i=1}^{K}\bar{\omega}^{*}\sum_{j=1}^{m}\alpha_{ij}\left\|A_j - B_i\right\|^2 \tag{1}$$

In the formula: F represents the sum of the squared distance between the cluster centers and the objective function; A_j represents the j data feature attributes, B_i represents the i feature cluster centers, and α_{ij} represents the weight coefficient. In the clustering process, the square of the distance between the cluster centers and the extreme value of the objective function need to be taken. When the k-means algorithm is applied to the extraction of complex power consumption characteristics of users, the original text data must be determined first, and then the cluster centers must be randomly selected. Use the k-means algorithm to find the F value. In the iterative process, as long as the F value keeps changing, it means that the clustering has not reached the optimum. At this time, the cluster center needs to be updated and the iterative process is repeated. The calculation formula for cluster centers is as follows:

$$O_{ij} = \Delta F/\Delta i\left\|A_j - B_i^{i-1}\right\| \tag{2}$$

The above formula mainly calculates the distance between the sample data and the feature clustering center. According to the above formula, the implementation steps are as follows:

1) Take k initial condensation points as the initial centroid;
2) Calculate the distance from the sample data to the feature cluster center;
3) According to the k-means clustering principle, all points are assigned to k feature cluster centers;
4) Calculate the centroid of k feature cluster centers;

5) Repeat steps 2), 3), and 4) until the center of mass no longer needs to be updated;
6) Realize feature recognition.

The power consumption characteristic data is obtained by sampling and processed, and sent to the data processing center to measure the power consumption of the users in combination with the total electricity consumption. With the support of deep learning algorithm, complex power consumption feature extraction is carried out to fill in the missing value of user power consumption, smooth noise data, delete off cluster points, and avoid unreliable data output. For the problem of missing data, the data mean and median can be used for interpolation, and the attribute value closest to the missing sample can be found in the record for interpolation. If the missing value is filled with constant, the data attribute should be analyzed in advance, and the data cleaning should be carried out based on this. In data clustering, we should first improve the data quality and reduce redundant data. Use metadata methods to avoid possible extraction problems in data sets. For data redundancy problems, it is necessary to use correlation analysis to detect whether the data is redundant, and store it in different measurement units, fully consider random variables, and map the original data In a small space, after data compression, it is judged whether the original data can be reconstructed, so as to regulate a large amount of data. Eliminate the noise data in the data, transform continuous attributes into categorical attributes, and divide the data into discrete intervals to analyze the data by clustering.

2.2 User Load Classification

With the development of power grid technology, the situation of users' power consumption becomes more complex, and a large number of power consumption characteristics are generated. Previous load feature extraction methods are affected by noise data, resulting in low extraction accuracy. To solve this problem, a typical user load feature extraction method based on deep learning is proposed [3]. In the deep learning algorithm, the principle of typical user load feature extraction is studied, and the data is cleaned, integrated and pre processed to avoid noise interference. Through the cluster decision-making power consumption feature points, the information gain of power consumption characteristics is obtained, and the complex power consumption characteristics are extracted to provide better and high-quality services for users. Load extraction and decomposition is to extract the components of the total load according to the characteristics of the load work imprint extracted from the measured data, and realize the load decomposition on this basis, including mathematical optimization and mode extraction. Load extraction problem can be described as the following mathematical problems:

$$\Phi = [A_l, A_2, \cdots, A_M] \tag{3}$$

$$A_i = \left[f_{i,1}, f_{i,2}, \cdots, f_{i,n}\right]^T \tag{4}$$

$$B = [F_1, F_2, \ldots, F_n]^T \tag{5}$$

$$X = [x_1, x_2, \cdots, x_M]^T \tag{6}$$

Assuming that the above formula is true, the solution of the following formula is satisfied:

$$\min d(C, B) = d\left(\sum_{i=1}^{M}(x_i A_i), B\right) \tag{7}$$

Among them, Φ is a feature matrix composed of M type electrical equipment, A_i is a feature matrix composed of all the load characteristics of i type electrical equipment, and f_{in} represents the n load feature of i type electrical equipment. B is a feature matrix composed of all the features in the entire electricity use scene, F_n represents the n load feature quantity in the electricity use scene, X is a state matrix composed of the working states of all M types of electrical equipment, and x_i takes 1 to indicate this type of electrical equipment is in the working state. x_i value of 0 means that this type of electrical equipment is off. $d(C, B)$ represents the distance between the possible combination of all load characteristics and the total load. Because the short-term load data of the power dispatching switch is affected by many factors, the data selection is not accurate. In order to solve this problem, it is necessary to select the load data that has a greater impact on it. Time and temperature have the greatest impact on the power load data [4]. In the process of intelligent short-term load forecasting combined with deep learning algorithm, the time and temperature data need to be integrated, and the power load data of a certain period of time can be predicted through the network, and the training sample selection process is shown in the figure (Fig. 1).

It can be seen from the figure that: the initial weight, input each learning sample of current N into the network in turn, calculate the input and output values of each layer, calculate the back propagation error of each layer, and record the number of trained samples M in time. If M is less than N, the learning samples need to be re input into the network to calculate the input and output of each layer; if M is equal to N, the weights of each layer should be corrected according to the weights, and the connection weights of hidden layers should be adjusted. According to the new weight, the error of each layer is recalculated. If the error is less than the set expected value, the training reaches the maximum number of times, and the sample selection is terminated. Otherwise, it returns to the sample input step.

Considering that among independent users, the types of home appliances are limited, the brand models of the same load are relatively fixed, the corresponding waveforms are fixed, and the operating environment and operating habits are relatively stable. Therefore, this article regards each independent user as the basic unit of constructing the feature library, and unlimited categories the classification problem is transformed into a limited category classification problem [5]. In order to ensure that the feature library has universal applicability to most independent users, the classification judgment conditions are set according to the operational characteristics of common loads, so as to transform the unsupervised classification problem into a supervised classification problem of load types. The Bayesian classification model is used to classify the load, the decomposed independent operating load characteristics are used as the posterior knowledge, and the posterior probability is converted into the prior probability of the load category to solve, as shown in the formula.

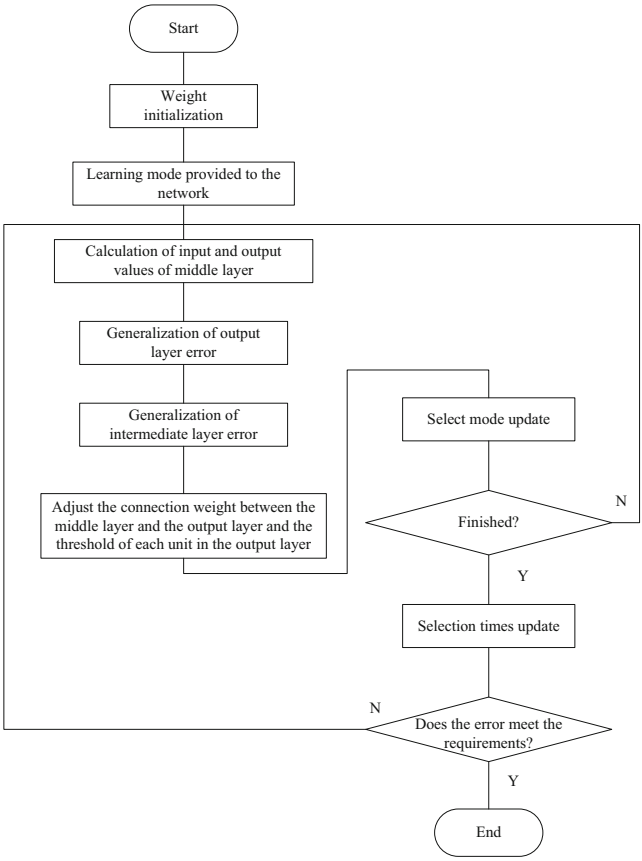

Fig. 1. Feature sample selection process

$$P(\omega_n|F_k) = \frac{P(F_k|\omega_n)P(\omega_n)}{P(F_i)} \quad n = 1, 2, \cdots, N \tag{8}$$

Where: ω_n is the classification of the k independent load signal U_k and I_k; F_k is the extracted unknown load characteristics; N is the total number of electrical appliances in the user; $P(\omega_n)$ is the probability of the occurrence of load category ω_n, $P(F_k|\omega_n)$ is the probability of feature F_k under the condition of known load type; $P(F_k)$ is the probability of switching on and off of load K in the table replaced by F_k after clustering. The prior probability $P(\omega_n)$ is converted to the posterior probability $P(\omega_n|F_k)$ by the obtained stability characteristic F_k, that is, the probability that the load k category belongs to ω_n when F_k is known, and the category with the largest probability is the label of load k, as shown in the formula:

$$L_t = L_k \arg\max P(\omega_n|F_k) \tag{9}$$

Where: L_k is the label of load k. In this way, all the separated independent loads are classified and processed, the unknown loads obtained by successive separations

can be labeled and their waveforms, characteristics, and categories are recorded in the feature library to complete the construction of the independent user load feature library. After forming the user's load characteristic database, the data in the database is used to continuously identify the user's independent load waveform after load separation in real time [6]. Since the high-frequency acquisition mode can better retain the integrity of the load waveform, the load identification quick optimization model is further established based on the waveform similarity. The steady-state current of the electrical equipment in normal operation has linear superposition, that is, the collected mixed signal can be estimated by the linear superposition of the current of the N-type electrical equipment. The current mixed current signal can be expressed as:

$$I(t) = L_t \sum_{k=1}^{N} a_i I_k(t) + n(t) \quad k = 1, 2, \cdots, N \tag{10}$$

Where: a_i is the start stop coefficient of load i, $a_i = 0$ is the load I is not opened at time t, and $a_i = w$ is that there are w such loads at the same time. The problem of load identification is transformed into a set of optimal weight coefficients a_1, a_2, ..., a_n, which makes the load superimposed current most similar to the real current, so as to determine the mixed load type in the collected current at this time and realize load extraction. The optimal weight coefficient at the current moment is determined by the minimum residual method:

$$\hat{I}_v = I(t) \sum_{i=1}^{N} a_q \hat{I}_q \tag{11}$$

$$\min d = \left\| I_k - \hat{I}_\tau \right\| \tag{12}$$

In the formula: when $a_q = 0$, it means that the corresponding electrical appliance is not turned on, when aq $a_q = 1$, it means that the corresponding electrical appliance is on, and $\hat{I}_q$ is the current of the electrical appliance q stored in the library. The key to load identification is to find the optimal weight coefficient to make the superimposed signal most similar to the real signal at the current moment [7]. The switching operation of an independent load is defined as an electrical event. When the above events occur in the circuit, the most obvious change in the collected data is the current intensity. Therefore, this paper samples the current and voltage according to a certain frequency to detect the event, and records the voltage and current waveforms in the last 2S in the buffer. The current intensity calculation of each cycle current signal can be expressed as follows:

$$i_{\text{vars}} = \min d \sqrt{\sum_{j=1}^{\tau} \frac{i^2(j)}{\tau}} \tag{13}$$

$$\Delta i = i_{\text{RMS}}(T+1) - i_{\text{vars}}(\eta\tau) \tag{14}$$

In the formula: τ is the number of sampling points in a cycle, $i^2(j)$ is the sampling value of the cycle current; T is the T-th cycle of the current. If the current intensity of one

cycle has a sudden change compared with the previous cycle, that is, $\Delta i > \eta$ and η are the thresholds for judging the sudden change of the steady-state current, and it can be considered that a load switching event has occurred. The fundamental phase angle of the steady-state current is determined by the initial phase of the voltage during measurement. Therefore, it is only necessary to ensure that the steady-state current is measured under the same initial phase angle voltage to meet the current superposition. The load current waveform reconstructs the mixed current waveform. Detect the voltage zero-crossing point that shows an upward trend in the corresponding steady-state terminal voltage waveform, and extract the steady-state current waveform I before the electric load k is switched on.

$$\begin{cases} U(j) > 0 \\ U(j-1) < 0 \end{cases} \tag{15}$$

Where: J is the sampling point of current stable period corresponding to the T-1 cycle when the formula is satisfied; U(J) is the steady-state voltage.

2.3 Implementation of Characteristic Extraction of Typical User Load

Deep learning is a method of machine learning based on representation of data. It belongs to the category of machine learning, can be said to be an upgrade on the basis of traditional neural network, about equal to neural network. Its advantage is to use unsupervised or semi supervised feature learning and hierarchical feature extraction algorithm to replace manual feature extraction. Because of its strong learning ability, wide coverage, good adaptability, good portability and other advantages, it is reliable to apply it to typical users' load feature extraction.

Due to the complexity of the data extracted from the user's power load feature, the extraction quality is low due to the interference of external noise in the process of acquisition and transmission. Therefore, it is necessary to preprocess it first, and label the unknown waveform stored in the feature library in this process [8]. Generally, typical user load features are switched on and off for many times, and the extracted load waveform is repetitive. It is not necessary to judge the types of all the extracted waveforms and store them in the feature library in the feature library establishment phase. Therefore, the waveforms extracted by deep learning method are clustered first [9]. In the initial operation stage, the number of load types, load operation modes and the number of modes are unknown. In this paper, the inherent characteristics of load steady-state operation are used for fast clustering. Therefore, the results of load characteristics based on deep learning can be expressed as follows:

$$F_t = \left\{P_k, Q_k, P_{Ft}, R_{\mathrm{THD},k}, r_k\right\} \tag{16}$$

Where: P_k and Q_k are active and reactive power respectively; P_{Fk} is power factor; R_{THD} and k are current distortion rate; r_k is Pearson coefficient between the current waveform and the constructed standard sine wave, and the peak value and frequency of the standard sine wave are the same as the current waveform. Since the numerical range of different features is different, normalize each feature according to the above formula to get.

$$\bar{\omega}^* = \frac{(\bar{\omega} - \bar{\omega}_{\min})F_t}{O_{ij}(\bar{\omega}_{\max} - \bar{\omega}_{\min})} \tag{17}$$

Where: $\bar{\omega}_{\max}$ and $\bar{\omega}_{\min}$ are the maximum and minimum values of features respectively; $\bar{\omega}$ is the value before and after feature normalization. Single step judgment can not guarantee the accuracy. Therefore, this paper uses the inherent characteristics of the load to pre classify the unknown waveform, narrow the load label range and reduce the burden of subsequent labeling. PFK, RTHD, K and rk are used to divide the load into linear and non-linear. According to the times of detecting such waveform in a short time, it is divided into continuous or intermittent operation. Finally, according to the duration of each switching the change degree is judged as fixed or non fixed operation time load, as shown in the formula:

$$T_{\mathrm{D}} = \bar{\omega}^* \sum_{p=1}^{p} \frac{T_d(p)}{P} \tag{18}$$

$$T'_{\mathrm{D}} \in [T_d(1-a), T_d(1+\alpha)] \tag{19}$$

Where: T_d is the average duration of the previous p times of the same load, $T_d(p)$ is the duration of the p-th load; T'_{D} is the duration of the detected load belonging to this type of load; α is the floating factor, if the above formula is satisfied, then This load is a fixed operating time category. Obtain current and voltage sampling data through data acquisition methods such as smart meters, and calculate characteristic values such as overall power and harmonics [10]. The difference feature extraction method is used to obtain the change feature quantity produced by the change of the electrical appliance state at any time, as the data input of the fuzzy clustering; the inter-cluster entropy value of the clustering result is calculated and iterated, and the optimal number of clusters is selected. Finally, calculate the similarity between each type of data and the characteristics of electrical appliances, and select the most similar electrical features as the type of electrical appliances represented by the data. Based on this, the user load feature extraction process is standardized, as shown in the following figure (Fig. 2):

Through the above process, the number and type of electrical appliances can be extracted under the condition that the situation of household appliances is unknown, which can solve the data problems for power consumption behavior analysis and smart home applications. Improve the accuracy of user load feature extraction.

3 Analysis of Results

In order to verify the practical application effect of typical user load feature extraction method based on deep learning, comparative experiments are carried out, and the mathematical optimization algorithm mentioned in the introduction is taken as the traditional method. The subjects of the experiment include residents, commerce, administration, transportation, industry and so on. The sampling current and voltage data are obtained by USB A/D data acquisition card, and the acquisition frequency is 1 kHz. The parameters of the experimental equipment are standardized, as shown in the table below (Table 1).

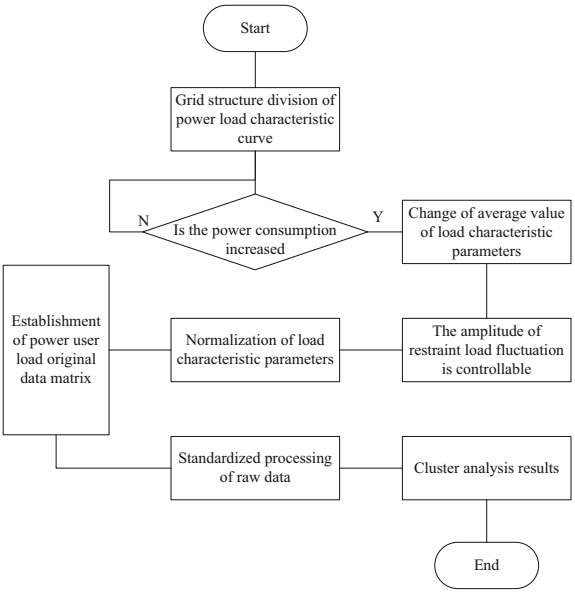

Fig. 2. Flow chart of user load feature extraction

Table 1. Experimental parameter settings.

	Database server	Application server	Client
Number of CPU	5	5	2
Type of CPU	-	-	P43.2G
Memory	16G	16G	512M
Hard disk	2 * 120G	2 * 320G	-
Network card	4 * 1000M	4 * 1000M	100M

During the experiment, one of the equipment was started and stopped randomly for 100 times, and the monitoring time was 10 min. The difference feature extraction method proposed in this paper is used to extract harmonic features as shown in the figure (Fig. 3).

It can be seen that the distribution of active and reactive power features of some electrical equipment is relatively concentrated, which is not conducive to the classification and extraction of clustering algorithms. After harmonic features are used to replace reactive power, the load features are clearly divided into 5 regions, which can be more effectively perform fuzzy clustering extraction. When the active and harmonic features are used, the entropy value between clusters is higher than that of the active and reactive features, and it is optimal when the number of clusters is 5, which is consistent with the actual situation. This shows that the difference feature extraction and fuzzy clustering methods used in this paper can effectively extract load features, and can perform

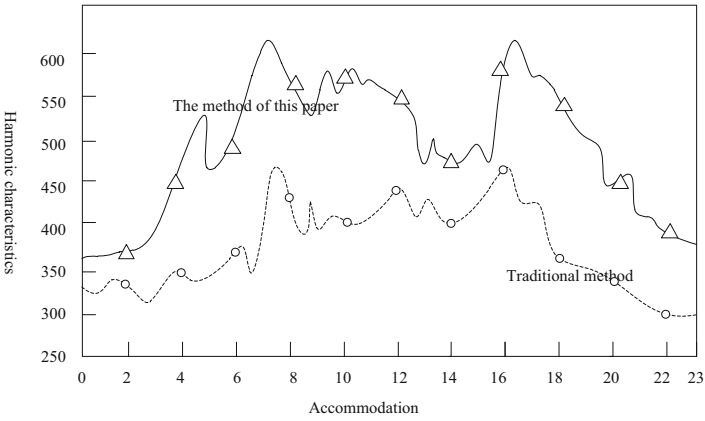

Fig. 3. Harmonic feature extraction results of electrical load

fuzzy clustering based on the inter-cluster entropy when the device type is unknown, and improve the accuracy of load extraction.

The load identification accuracy of typical users such as residents, commerce, administration, transportation and industry is shown in the following table:

Table 2. Correct rate of load identification.

Electrical type	Recognition accuracy/%	
	Traditional method	The method of this paper
Resident	80.0	100.0
Business	73.7	100.0
Administration	80.0	100.0
Traffic	65.4	94.1
Industry	82.0	98.0
Resident	76.2	98.4

It can be seen from Table 2 that the correct rate of load identification of this method is higher. In order to further verify the performance of the method in this paper, high-power home appliances are used to detect load characteristics. Some equipment starts and stops randomly for a total of 200 times, and the monitoring time is 1000 s. The load characteristic data obtained through sampling is normalized. The distribution of power characteristics of a variety of electrical equipment is relatively concentrated, especially the load with lower power, which is not easy to monitor and extract the load. Due to the large harmonic current characteristic value of low-power load, the load characteristic distribution is more scattered, which is more conducive to further fuzzy clustering extraction. Based on this analysis of the test results, the details are as follows:

Table 3. Load identification accuracy of high power household appliances.

Electrical type	Recognition accuracy/%	
	Traditional method	The method of this paper
Resident	75.4	98.4
Business	82.1	100.0
Administration	79.6	100.0
Traffic	80.6	99.2
Industry	81.3	96.5
Resident	85.4	100.0
Resident	77.2	96.4
Business	70.6	100.0
Administration	78.4	100.0
Traffic	79.9	98.9

According to the analysis of Table 3, the extraction accuracy of typical user load feature extraction method based on deep learning has been significantly improved compared with the traditional method.

4 Concluding Remarks

This paper mainly proposes a typical user load feature extraction method based on deep learning. The k-means algorithm is used to cluster the user load data, the typical user load sample data is selected from the clustering results, and the user load categories are divided, and the typical user load characteristics are extracted according to the classification results combined with the deep learning method. Finally, simulation experiments show that the method in this paper can achieve effective and accurate load extraction.

References

1. Angkoon, P., Rami, N.K., Erik, S.: Feature extraction and selection for myoelectric control based on wearable EMG sensors. Sensors **18**(5), 1615–1618 (2018)
2. Aliakbary, S., Habibi, J., Movaghar, A.: Feature extraction from degree distribution for comparison and analysis of complex networks. Comput. J. **58**(9), 2079–2091 (2018)
3. Christ, M., Braun, N., Neuffer, J., Kempa-Liehr, A.W.: Time series FeatuRe extraction on basis of scalable hypothesis tests (tsfresh – a Python package). Neurocomputing **307**, 72–77 (2018). https://doi.org/10.1016/j.neucom.2018.03.067
4. Lijian, Z., Chen, Z., Zuowei, W., et al.: Hierarchical palmprint feature extraction and recognition based on multi-wavelets and complex network. IET Image Proc. **12**(6), 985–992 (2018)
5. Taguchi, Y.-H.: Tensor decomposition-based and principal-component-analysis-based unsupervised feature extraction applied to the gene expression and methylation profiles in the brains of social insects with multiple castes. BMC Bioinform. **19**(S4), 99 (2018)

6. Woo, J.-H., Kim, C.-W., Bang, T.-K., Lee, S.-H., Lee, K.-S., Choi, J.-Y.: Experimental verification and electromagnetic characteristic analysis of permanent magnet linear oscillating actuator using semi 3D analysis technique with corrected stacking factor. IEEE Trans. Appl. Supercond. **30**(4), 1–5 (2020). https://doi.org/10.1109/TASC.2020.2986737
7. Fu, W., Liu, S., Srivastava, G.: Optimization of big data scheduling in social networks. Entropy **21**(9), 902 (2019)
8. Liu, S., Li, Z., Zhang, Y., et al.: Introduction of key problems in long-distance learning and training. Mob. Netw. Appl. **24**(1), 1–4 (2019)
9. Liu, S., Lu, M., Li, H., et al.: Prediction of gene expression patterns with generalized linear regression model. Front. Genet. **10**, 120 (2019)
10. Takeuchi, K., Matsushita, M., Makino, H., Tsuboi, Y., Amemiya, N.: A novel modeling method for no-load saturation characteristics of synchronous machines using finite element analysis. IEEE Trans. Magn. **57**(2), 1–5 (2021). https://doi.org/10.1109/TMAG.2020.3010859