Low voltage abnormal user identification based on improved fish swarm algorithm

Bing Kang\textsuperscript{1,2*}, Chuan Liu\textsuperscript{2}, Min Sun\textsuperscript{1}, Tianqi Meng\textsuperscript{1}, Jun Zhou\textsuperscript{2}, Zongyao Wang\textsuperscript{2}, Zhihao Xu\textsuperscript{2}

\textsuperscript{1}State Grid Jiangxi Electric Power Research Institute, Nanchang, 330096, China
\textsuperscript{2}Nanchang Institute of Technology, Nanchang, 330099, China
kangbing@nit.edu.cn

Abstract. The power consumption readings of sub meter and total meter of distribution transformer of low-voltage users follow the law of conservation of energy. The meter power loss rate of abnormal low-voltage users must also be abnormal. This paper studies the solution of the meter power loss rate under the four abnormal power consumption scenarios of single (multi) user and full (partial) period. The traditional linear solution method has accurate identification effect for the abnormal power consumption scenario of full period, but it cannot identify the abnormal power consumption scenario of partial period. In this paper, an improved artificial fish swarm algorithm is proposed. By adjusting the fixed step to the adaptive step, the power loss rate of each sub meter is obtained, and the abnormal power users are pinpointed. The research results are verified by simulation examples on IEEE European Low Voltage Test Feeder. The results show that the improved artificial fish swarm algorithm in this paper can identify abnormal power users for the above four abnormal electric field scenarios. The algorithm provides a new alternative for the identification of abnormal low voltage users.

1. Introduction
With the rapid development of society, the demand for electric energy increases rapidly. Electric energy is not only an important support for the development of the national economy, but also an important guarantee for electric power companies to maintain their own development. The abnormal power consumption behavior of low-voltage users not only damages the interests of electric power companies, but also seriously affects their healthy development. Meanwhile, it also brings great hidden dangers to the security of the power grid and power consumption.

Nowadays, there are various low-voltage users' abnormal power consumption behaviors, and the detection of abnormal power consumption behavior cannot be achieved solely by relying on the abnormal monitoring function of the electric power meter itself. In recent years, with the development of big data application innovation strategy, a comprehensive construction of power user electricity information collection system is created, and the system collects massive power measurement information and electricity state information. The huge electricity information data contains users’ electricity consumption state through the existing massive data research. By establishing abnormal electricity behavior analysis model, realizing suspicious electricity user screening, improving electricity information collection system prevention, investigating abnormal timeliness and pertinence, the system gives full play to the role of electricity information collection in abnormal electricity behavior identification.
The traditional identification methods of abnormal electricity behavior are mainly as follows: 1. The power theft inspectors check power on all kinds of customers from time to time; 2. Consulting the electricity marketing system investigation and select some key users for on-site inspection; 3. Installing the new anti-theft intelligent electricity meter or the anti-theft electricity meter box. The above abnormal identification methods of electricity behavior have high requirements on the experience of inspectors and require a lot of manpower and material resources. Therefore, using massive basic data to model and analyze user electricity behavior, and using the algorithm model to judge whether users have abnormal electricity behavior through big data, artificial intelligence and other technologies has received more and more attention. At present, using clustering algorithm to classify the electricity load curve of users is a rather popular method. Paper [1] uses a fuzzy C mean (fuzzy C-means, FCM) algorithm to cluster and analyze the users’ data, which include residents, business, industries and government. Paper [2] combines linear loss correlation analysis with clustering by fast search and find of density peaks, CFSFDP, for the daily load curve. This method can verify the identification of abnormal electricity users by more than 70% on data gathered from Irish residents and small and medium-sized enterprises. Paper [3] uses the outlier characteristics formed from the voltage current data measured at the split voltage or diversion theft, and clusters and analyzes the current voltage data of low-voltage users in high-loss station to identify abnormal users using under-pressure method and undercurrent method. To sum up, the cluster-based abnormal electricity user identification method generally uses the characteristic indicators such as the trend decline of electricity consumption, abnormal daily load curve and low installation capacity utilization rate for feature optimization and cluster analysis [4-5]. Due to the significant differences in electricity behavior characteristics of users in different industries, the practicability of identifying electricity abnormalities based on daily load curve clustering needs to be tested in engineering practice. In addition to the above clustering algorithm, paper [6] uses support vector machine, SVM, as a classifier and compared with the abnormal power consumption user identification algorithm. Paper [6] suggests electricity anomaly identification method based on sparse random forest model, and measure model performance by TPR, FPR. Paper [8-9] adopt deep convolutional neural network, CNN, and stack-based correlation autoencoders to identify abnormal electricity users using load time series data, respectively, and achieved good identification effect on the dataset used for the test.

In this paper, the algorithm model of low voltage users based on improved fish algorithm is testified for identification and verification of abnormal power consumption behavior in four different abnormal power consumption scenarios: single user in full time period, single user in partial time period, multiple users in full time period, and multiple users in partial time period. Then, the method in this paper is compared with the least square method.

2. Methodology

2.1. Principle of conservation of electric energy in station area

![Fig. 1. Schematic diagram of power cluster in low voltage substation area](image-url)
Suppose there are \( n \) electricity users under a certain area, as shown in Fig.1, \( M_i \) is the outlet power meter of the distribution transformer in the station area; \( M_1, M_2, M_3, \ldots, M_n \) are the power meter reading of the user 1 to the user \( n \) respectively. And \( \delta_1, \delta_2, \delta_3, \ldots, \delta_n \) are the line power loss rate corresponding to the user respectively. Since the active electric energy of the station area is conserved, we can have the following relationship:

\[
M'_n = M'_1 + M'_2 + \cdots + M'_n
\]  

(1)

Further simplify formula (1): assuming that the meter error \( \delta_1, \delta_2, \delta_3, \ldots, \delta_n \) is fixed within a certain load range, then the readings of the power meter \( M_1, M_2, M_3, \ldots, M_n \) within \( m \) time period are placed into equation (1), and the following system of equations can be obtained:

\[
\begin{align*}
M'_1 &= M'_1 + M'_2 + \cdots + M'_n \\
M'_2 &= M'_1 + M'_2 + \cdots + M'_n \\
\vdots & \quad \vdots \\
M'_n &= M'_1 + M'_2 + \cdots + M'_n \\
\end{align*}
\]  

(2)

In formula, \( M'_1, M'_2, M'_3, \ldots, M'_n \) is the readings of the power meter \( M_1, M_2, M_3, \ldots, M_n \) during the first period, \( M'_2, M'_3, M'_4, \ldots, M'_n \) are the second period and the readings of the power meter \( M_1, M_2, M_3, \ldots, M_n \). Therefore, the power loss rate \( \delta_1, \delta_2, \delta_3, \ldots, \delta_n \) can be obtained by solving the \( n \) element equations of Eq. (2).

2.2. Adaptive step size artificial fish swarm algorithm

The text follows on from the subsubsection heading but should not be in italic. In water, fish can often follow other fish to find more nutrients or by themselves. So the place where the densest the fish is the most nutrients it has. Artificial fish algorithm is based on this characteristic and construct artificial fish to imitate fish foraging, clustering and following behavior, so as to achieve excellent performance[10].

Due to the rather large step size, the range of artificial fish movement is also wide, which is conducive to rapid convergence. However, this might lead to the skipping of the optimal point in artificial fish swarm, which is not conducive to convergence. The later part of the algorithm has shorter step size which help accurate searching. The artificial fish step length adjustment mechanism is designed to make the initial artificial fish group algorithm have a large step length, so as to achieve rapid convergence and the step length in the later stage is shorter which cab improve the search accuracy.

Set \( f_i(i) \) as the distribution function of the current artificial fish step length, \( i \) as the current iteration times and step as the initial moving step. With the right attenuation of the Poisson distribution, the artificial fish moving step is attenuated with the iteration times and \( \lambda \) is the custom constant;

\[
f_i(i) = \begin{cases} 
\text{step} & i = 1 \\
\frac{\lambda^i}{i!} \times \text{step} & 1 < i \leq i_{\text{max}}
\end{cases}
\]  

(3)

On the basis of formula (1), when a user experiences abnormal electricity consumption, \( M'_1 > M'1 (1 + \delta_1) + M'2 (1 + \delta_2) + \cdots + M'n (1 + \delta_n) \) is obtained. Correct the \( \delta_1, \delta_2, \delta_3, \ldots, \delta_n \) and set the modified power loss rate of the line as \( \delta'_1, \delta'_2, \delta'_3, \ldots, \delta'_n \). Then we have:

\[
M'_n = M'_1 (1 + \delta'_1) + M'_2 (1 + \delta'_2) + \cdots + M'_n (1 + \delta'_n)
\]  

(4)

The objective function of power balance equation for distribution grid (4)
\[ F = |M_0 - M_e| \\
= |M_{001} - M_{e01}| + |M_{002} - M_{e02}| + \ldots + |M_{096} - M_{e96}| \quad (5) \]

In formula: \( F \) is the error in the power calculated values at 96 sampling points a day and the number of measured energy values at 96 sampling points a day after correcting the line power loss rate. \( M_{001}, M_{002}, \ldots, M_{096} \) are the corrected number while \( M_{e01}, M_{e02}, \ldots, M_{e96} \) are the actual number.

The objective function is solved by the adaptive step artificial fish group algorithm to identify low-voltage abnormal users. First initialize the position of the artificial fish individuals in the adaptive step artificial fish population algorithm.

Initialize the position state of the individual artificial fish as a vector \( X = (x_1, x_2, \ldots, x_n) \). The vectors \( x_1, x_2, \ldots, x_n \) are a set of possible line power loss rates corresponding to users. The food concentration of the current position of artificial fish is \( Y = f(x) \) corresponding to the error in the target function. The higher the food concentration, the lower the error. Set the initial parameters of artificial fish number as \( N \), fish vision \( \text{visual} \), fish movement step \( \text{step} \), maximum foraging attempts \( \text{try\_number} \), crowding factor \( \sigma \), and maximum iteration times \( i_{\max} \); Randomly generate the initial fish group and each artificial fish represents a set of possible line power loss rate.

Calculate the optimal individual and count the corresponding food concentration function value into the bulletin board. Select the power data collection column of 96 points a day of \( n \) users from the target function. Calculate the food concentration function value, compare and select the optimal individual with the minimum error value which also means the highest food concentration. Then write the value on the bulletin board and adjust the artificial fish visual size length and implement foraging behavior and following behavior.

When artificial fish simulate foraging and following behavior respectively, the behavior with larger food concentration is actually implemented. After performing a foraging behavior and following behavior, each artificial fish checks its own state and food concentration, and compares it with the value recorded on the bulletin board. If it is better than the value recorded on the bulletin board, then the bulletin board is updated. Then judge whether the maximum number of iterations is reached, if the output result is reached, obtain a set of line power loss rate columns that meet the requirements. Finally, the number of power loss rate is analyzed, and the lines with high electricity loss rate is deemed as low voltage abnormal user.

**2.3. Approach process**

In this paper, abnormal users are divided into single user (multi-user) whole time and single user (multi-user) part time. Through the IEEE European Low Voltage Test Feeder standard test node, the power loss rate of each meter is calculated using law of conservation of energy.

As to the limitations of the least squares method, this paper proposes an improved artificial fish swarm algorithm to solve the linear power loss rate of part time single user (multi-user). And finally we find out the final abnormal power users according to the distribution of each user’s linear power loss rate. The research idea process of this paper is shown in Fig.2.
3. Analysis of simulation example

3.1. Four abnormal power consumption activities: single-user whole time power consumption abnormality, single-user part time power consumption abnormality, multiple-user whole time power consumption abnormality and multiple-user part time power consumption abnormality.

In this experiment, some nodes in the IEEE European Low Voltage Test Feeder model were used. There are 32 loads, one distribution transformer and one power supply, and with electric energy meter at the distribution low voltage side and each load. The topology of the station area is shown in Fig. 3.

Run the simulation and output the power meter data at 96 times in a day (collect them every 15 minutes), and get the electricity consumption data of each load and distribution outlet within 95 time periods. The electricity data of 32 users for 95 part time periods is transformed into X matrix, and the electricity data of 95 whole time periods of the total table into Y matrix, then solve for δ with the least squares method.

3.1.1. Multiple-user whole time abnormal power consumption. The full time refers to the user’s power behaviour at 96 times of a day (95 periods). The linear abnormal power consumption refers to the user who transforms the power meter to reduce the current or voltage by a certain proportion, or reduce the number of power representations by a certain proportion by other means. The actual abnormal users are mostly linear abnormal electricity users.
User electricity consumption numbered 20, 25, and 30 are multiplied by 0.5, 0.6, and 0.7 to simulate abnormal electricity consumption. Put these data into the formula (2) and get $\delta$.

3.1.2 Single-user partial time abnormal power consumption. Some periods refer to 96 moments of day (95 periods), selected as linear abnormal power behavior. Electricity data of the 30th to 60th time periods of the user numbered 31 is multiplied by 0.5 to simulate linear abnormal power consumption in partial periods. Put the data of abnormal setting power consumption into formula (2) and get $\delta$.

The result is shown in Fig. 5: the power consumption coefficient of multiple normal users is obviously abnormal, while the user line loss coefficient of the number 31 is 1.05, so the abnormal power consumption user cannot be judged from the result. According to the experiment, the least squares method cannot identify the abnormal electricity users in some time periods.

Now with the adaptive step size artificial fish swarm algorithm, we can have the data as shown in Fig. 6.
Fig. 6. Improved artificial algorithm for power loss rate in single user partial period abnormal power consumption scenario

The user line loss coefficient numbered 31 in Fig. 6 is 1.28, which is significantly higher than the other normal users. Compared with the least squares method, the improved fish group algorithm can identify individual users of some periods. At the same time, in order to further verify the applicability of the algorithm, the user numbered 23 with abnormal power consumption is added on the basis of the abnormal power consumption of part time singular user. Multiply the power consumption data in the 30th to 60th time period by the simulated linear abnormal power consumption of 0.4, and finally use the algorithm to obtain the line loss coefficient of the user.

Fig. 7. Improved artificial algorithm for power loss rate in multi-user partial time abnormal power consumption scenario

Fig. 7 shows that the line loss coefficient of abnormal power users during parts of the numbers 23 and 31 is significantly higher than other normal users. Thus, the algorithm still has a good effect on identifying the abnormal electricity behaviour of some periods of multiple users.

3.2. Approach comparison

Add the data from the meters to formula (2) and calculate the power loss coefficient $\delta^{(1)}$ with the least squares method. Experiments show that the conventional least squares method is only suitable for whole time period power consumption abnormality instead of part time period power consumption abnormality. The adaptive step size artificial fish swarm algorithm proposed in this paper is suitable in both scenarios. The accuracy of the algorithm above is shown below:
Table 1. Comparison of least squares method and adaptive step size artificial fish swarm algorithm

| Method                              | Single user whole time period power consumption abnormality | Multiple users whole time period power consumption abnormality | Single user part time period power consumption abnormality | Multiple users part time period power consumption abnormality |
|-------------------------------------|-----------------------------------------------------------|---------------------------------------------------------------|----------------------------------------------------------|-------------------------------------------------------------|
| the least squares method            | 100%                                                      | 94.63%                                                        | 1.43%                                                    | 0.27%                                                       |
| The adaptive step size artificial fish swarm algorithm | 100%                                                      | 98.27%                                                        | 70.13%                                                   | 69.81%                                                      |

From Table 1, it can be seen that the traditional least squares method is comparable to the algorithm proposed in this paper in the full time electricity consumption anomaly scenario, and the algorithm proposed in this paper is more accurate in the multi-user full time scenario, while in the partial time electricity consumption anomaly scenario, the accuracy of the algorithm proposed in this paper is much better than the traditional algorithm. The accuracy of the traditional algorithm in the partial time anomaly electricity consumption scenario of the traditional algorithm is only 1.43% and 0.27% for single-user and multi-user, respectively, while the accuracy of the proposed algorithm can reach about 70%, thus providing a new and effective algorithm for the identification of partial time abnormal electricity consumption scenarios.

4. Conclusions
This paper presents a low-voltage anomalous user identification method based on the improved fish group algorithm, and verifies its feasibility through the experimental data of the IEEE standard test node simulation. The results show that the algorithm can well identify the abnormal power consumption of many users and compensate for the defect that traditional least squares which is only applicable to linear abnormal electricity scenarios in whole time. Moreover, the stability of this algorithm identification is a key point to be further studied in future work.

Acknowledgement
This paper is sponsored by Science and Technology Project of Jiangxi Provincial Department of Education. (GJJ190969)

References
[1] Dos Angelos E W S, Saavedra O R, Cortes, O.A.C, et al. Detection and Identification of Abnormalities in Customer Consumptions in Power Distribution Systems[J]. IEEE Transactions on Power Delivery, 2011, 26(4): 2436-2442.
[2] Zheng K, Chen Q, Wang Y, et al. A Novel Combined Data-Driven Approach for Electricity Theft Detection [J]. IEEE Transactions on Industrial Informatics,2018, PP:1-1.
[3] Cheng Chao, Zhang Hanjing, Jing Zhimin, et al. Anti-power theft study based on outlier algorithm and electricity information collection system [J]. Power System Protection and Control, 2015 (17): 69-74.
[4] Passos Junior L A, Oba Ramos C C, Rodrigues D, et al. Unsupervised non-technical losses identification through optimum-path forest[J]. Electric Power Systems Research,2016,140(nov.): 413-423.
[5] Zhuang Chijie, Zhang Bin, Hu Jun, and other abnormal power mode detection of power users based on unsupervised learning [J]. Chinese Electrical Engineering Journal, 2016, 36 (2): 379-387.
[6] Nagi, Jawad, Keem, et al. Nontechnical Loss Detection for Metered Customers in Power Utility Using Support Vector Machines.[J]. IEEE Transactions on Power Delivery, 2010, 25(2):1162-1171.
[7] Xu Gang, Yuanpeng, Dai Tenghui. abnormal behavior detection under sparse random forest [J]. Grid technology, 2017,41 (06): 1964-1973.

[8] Zheng Z, Yatao Y, Niu X,et al. Wide & Deep Convolutional Neural Networks for Electricity-Theft Detection to Secure Smart Grids[J]. IEEE Transactions on Industrial Informatics, 2017:1-1.

[9] Hu Tianyu, Guo Qinglai, Sun Hongbin. Power theft detection based on stacking of related autoencoders and support vector machines [J]. Automation of power system, 2019, 43 (01): 119-125.

[10] Li Xiaolei. A new intelligent optimization method- -artificial fish group algorithm [D]. Zhejiang University, 2003.

[11] Ding Keliang, Shen yunzhong & Ou jikun. Methods of Line-fitting based on total least-square [J]. Journal of Liaoning Technical University(Natural Science), 2010,29(01):44-47.