Research on the strategy of locating abnormal data in IOT management platform based on improved modified particle swarm optimization convolutional neural network algorithm

Yuan Liu[1], Qiang Li[1], Di Cai[1], and Weizhi Lu[1]

[1]State Grid Henan Information & Telecommunication Company

January 18, 2023

Abstract

The IOT management platform is used to handle and transmit data from many types of power system terminal devices. The current IOT management platform has a low data processing efficiency and a high mistake rate when it comes to finding anomalous data. Furthermore, the effective selection and optimum decision of the convolutional neural network's structural parameters has a significant impact on prediction performance. Based on this, the paper proposes a decision algorithm for locating anomalous data in an IOT integrated management platform using a convolutional neural network (CNN) and a global optimization decision of key structural parameters of a convolutional neural network using an improved particle swarm optimization (APSO) algorithm. First, an index model is created to determine if the data retrieved from the IOT management platform is anomalous or not. Second, the structure of the convolutional neural network-based decision method for finding anomalous data is examined. Following that, an enhanced particle swarm optimization technique is developed to optimize the structural parameters of the convolutional neural network, and an APSO-CNN with improved performance for anomalous data localization is generated. Finally, the established algorithm's correctness, feasibility, and efficacy were evaluated using the Adam optimizer. The results reveal that the established APSO-CNN-based decision algorithm for anomaly data localization offers considerable benefits in terms of accuracy and running time, with extremely interesting application potential.

Hosted file

Research on the strategy of locating abnormal data in IOT management platform based on improved modified particle swarm optimization convolutional neural network algorithm

Liu Yuan [1] | Li Qiang[1] | Cai Di[1] | Lu Weizhi[1]

[1]State Grid Henan Information & Telecommunication Company, Zhengzhou, Henan 450052

Abstract

The IOT management platform is used to handle and transmit data from many types of power system terminal devices. The current IOT management platform has a low data processing efficiency and a high mistake rate when it comes to finding anomalous data. Furthermore, the effective selection and optimum decision of the convolutional neural network's structural parameters has a significant impact on prediction performance. Based on this, the paper proposes a decision algorithm for locating anomalous data in an IOT integrated management platform using a convolutional neural network (CNN) and a global optimization decision of key structural parameters of a convolutional neural network using an improved particle swarm optimization (APSO) algorithm. First, an index model is created to determine if the data retrieved from the IOT management platform is anomalous or not. Second, the structure of the convolutional neural network-based decision method for finding anomalous data is examined. Following that, an enhanced particle swarm optimization technique is developed to optimize the structural parameters of the convolutional neural network, and an APSO-CNN with improved performance for anomalous data localization is generated. Finally, the established algorithm's correctness, feasibility, and efficacy were evaluated using the Adam optimizer. The results reveal that the established APSO-CNN-based decision algorithm for anomaly data localization offers considerable benefits in terms of accuracy and running time, with extremely interesting application potential.

Key words: IOT management platform, sample data analysis, anomaly data localization, convolutional neural network, improved particle swarm optimization

1 | INTRODUCTION

With the continual advancement of technology and the advancement of social play, the operating stability and quality of power supply of the electric power system are receiving increasing attention [1]. However, the failure rate of power transmission lines remains very high due to physical structural loss of current power transmission equipment [2] and many complicated uncertainties in long-distance power transportation, which has a significant impact on regular production operations [3]. Based on this, a full IOT management platform capable of multi- dimensional data transfer and electric power system anomaly monitoring has been created [4]. The IOT management platform, as the basic support platform of the intelligent IOT system, is used for the unified management and data transmission of multiple categories of power system terminal devices, in order to realize standardized data access and sharing of terminal devices in various fields such as transmission, substation, distribution, customer side, and supply side [5]. When it comes to the access characteristics of various professional devices such as equipment side, customer side, and supply chain, the IOT management platform still suffers from insufficient device access flexibility management and insufficient control of heterogeneous devices [6]. In terms of lean and intelligent device operation and maintenance management, the IOT management platform has flaws such as insufficient monitoring of the entire link of devices and insufficient analysis and statistics of operation and alarm data, making it difficult to grasp device operation status and deal with abnormal problems of devices in a timely manner [7]. As a result, the investigation of IOT management platform data analysis and processing by corresponding control algorithms [8] and autonomous decision making of anomalous nodes [9] has significant practical prospects and engineering value.

The abnormal transmission data of the power transmission system in daily operation and maintenance is an essential data resource for the power system's abnormal monitoring equipment [10], which mostly originates from each abnormal link of each terminal category in real production and operation [11]. The IOT management platform may make the data features at the abnormal nodes more evident by processing and analyzing the abnormal data, allowing for faster decision making and location of power system issues [12]. However, due to the power system's extraordinarily enormous size, how to properly analyze the current data swiftly necessitates the direction of matching control algorithms. Existing data processing and analysis techniques include fuzzy clustering [13], K-means clustering [14-16], neural networks [17-20], and others. The literature [21] developed a deep neural network algorithm and clustering analysis-based approach for detecting abnormalities in sensor network data. The test results validated the developed algorithm's efficacy in recognizing and separating ambient noise and abnormal events in the network. The literature [22] proposes a network anomaly data stream mining approach based on enhanced clustering analysis. Real-time querying of data streams was enabled by developing a pre-processing model for anomalous network data streams [24]. The literature [25] suggested a data mining and neural network-based system for monitoring and evaluating abnormal data. The literature [23] developed a BP neural network algorithm-based technique for detecting irregular electrical energy metering data. Electrical energy measurement data are acquired and utilized to determine active power losses in this work. The technique described above offers substantial advantages in low-dimensional data processing and analysis. However, when utilized directly for parallel processing and analysis of power operation data based on the IOT management platform, difficulties such as sluggish processing speed and big decision mistakes may arise, and the benefits may be lost. As a result, more study is needed to develop a control algorithm with strong data processing performance and rapid operating speed [26].

In summary, the study develops a decision method based on graph convolutional neural network for identifying anomalous data in an IOT integrated management platform, which is utilized to increase locating accuracy and reduce data processing time. The study examines the index model used to evaluate data anomalies, as well as the structure and working path of the convolutional neural network-based decision algorithm for identifying anomalous data. Meanwhile, to prevent the inertia mistake produced by setting parameter values based on the designer's expertise, the study proposes the APSO method for global optimization of CNN structural parameters in light of the variety of structural parameter selection in convolutional neural networks. Finally, the established APSO-CNN algorithm's accuracy, practicality, and efficacy are examined in the MATLAB platform. The results show that, when compared to the traditional neural network algorithm for locating anomalous data, the designed APSO-CNN-based decision algorithm for locating anomalous data can significantly reduce the data processing pressure of the IOT integrated management platform and has a broad application prospect.

The rest of the paper is organized as follows: Section 2 describes the evaluation criteria and theoretical analysis of anomalous data in power systems, Section 3 develops the APSO-CNN algorithm for data processing and anomaly monitoring, and Section 4 conducts simulation tests and feasibility analysis of the established data processing algorithm guided by example data. Finally, Section 5 summarizes the paper's primary work and suggests future study areas.

2 | Abnormal data evaluation model

The idea behind monitoring and autonomous location of abnormal data in power transmission systems is to compare differences in data characteristics between groups to determine the probability and category of fault occurrence, and to output the corresponding information to assist managers in dealing with the source of the fault. The process of processing and evaluating the detected data should be carried out using the feature variables and feature values associated with the data. Creating a realistic data feature model aids in the speedy localization of aberrant data.

2.1 Rate of load change

After the operation data of each terminal device is uniformly transmitted to the IOT management platform for the power transmission system, it is possible to determine whether there is an abnormality in the group of data by counting the relationship between the load change rate of each group of data to be detected and the system's historical compound change rate range. The difference between the daily load curve and the characteristic curve of the data to be discovered can be represented as

$$D_a = \max\left|Y_a(x) - Y_t(x)\right|, x = 1, 2, \cdots, N \tag{1}$$

where D_a indicates the distance between the daily load curve and the characteristic curve of the data to be tested, $Y_a(x)$ indicates the corresponding value of each node of the daily load curve of the data to be tested, $Y_t(x)$ indicates the corresponding value of each node of the characteristic curve, N indicates the number of nodes.

From the difference between the daily load value and the characteristic value of the data to be tested, as described in equation (1), the corresponding rate of load change is calculated.

$$\beta_a(x) = \frac{Y_a(x) - Y_t(x)}{Y_a(x)} \tag{2}$$

where $\beta_a(x)$ represents the rate of load change for each sampling point of the data to be tested.

The load change rate of the data to be tested calculated by equation (2) can characterize the change of each sampling point on the daily load curve of the data to be tested, and compare it with the fluctuation range of the historical load change rate $\left[\beta_{a\min}(x), \beta_{a\max}(x)\right]$. When $\beta_a(x)$ is within the fluctuation range, the set of data to be tested is considered not abnormal, otherwise the set of data is marked and proceeds to the next processing stage.

When an anomaly is discovered at a sampling node of the data to be tested, the data is deemed to be corrected using the following equation based on the similarity of characteristics between the data to be tested and the characteristic curve in both horizontal and vertical dimensions.

$$Y_{ac}(x) = \frac{Y_t(x)}{2} \cdot \left[\frac{Y_a(x-1)}{Y_t(x-1)} + \frac{Y_a(x+1)}{Y_t(x+1)}\right] \tag{3}$$

where, $Y_{ax}(x)$ is the load change rate of the corrected data to be detected at the corresponding sampling node. $x-1$ and $x+1$ denote the previous node and the next node of the current sampling point, respectively.

When comparing the horizontal and vertical similarity of the data to be tested with the characteristic curve, if the sampling nodes of the data to be tested show continuous anomalies from node x to node y, the data correction is performed by the following equation.

$$Y_{ac}(x) = \frac{Y_t(x)}{2} \cdot \left[\frac{Y_a(x-1)}{Y_t(x-1)} + \frac{Y_a(y+1)}{Y_a(y-1)}\right] \tag{4}$$

2.2 Degree of data outliers

The number of outliers in the data acquired by the IOT integrated management system while monitoring each device of the power transmission structure may be used to describe the presence of abnormalities in the group of data to be discovered. The following equation may be used to compute the degree of outlier data of a node in the data set to be discovered.

$$S(i) = \sum_{j \in Q} \frac{\sqrt{\min\left\{\left(x_i^k - x_j^p\right)^2 \middle| k - h \le p \le k\right\}}}{d\left(s_i, s_j\right)} \tag{5}$$

where i denotes a data node, Q denotes the data set with the node i as the center of the circle, x_i^k and x_j^p denote the data of the node i after the normalization process of p , respectively, d denotes the distance from the node i to , j h denotes the total number of times.

The normalization process of x_i^k can be expressed as

$$x_i^k = \frac{\partial\left(m_{\max} - m_i\right)}{m_{\max}} \times 100\% \tag{6}$$

where m_i represents the data of node i at the first k sampling, $m_{\max}$ represents the maximum similar value of node i in the range at the first k sampling, ∂ is the range control parameter, which is used to ensure the normalization calculation process is carried out properly.

2.3 Information entropy

The term "information entropy" relates to the relevant theory of probability statistics, and the entropy value is utilized as a judgment indication of a random event's uncertainty. A greater entropy number implies more disorder in the data to be identified, and vice versa shows a high degree of orderliness in the set of data. As a result, when anticipating the abnormal nodes of the grid system operation data of the IOT integrated management platform, the information entropy may be utilized to define whether the data to be detected are abnormal or not. The following equation can be used to express the information entropy utilized in this study.

$$H(x) = -\sum_i p_i \log_2 p_i \left(i = 1,2,\cdots,n\right) \tag{7}$$

where $H(x)$ denotes the information entropy of the data to be detected x, which is used to characterize the disorder of the set of data, p_i denotes the probability of the anomaly of the i_{th} data node, which satisfies $\sum_i p_i = 1$.

According to the above analysis, the information entropy can be used as an index for judging whether an abnormality occurs in the electric power supply system. According to the power operation and maintenance data stored in the IOT management platform, the information entropy of abnormal data can be expressed as follows.

$$H(y) = -\sum_v \frac{p_v \log_2 p_v}{\alpha} \left(v = 1,2,\cdots,m\right) \tag{8}$$

where $H(y)$ denotes the information entropy when the power operation data is anomalous, p_v denotes the probability that the v_{th} power operation data is anomalous, which also satisfies $\sum_v p_v = 1$. α denotes the calculation factor.

2.4 Judgment index of correct decision rate

When the historical power operation data of the IOT management platform is extracted and predicted, the proportion of abnormal node data is very small when considering the scale of the power supply system and the high dimensional characteristics of the actual operation and maintenance data of various terminal devices. As a result, when the fraction of abnormal data in the overall database is relatively low, the standard prediction accuracy theory cannot adequately estimate the success of the specified abnormal data localization technique. As a result, as indicated in the following equation, this research creates a thorough evaluation index to evaluate the correctness of the decision algorithm proposed in the study for finding anomalous data.

$$\lambda_c = \frac{f_{cc}}{f_{cc} + f_{cw}} \times 100\% \tag{9}$$

$$\lambda_w = \frac{f_{ww}}{f_{wc} + f_{ww}} \times 100\% \tag{10}$$

$$\lambda_{all} = \frac{f_{cc} + f_{cw}}{f_{cc} + f_{cw} + f_{wc} + f_{ww}} \times 100\% \tag{11}$$

where λ_c indicates the proportion of normal data among all normal data, λ_w indicates the proportion of abnormal data among all abnormal data, λ_{cc} indicates the number of data that are actually normal and the decision result is also normal, λ_{cw} indicates the number of data that are actually abnormal but the decision result is also normal, λ_{ww} indicates the number of data that are actually abnormal and the decision result is also abnormal. λ_{wc} indicates the number of data that are abnormal but the decision result is normal, λ_{all} indicates the overall accuracy rate.

3 | APSO-CNN algorithm

The structure and complexity of the electric power supply system, as well as the quantity of different types of terminal equipment, are enormous. As a result, for the IOT integrated management platform, determining the abnormal fault state of a power system node and making it obvious is a significant technique to enhance power supply quality and raise economic efficiency. Based on this, the study develops an APSO-CNN-based decision algorithm for finding aberrant data in the power system, the major phases of which are depicted in Fig.1.

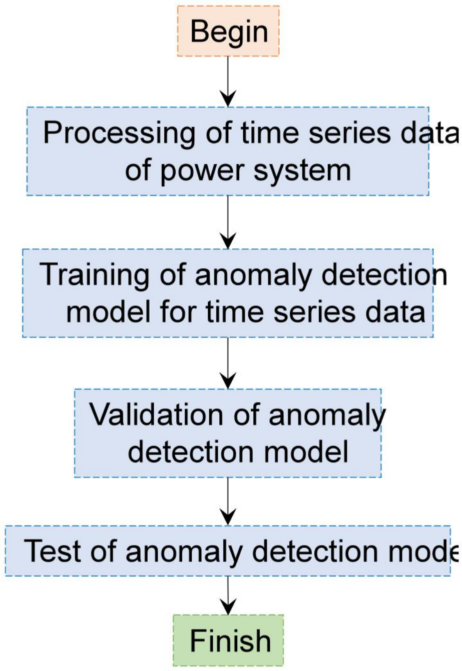

Fig.1 Anomaly data localization and decision steps of APSO-CNN algorithm

As seen in Fig. 1, when using the APSO-CNN algorithm to locate and decide abnormalities in power system operating data exported from the IoT integrated management platform, the data processing, model training, model validation, and model prediction phases are covered.

3.1 Data processing

Being used for analysis, the power system operating data given by the IOT integrated management platform needs be analyzed. Because of the redundant nature of the power supply system, some of the power operation data may not be complete time series, which might have an undesirable effect on the neural network model's training and testing. As a result, before training the graph convolutional neural network model, the raw power time series data supplied from the IOT integrated management platform is processed. The following requirements should be followed when processing data.

(1) Uniformity of time series dimensions

To ensure the neural network's training speed and forecast accuracy, the amount and dimensionality of the data to be identified must be consistent. As a result, the first exported data from the IOT management platform is processed in accordance with the requirement of time series dimensional uniformity, which includes information on the format, units, dimensions, and length of the values.

(2) Correction of incorrect data

The recorded power data may contain mistakes in data types, data formats, and misalignment values due to the complexity of the power system and the comparatively sluggish data processing speed of the IOT management platform. As a result, the processing of such data is critical. Duplicate and missing data are examples of misplaced values. This article classifies duplicate data according to specified standards, evaluates whether there is duplicate data in the group to be tested by the similarity of neighboring data, and selectively eliminates some duplicate data under the premise of guaranteeing the overall data structure's logic. This study decides to supplement the null values for missing data using the associated data rules. By analyzing the distribution of data features using the mean or median method to supplement the null data.

(3) Deletion of major error data

According to the idea of (2) for the processing of numerically misaligned data sequences, when there are missing or duplicated data of large size to be detected, then the overall deletion is considered to ensure the accuracy and credibility of the overall data prediction. In addition, the removal of irrelevant data will also reduce the dimensionality and computational pressure of data processing, which is beneficial to the fast convergence of the graph convolutional neural network model.

3.2 Structure of CNN

Compared with the anomalous data feature localization method developed by the designer's artificial experience, CNN can extract deeper features of the anomalous data sequence of the joint management platform. the convolutional layer of CNN can directly perform feature extraction on the initial sample data sequence, avoiding the problems of inadequate feature extraction and high error rate of the manual feature extraction method, and ensuring the accuracy of anomalous data localization and decision making. CNN can be used by multiple the structure of CNN mainly includes input layer, convolutional layer, pooling layer, fully connected layer, Softmax classification layer and output layer, as shown in Fig.2.

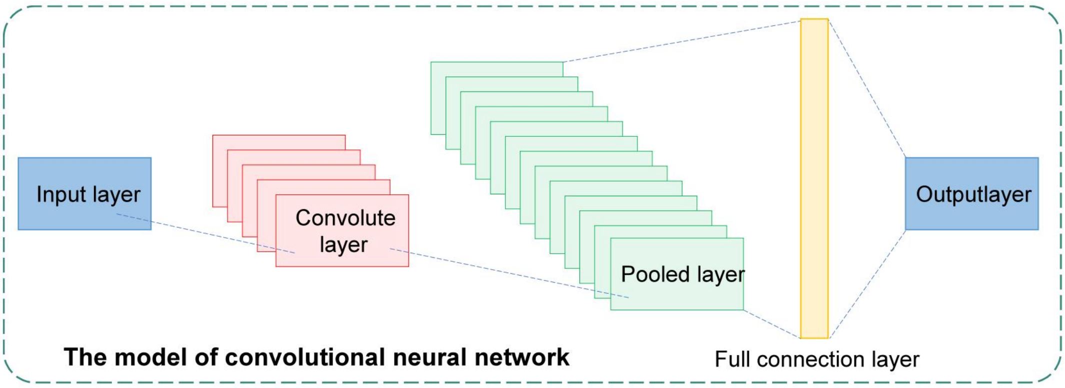

Fig.2 Basic structure of the designed CNN

The CNN model designed in the article for the application of anomaly data localization in the IOT management platform is designed with two convolutional layers, two pooling layers, one activation layer, and one fully connected layer, according to the role of each layer structure and the processing characteristics of the data. In the real connection, the architecture of the convolutional and pooling layers intersects. The pooling type is VALID, and the kernel of the two pooling layers is specified as 3*1. The role of each layer structure in processing the initial data sequence to be detected output from the IOT management platform may be generally defined as follows.

(1) Input layer

The CNN's input layer receives the data sequence to be detected from the IoT management platform. When inputting the CNN model, the data sequence to be detected must first be pre-processed as stated in Section3.1 and turned into a data sequence that meets the CNN model format requirements.

(2) Convolutional layer

The convolutional layer can be regarded as a set of N*N filters combined in a certain order, with multiple mapping features. After receiving the data sequence to be detected transmitted from the input layer, the convolution layer is calculated according to the parameters of convolution kernel size and convolution kernel move step set by the implementation, and the sum of the parameters in the convolution kernel and the corresponding position points after multiplying them in sequence is used as the result of the convolution layer. After calculating the convolution in the order of arrangement, the final output data, i.e., the convolutional features of the convolutional kernel size, is input to the next layer of the structure. The process can be described by the following equation.

$$A_l^i = f_{(l)} \sum_{x=1}^{m} \left(A_x^{i-1} \otimes W_{xl}^i + q_l^i \right) \tag{12}$$

where A_l^i denotes the l_{th} neuron at the output of the i layer, $f_{(l)}$ denotes the activation function at the layer i, m denotes the number of neurons, W_{xl}^i denotes the weight coefficient between the x neuron and the l neuron at the output of the i_{th} layer, q_l^i denotes the bias of the l neuron at the output of the i layer, $\otimes$ denotes the sign of the convolution operation.

The activation function $f_{(l)}$ in the above equation can be set to different forms to calculate the convolutional features of the convolutional layer. The following equation is chosen in the article as the formula for the activation function $f_{(l)}$.

$$\tanh(A) = \frac{e^A - e^{-A}}{e^A + e^{-A}} \tag{13}$$

(3) Pooling layer

The pooling layer, which contains both maximum pooling and average pooling, is often situated between the convolutional layers and is used to minimize the spatial size of the feature data acquired from the convolutional layers. Because the data samples created by the IOT management platform during everyday operations are exceedingly big, the influence of unknown disturbance and noise elements is significant. As a result, the article opts for maximal pooling for future data modification. The largest value in the pooling window is the pooling output corresponding to the location of the input data sequence. The maximum pooling calculation technique has the following formula

$$A_l^i = \max\left(\sum_{x=1}^{m} \left(A_x^{i-1} \right) \right) + q_l^i \tag{14}$$

(4) Fully connected layer

Each node in the fully connected layer is linked to all nodes in the preceding layer, allowing the data features generated from the previous layer to be fitted and computed, and the results of the fitting to be fed to the Softmax classification layer to complete additional classification processes. This layer has the most structural parameters due to the structural properties of the completely linked layer. The completely linked layer's work may be stated as follows:

$$A_l^i = f\left(A_x^{i-1} \times D_{lo}^i + q_l^i \right) \tag{15}$$

where D_{lo}^i denotes the learnable parameters of the i layer.

(5) Softmax classification layer

The Softmax classification layer is used to output the probability of occurrence of the corresponding fault class of the data sequence to be detected in the IOTC management platform and can be expressed by the following equation.

$$P_k = \frac{e^{b_k}}{\sum_{x=1}^{n} e^{b_x}} \tag{16}$$

where P_k denotes the probability that the data sequence to be detected in the IoM platform corresponds to the fault category of k, n denotes the total number of fault categories of the data sequence to be detected, b^k denotes the neurons to be activated in the output layer.

In addition, the value of the loss function can be reduced by the gradient descent method shown in the following equation when CNN is applied to the abnormal fault prediction and localization of data sequences with the IOT management platform as the carrier. The commonly used loss functions are mean square error function and cross-entropy function in two categories, and the cross-entropy function is selected in the article.

$$Q_s = \sum_{x=1}^{n} Q_{d\_x} \log\left(P_k\right) \tag{17}$$

where Q_s denotes the loss function value, Q_{d_x} denotes the expected value.

According to the role of each layer structure and the processing characteristics of the data, the CNN model designed in the article for the application of anomaly data localization in the IOT management platform is designed with two convolutional layers, two pooling layers and one activation layer each, and one fully connected layer. The structures of the convolutional and pooling layers cross each other in the actual connection. Among them, the kernel of the two pooling layers is defined as 3*1 and the pooling type is defined as VALID.

3.3 Improved PSO algorithm

The local and global search ability of the PSO algorithm depends heavily on the control parameters of the particles. A larger value of inertia weight enhances the global search ability of the algorithm, and a smaller value of inertia weight enhances the local search ability of the algorithm, compared with the social cognitive ability, a larger self-cognitive ability will lead the particles to search in the whole search space, thus enhancing the local search ability of the algorithm, compared with the self-cognitive ability, a larger social cognitive ability will lead the particles to search locally, thus enhancing the global search ability of the algorithm. The global search capability of the algorithm is enhanced by the larger social cognitive ability, which leads to local search of particles compared to the self-cognitive ability. The optimization process of the traditional PSO algorithm can be described as follows.

(a) Initialize a particle swarm of size N and set its initial position and velocity.

(b) Calculation of the fitness value for each particle by reference.

(c) For each particle, the adaptation value is compared with the best position it has experienced p_{is}. If it is better, the position is recorded as the current best position.

(d) For each particle, the adaptation value is compared with the adaptation value of the best position experienced globally at p_{gs}. If it is better, it is taken as the current global best position.

(e) According to Eq. The position and velocity of each particle v_{is} are updated for.

$$\begin{cases} v_{is}\left(t+1\right) = v_{is}\left(t\right) + a_1 r_{1s}\left(t\right)\left(p_{is}\left(t\right) - x_{is}\left(t\right)\right) + a_2 r_{2s}\left(t\right)\left(p_{gs}\left(t\right) - x_{is}\left(t\right)\right) \\ x_{is}\left(t+1\right) = x_{is}\left(t\right) + v_{is}\left(t+1\right) \end{cases} \tag{18}$$

where a_1 and a_2 are learning factors, r_1 and r_2 are mutually independent pseudo-random numbers that obey a uniform distribution on $\left[0,1\right]$.

(f) The corresponding solution is output when the termination condition is satisfied, otherwise go to step (b).

Considering the performance limitation of the traditional PSO algorithm between global optimal solution and local optimal solution, in order to improve the performance of PSO, an adaptive policy-based particle swarm optimization algorithm (SAPSO) is proposed so that it can well balance the local and global search ability of particles with the following adaptive update formula.

$$\omega(t+1) = (\omega_{\max} - \omega_{\min}) \times \delta + \omega_{\min} \tag{19}$$

$$c_1(t+1) = (c_{1s} - c_{1f}) \times \delta + c_{1f} \tag{20}$$

$$c_2(t+1) = (c_{2s} - c_{2f}) \times \delta + c_{2f} \tag{21}$$

$$\delta = 1 - \frac{t}{t_{\max}} \tag{22}$$

where $\omega_{\max}$ and $\omega_{\min}$ are the maximum and minimum values of inertia weights, respectively, c_{1s} and c_{1f} are the initial and final values of self-perception parameters, c_{2s} and c_{2f} are the initial and final values of cognitive parameters, $t_{\max}$ denotes the maximum number of iterations, in the adaptive strategy, $\omega_{\max} = 0.9$, $\omega_{\min} = 0.4$, $c_{1s} = c_{2f} = 2.5$, $c_{1f} = c_{2s} = 0.5$, take the value of 1000 for $t_{\max}$.

3.4 Optimization process of CNN structure parameters by APSO

According to the previous description, the structural parameters of the CNN model are relatively large, and it is difficult to obtain relatively efficient CNN performance by assigning parameters based on human experience. Therefore, based on the performance advantages of the APSO algorithm in global and local search, the article optimizes the structural parameters of the CNN through the APSO algorithm to obtain a more accurate model for locating anomalous data in the IoT management platform.

The APSO algorithm can initialize and update the velocity and initial position of the particles during each iteration, and calculate to obtain the individual extremes and global extremes for the current particle distribution. Therefore, when the APSO algorithm is used to optimize the structural parameters of the CNN, a certain number of particles are first generated randomly and initialized with the initial position and changing velocity of the particles. The parameters such as filter size, convolutional kernel size, and pooling layer size of each layer structure of the CNN model are optimized using the APSO algorithm. The assignment of the above structural parameters is carried out at the end of each iteration of the optimization process, and the value of the loss function trained by the CNN model is calculated by combining these parameters. Through multiple iterations, the optimal values of the above structural parameters are determined as the final parameters of the CNN model used for anomaly data localization decision in the IoT management platform. The specific optimization structure is shown in Fig.3.

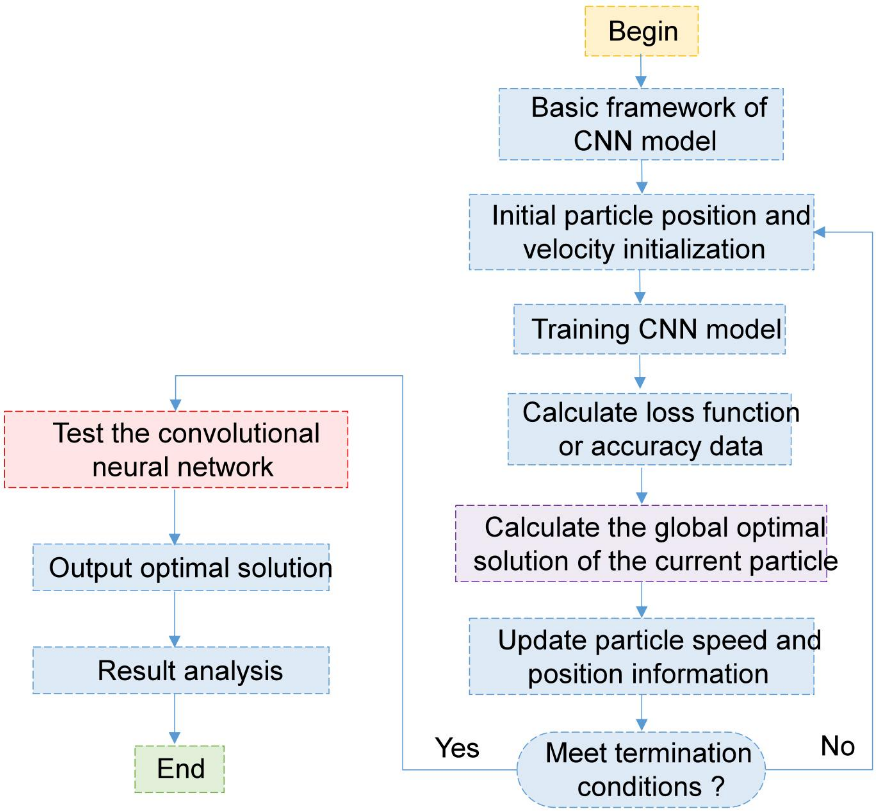

Fig.3 CNN model for locating abnormal data in IOT management platform based on APSO optimization

According to Figure 3, the main structural parameters of the CNN model designed in Section 3.2 are optimized with the help of the designed APSO algorithm. Assuming that the two convolutional layers are denoted as C1 and C2, respectively, five groups of particles are randomly generated when the structural parameters are optimized using APSO. The above five groups of particles are used to represent the number of filters in convolutional layer C1, the number of filters in convolutional layer C2, the number of neurons in the hidden layer, the value of the loss function in the convolutional layer, and the value of the loss function in the fully connected layer, respectively.

Define each of the above five groups of particles as $\gamma_i\left(i=1,2,\cdots 5\right)$. The first iteration of the optimization starts by initializing γ_i in the respective allowed variation range. The variation range of $\gamma_i\left(i=1,2,3\right)$ is (20,150), the variation range of $\gamma_i\left(i=4,5\right)$ is (0.1,0.99). The accuracy of anomalous data localization is used as the fitness function for global search, and the CNN model is endowed with autonomous search capability to obtain high accuracy characteristics in anomalous data localization.

4 | Simulation test

The method was simulated and examined with the aid of Adam optimizer to test the accuracy, practicality, and effectiveness of the idea of APSO-CNN algorithm based on abnormal operation data choice of power system presented in the article. In addition, the traditional neural network (unoptimized CNN) method is utilized as a comparison for simulation testing in order to evaluate the quick localization performance when using APSO-CNN to anomaly data localization decision.

Before the simulation begins, the time series of power system operating data acquired from the IOT integrated management platform are categorized, with 70% utilized for APSO-CNN training, 20% for model validation, and 10% for the final test. The maximum number of iterations for improving the structural parameters of the CNN via APSO is set to 50, with a particle count of 10 per group. To increase the accuracy and confidence of the data, 10 sets of tests were done for each data set during the simulation and summed as the final anomaly data prediction findings. As illustrated in Fig.4, the loss function values of the training and testing processes are continually lowered.

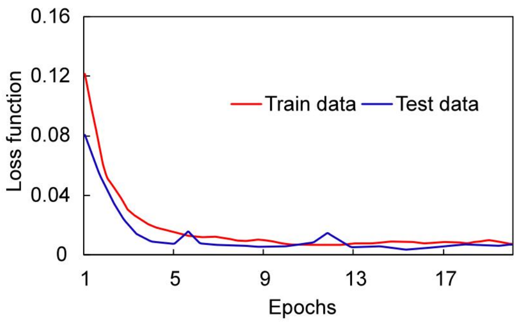

Fig.4 Variation in the values of the loss function

Fig.5 depicts the combined accuracy of the APSO-CNN-based anomaly data localization algorithm described in the paper versus the unoptimized classical CNN localization method. According to Fig.5, when analyzing the success rate of 10 sets of data to be identified, the APSO-CNN algorithm is consistently more accurate than the standard neural network in the decision making of anomalous nodes for power system operating data. Furthermore, because APSO-CNN is adaptable to sample data, prediction accuracy improves as the number of trials increases. This benefit, however, is not represented in the typical neural network technique. The typical neural network's prediction accuracy even declines significantly as the number of trials increases. Fig.5 demonstrates the existing APSO-CNN-based anomaly data determination algorithm's great performance and broad applicability for quick localization of anomalous nodes in power system operating data.

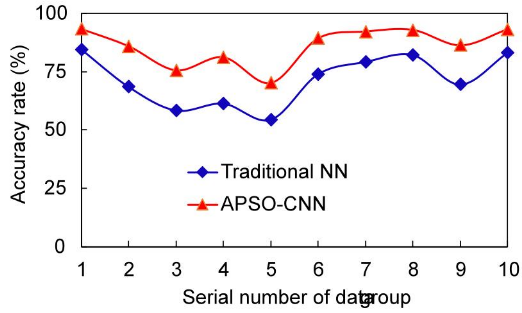

Fig.5 Combined accuracy of the two algorithms when applied to anomalous data localization

To further validate the benefits of the established anomaly data localization decision algorithm in terms of comprehensive usage performance, the paper conducted a comparative analysis of the time consumed in completing the process of detecting the required data and the number of groups, and the results are shown in Fig.6. The simulation time of the APSO-CNN based anomaly data localization choice method in each group is easily found to be less than that of the classic neural network technique. Given the large number of terminal devices linked to the IOT integrated management platform, timely decision of aberrant data is critical, as it will considerably decrease the data processing strain on the IOT integrated management platform and increase its work performance.

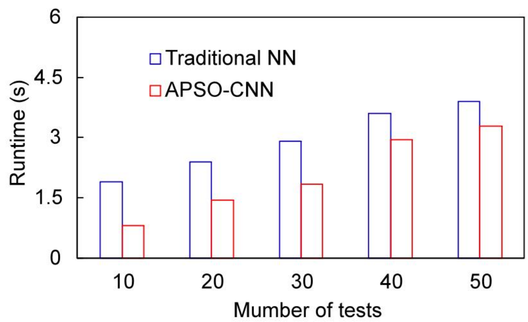

Fig.6 Time consumed by simulation

5 Conclusion

The paper develops an APSO-CNN-based anomalous data location determination algorithm for the IOT integrated management platform, which is utilized to increase location accuracy and reduce data processing time. The MATLAB simulation results reveal that the existing APSO-CNN-based anomalous data placement choice method has considerable advantages in terms of accuracy and running time. In comparison to the traditional neural network algorithm, the abnormal data localization decision algorithm with APSO-CNN can effectively reduce the data processing pressure of the IOT integrated management platform while ensuring normal power system operation and rapid localization of abnormal faults. We will continue to focus on the appropriate efficient data processing algorithms in future research to improve the working performance of the IOT integrated management platform.

REFERENCES

[1] Lv Fei, He Mingyi, Zong Weikang, Wang Lixu, Zhang Pan. Research and Application of Key Technologies of Internet of Things Access Service for Power Distribution [J]. Single Chip Microcomputer and Embedded System Application, 2022,22 (09): 37-40+66

[2] Liu Donglan, Zhang Hao, Wang Rui, et al. Design and application of full scene simulation detection system of intelligent Internet of Things system for energy Internet [J]. Shandong Power Technology, 2022,49 (02): 1-6+46

[3] Wei Xunhu. Discussion on Intelligent IoT System Framework Based on Power Internet of Things [J]. Science and Technology Innovation and Application, 2022,12 (04): 33-35. DOI: 10.19981/j.CN23-1581/G3.2022.04.09

[4] Reimers Clare E., Wolf Michael, Alleau Yvan, Li Cheng. Benthic microbial fuel cell systems for marine applications [J]. Journal of Power Sources,2022,522.

[5] Wang Zhaojun,He Yuhan,Li Xiao,Zhao Xi,Guo Hongxia,Guo Hongmei,Wang Zhelong,Liu Lijun,Ren Dawei,Liu Zhimei. research on New Energy Internet Socket and Power Management Platform System[J]. IOP Conference Series: Earth and Environmental Science,2021,632(4).

[6] Yongqing Zheng,Han Yu,Yuliang Shi,Kun Zhang,Shuai Zhen,Lizhen Cui,Cyril Leung,Chunyan Miao. PIDS: An Intelligent Electric Power Management Platform[J]. Proceedings of the AAAI Conference on Artificial Intelligence,2020,34(08).

[7] Microwave and Wireless Technologies, New Microwave and Wireless Technologies Findings from Ghent University Described (SIW antennas as hybrid energy harvesting and power management platforms for the internet of things) [J]. Telecommunications Weekly,2016.

[8] Semiconductor Integrated Circuits Companies, Semtech to Demonstrate Innovative Wireless Charging and Power Management Platforms at APEC 2016[J] . . Computers, Networks & Communications,2016.

[9] Zhang Chen Guang,Jiang Chao. Research on Integrated Management Method for Intelligent Community[J]. Advanced Materials Research,2012,614-615(614-615).

[10] Tower Semiconductor and Triune Systems Join on Power Management Platform[J]. Wireless News,2009.

[11] Zhiwei Wang,Sisi Zheng. Research on the Method of Abnormal Data Location in Large Scale Network Trading Platform[C]//.Information, Business and Management (Lecture Notes in Management Science Volume 49) under. ,2015:183-187.

[12] Shaoqing Ye,Lidong Lin,Peng Guo. Abnormal Data Detection and Cleaning of Wind Turbines Based on Average Confidence Interval Method[C]//. Proceedings of the 34th Chinese Conference on Control and Decision Making (11). [Publisher unknown], 2022:8-12. doi:10.26914/c.cnkihy.2022.022643.

[13] Qiao, Caicai,Wu, Chengmao,Li, -Changxing,Wang, Jiaye. Robust fuzzy clustering algorithms integrating membership guided image filtering [J]. Signal, Image and Video Processing,2022(prepublish).

[14] Awad Fouad H.,Hamad Murtadha M.. Improved k-Means Clustering Algorithm for Big Data Based on Distributed SmartphoneNeural Engine Processor[J]. Electronics,2022,11(6).

[15] Cao Wensi,Li Chen,Li Zhaohui,Wang Shuo,Pang Heyuan. research on Low-Resistance Grounding Fault Line Selection Based on VMD with PE and K-Means Clustering Algorithm[J]. Mathematical Problems in Engineering,2022,2022.

[16] Tuyishimire Emmanuel,Mabuto Wadzanai,Gatabazi Paul,Bayisingize Sylvie. Detecting Learning Patterns in Tertiary Education Using K-Means Clustering[J]. Information,2022,13(2).

[17] Hina Raja,M. Usman Akram,Arslan Shaukat,Shoab Ahmed Khan,Norah Saleh Alghamdi,Sajid Gul Khawaja,Noman Nazir. Correction to: Extraction of Retinal Layers Through Convolution Neural Network (CNN) in an OCT Image for Glaucoma Diagnosis[J]. Journal of Digital Imaging,2021,34(4).

[18] Zhenyu Luo,Feifan Huang,Huan Liu. PM 2.5 concentration estimation using convolutional neural network and gradient boosting machine[J]. Journal of Environmental Sciences,2020,98.

[19] Cerebrovascular Diseases and Conditions - Stroke, New Stroke Data Have Been Reported by Researchers at University of New Brunswick (Antlion Re- sampling Based Deep Neural Network Model for Classification of Imbalanced Multimodal Stroke Dataset)[J]. Computers Networks & Communications,2020.

[20] Imaging Technology - Digital Imaging, Findings on Digital Imaging Described by Researchers at Yozgat Bozok University [Automatic Segmentation of Meniscus in Multispectral MRI Using Regions with Convolutional Neural Network (R-CNN)] [J]. Journal of Engineering,2020.

[21] Li Min,Sharma Anil. Abnormal Data Detection in Sensor Networks Based on DNN Algorithm and Cluster Analysis[J]. Journal of Sensors,2022,2022.

[22] Jia, Xiaoqiang. Research on network abnormal data flow mining based on improved cluster analysis [J]. Distributed and Parallel Databases,2021(prepublish).

[23] Tao Li,Lei Chen. space event detection method based on cluster analysis of satellite historical orbital data[J]. Acta Astronautica,2019,160(C).

[24] Chen Yanyan. Abnormal Data Monitoring and Analysis Based on Data

Mining and Neural Network [J]. Journal of Sensors,2022,2022.

[25] Liu Shengnan,Ye Zhanqing,Yi Ling,Li Min. Identification method of abnormal data of electric energy measurement based on BP neural network algorithm[J]. Journal of Physics: Conference Series,2022,2260(1).

[26] Chen Hansi,Liu Hang,Chu Xuening,Liu Qingxiu,Xue Deyi. Anomaly detection and critical SCADA parameters identification for wind turbines based on LSTM-AE neural network[J]. Renewable Energy,2021,172.