sciendo
ISSN 1335-8871
MEASUREMENT SCIENCE REVIEW
Journal homepage: https://content.sciendo.com

Research of Online Error Prediction and Parallel Algorithm for Voltage Transformer in Smart Grid

Xianfeng Xin[1], Rongye Chen[2], Cong Lin[3], Qingchan Liu[3], Zhaolei He[3], Tengbin Li[3], Guangrun Yang[3], Orest Kochan[4,5]*, Mykhailo Karpa[5]

[1]*Shandong Taikai Detection Co., Ltd, Taian, 271000, China*
[2]*Shandong Taikai Instrument Transformer Co., Ltd, Taian, 271000, China*
[3]*Metering Center of Yunnan Power Grid Co., Ltd., Yunnan Power Grid Co., Ltd., Kunming, 650200, China*
[4]*School of Computer Science, Hubei University of Technology, Wuhan, 430068, China*
[5]*Department of Measuring-Information Technologies, Lviv Polytechnic National University, Bandery Str. 12, 79013, Lviv, Ukraine, orest.v.kochan@lpnu.ua, mykhailo.y.karpa@lpnu.ua*

Abstract: Online voltage transformer error prediction is an important research direction in the field of smart grids. This article mainly focuses on the online error prediction and the parallelization method of voltage transformers. First, an optimized multi-layer perceptron model based on the sparrow search algorithm (SSA) is proposed. The weight initialization process is optimized using the SSA to improve the prediction accuracy of the multi-layer perceptron. Considering the massive amount of data in real-world scenarios, a distributed sparrow search optimization algorithm for the multi-layer perceptron model was then developed, and the acceleration and scalability were tested on different data scales. In addition, transformer error prediction experiments were conducted to demonstrate the performance of the proposed algorithm.

Keywords: online error prediction of voltage transformer, sparrow search algorithm, multilayer perceptron, distributed platform

1. INTRODUCTION

Transformers play a crucial role in power systems by measuring current and voltage signals within the electrical grid. These devices, which are essential for the functioning of the power infrastructure, consist of measurement windings, protection windings, and metering windings. The metering windings are responsible for supplying precise current and voltage signals to the electricity meters, which in turn use these signals for accurate energy trading settlements. The accuracy of transformer output signals during operation has a direct impact on the economic interests of power supply, transmission, and consumption entities. Traditionally, the accuracy of transformer output signals is evaluated under offline conditions, especially during power outages. Transformers with higher accuracy grades are used as reference standards, and their signals are compared with those of the transformers under evaluation to determine their accuracy. However, as China progresses in the development of its smart grid, the planned downtime window for maintenance is continuously decreasing. The traditional operational model, which relies on power outages, is proving inadequate to meet the evolving demands of the grid. The growing conflict between the reliability of electricity consumption and the operational mode of transformer error assessment is becoming increasingly apparent to citizens. Therefore, there is an urgent need for the development of methods to assess transformer errors in a non-power outage state [1], [2].

The evolution of transformer error assessment technologies has been continuously updating and improving. In the early days of the power industry, transformers entering the grid had to undergo error detection during planned power outages. In this traditional method, which is still widely used in power grids, transformers with higher accuracy are used as reference standards and signals are compared to determine the accuracy of the transformers under evaluation. To mitigate the negative impact on grid operation during fixed intervals, some scientists have introduced the concept of "live testing" [3]. In this method, more accurate standard transformers are cut in on the high-voltage side of the operating grid to detect errors in real-time while the system is energized, thus avoiding economic losses due to power outages. However, "live testing" poses significant safety risks, including potential arc flash incidents and explosions of standard transformers during switching operations. In addition, standard transformers are not designed for prolonged operation in real power grids, making "live testing" impractical for widespread use [4]. Therefore, researchers

DOI: 10.2478/msr-2024-0033
Corresponding author: orest.v.kochan@lpnu.ua (O. Kochan)

worldwide have begun to explore evaluation techniques for running transformers without the need for standard physical instruments. Professor A.G. Phadke's team at Virginia Tech has proposed a physical modeling technique based on an equivalent circuit model. This method involves offline detection to identify a small number of accurate transformers as reference transformers [10]. A precise circuit equation is then established that includes both the reference transformer and the transformer under evaluation. The measured values of the reference transformer are used as known quantities to determine the error of the transformer under evaluation. However, the difficulty of determining reference transformers in practical scenarios leads to a low accuracy of the results. Based on this, Professor Thorp J of Virginia Tech has proposed an online estimation method for operating errors in transformers at both ends of transmission lines [11]. This method uses node current laws to establish linear equations that incorporate voltage and current transformer measurements, transmission line impedances, and other parameters. A set of transformers at both ends of the transmission line is called reference transformers. The measured values of these reference transformers, along with known line impedance parameters from acceptance documents, are used to solve for the unknowns in the equations. However, the discrepancy between the impedance parameters from the documents and the actual values often leads to a lower accuracy of the calculated results. The China Electric Power Research Institute has proposed an online group calibration technique for voltage transformers [5]-[6]. It proposes to use electromagnetic voltage transformers as temporary reference devices and calibrate voltage transformers with the same voltage and rating. However, this method has several application limitations, making it less suitable in many substations. In collaboration with researchers from Huazhong University of Science and Technology, a method based on principal component analysis was introduced to evaluate the measurement performance of voltage transformers in operation [7]. This approach uses principal component analysis to extract statistical features from the measurement data during operation to analyze the error of operating voltage transformers. However, this method has special requirements on the number and operating mode of voltage transformers. Yang Xuedong from Chongqing University proposed an online monitoring method based on the wavelet fractal theory [8]. In this method, the output signals of transformers are transformed, decomposed, and reconstructed by utilizing the advantages of wavelet transform in denoising and detecting singular values. Fractal theory is then used to calculate the box dimension of the processed signal, which enables the analysis of transformer errors. However, this method faces challenges in capturing small, gradual variations typical of error signals [9].

In order to achieve an accurate analysis of weak and small signal gradients, the project team is researching the online error prediction of voltage transformers and its parallelization for the assessment of operational transformer errors. The rest of the paper is organized as follows. In Section 2, we give an overview of the principles of transformer operation and online monitoring. Then, we propose a voltage transformer error prediction model based on a MLP optimized with the sparrow search algorithm (SSA) and design the distributed architecture in Section 3. In Section 4, a large number of experiments are conducted to demonstrate the effectiveness of the proposed algorithm. Finally, we draw conclusions in Section 5.

2. SUBJECT & METHODS

To measure the voltage or current information in the grid, we cannot measure high voltages and high currents directly on the grid. Therefore, we need to convert the grid information into safe and measurable signals via a transformer. The structure of the transformer is mainly a sensor between the primary and secondary side of the grid. The main difference between the transformer and the traditional transformer is the difference in the sensing coil. There are many types of transformer classification, based on the presence or absence of power requirements, the detection of differences in the object, the measurement of different roles, can be divided into a variety of different types of transformers, different classifications of the transformer may have a very different way of measurement, this paper focuses on the capacitance of the transformer for research.

A. Principle of operation of capacitive transformers

The capacitor voltage transformer (CVT) consists of the capacitor voltage divider (CVD) and the electromagnetic unit. The high voltage part of the CVT is taken over by the CVD, which significantly reduces the insulation requirements of the electromagnetic unit. The main components of the electromagnetic unit are the intermediate transformer, the damper, and the series reactor. The compensating reactor is connected in series after the CVD unit to make the CVT resonate in series at the rated operating frequency and thus improve the accuracy and load carrying capacity of the CVT. The fast saturation damper consists of a reactor and a damping resistor connected in series, in the event of a voltage error exceeding the rated value, the reactor can be quickly saturated, the inductance value decreases rapidly, and the circuit current loses a large part of the damping power in the damping resistor, so that the ferromagnetic resonance can be suppressed. The open structure of the CVD part of the CVT makes it susceptible to the action of the external electric field when it generates a voltage divider, and the Trench company provides a typical CVT profile, as shown in Fig. 1. According to Fig. 1, the CVT mainly contains a capacitive voltage divider (C1, C2) and an electromagnetic unit consisting of an intermediate voltage transformer T, a compensation reactor L, a damper D, etc., and the carrier device can be connected between *N* and *X*. The electromagnetic unit part of CVT (including intermediate transformer, series reactor, damper, etc.) is encapsulated in the oil-filled bottom box. The electromagnetic unit of the CVT is virtually unaffected by external electromagnetic fields due to the electromagnetic shielding effect of the oil-filled base case. The basic CVT wiring is shown in Fig. 2, where C1 is the high-voltage capacitor, C2 is the medium-voltage capacitor, L is the compensation reactor, T is the intermediate transformer, 1a1n, 2a2n, 3a3n are the secondary winding terminals, dadn is the residual winding terminal, and ZB1 and ZB2 are the dampers.

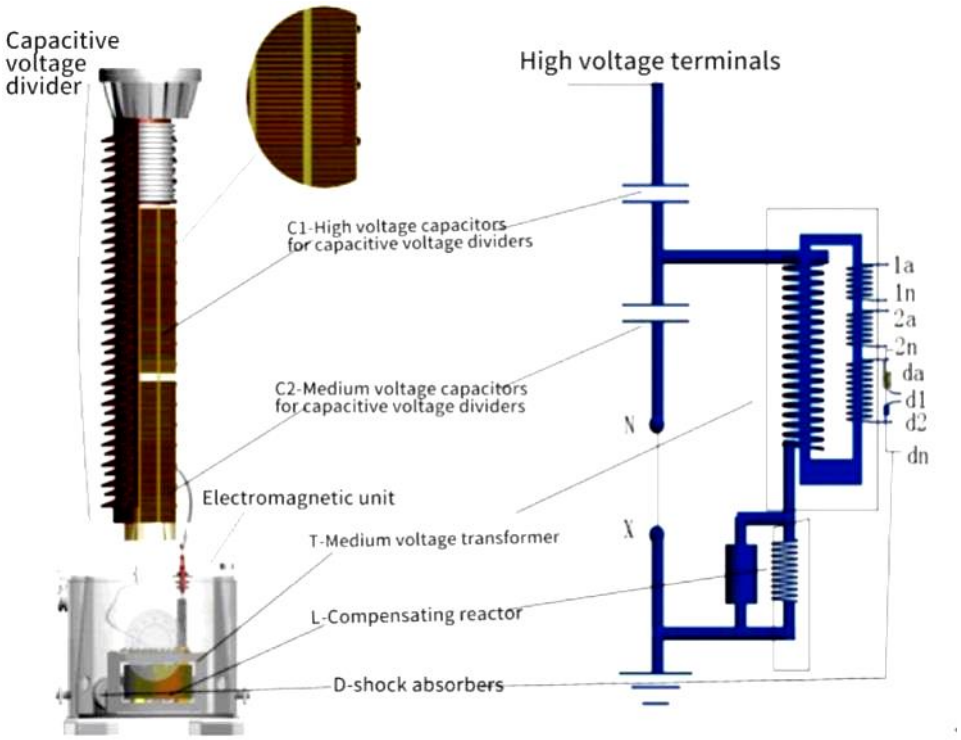


Fig. 1. Schematic diagram of a CVT profile.

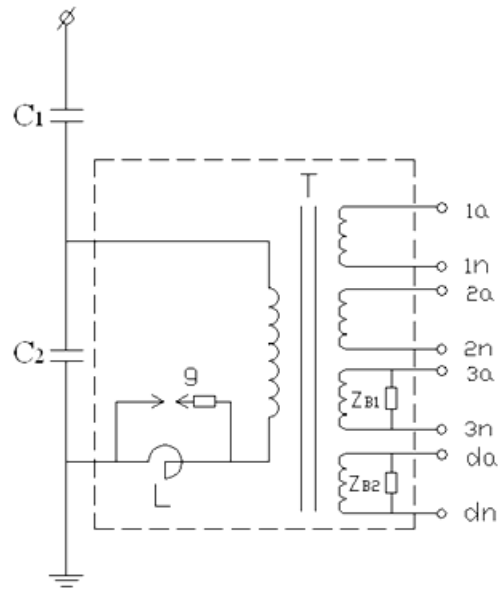


Fig. 2. CVT schematic circuit diagram.

The CVD unit reduces the high voltage at the input to a medium-low voltage suitable for the secondary side. Its main body is a voltage divider capacitor in series. Each section of the capacitor is approximately 10000 pF-25000 pF, with the different voltage levels of the CVT representing different capacitor capacity. The 110 kV CVT consists of a single capacitor section, the total capacitor is approximately 10000/20000 pF; the 220 kV CVT consists of 2 capacitor sections, each section of the capacitor is approximately 10000/20000 pF, total approximately 5000/10000 pF; the 330 kV CVT consists of 3 capacitor sections, each section of the capacitor is approximately 15000 pF, totaling 5000 pF; the 500 kV CVT consists of 3 or 4 capacitor sections, each section of the capacitor is approximately 15000/20000 pF, totaling 5000 pF; the 750 kV CVT consists of 4 sections of the capacitor, each section of the capacitor is approximately 20000 pF, totaling 5000 pF; the 1000 kV CVT has 5 capacitor sections, each consisting of about 25000 pF each, totaling 5000 pF.

The CVT unit can be considered as a two-port network, with a high voltage and ground at the input and a medium voltage and ground at the output. When the voltage at the input is U_1, the open circuit voltage U_C (intermediate voltage) at the output is the equivalent electromotive force and U_C can be expressed as follows:

$$U_C = U_1 C_1 / (C_1 + C_2) = U_1 / K_C \tag{1}$$

where $K_C = (C_1+C_2)/C_1$ is the voltage division ratio of the CVD part. The output impedance obtained by short-circuiting the input is equivalent to the internal impedance, i.e., the high-voltage capacitor C_1 and the medium-voltage capacitor C_2 are connected in parallel. Since there is a dielectric loss in actual operation, the internal impedance Z_C at the rated frequency f_N can be expressed as follows:

$$Z_C = R_C + jX_C \tag{2}$$

where $X_C = 1/\omega N (C_1+C_2)$. Since Z_C is large, the CVT does not have a high load carrying capacity. Consider connecting the equivalent capacitive reactance XL in series with the output circuit so that it resonates with the equivalent capacitance (C_1+C_2) at the rated frequency to improve the load capacity of the CVT.

The CVT error consists of the ratio difference (%) and the phase difference δ(′). The ratio difference results from the deviation between the actual ratio and the rated ratio, the expression of which is shown in (3):

$$f_U = (k_N U_2 - U_1) / U_1 \quad (3)$$

where k_N is the rated ratio of the CVT, U_1 is CVT primary side voltage, and U_2 is CVT secondary side voltage. The reason for the phase difference is that the phase angle between the output voltage and the input voltage does not match. The phase difference is expressed by subtracting the phase angle of the primary voltage from the phase angle of the secondary voltage. If the value is positive, this means that the secondary voltage is ahead of the primary voltage and, conversely, that it is behind the primary voltage. The expression for the calculation is shown in (4):

$$\delta_U = \varphi_{U2} - \varphi_{U1} \quad (4)$$

where φ_{U1} is the primary voltage phase; φ_{U2} is the secondary voltage phase, respectively.

B. Online monitoring of transformers

The online monitoring of transformers is to detect the secondary output signals of the metering windings of several groups of operating transformers in the station, realize the construction of a "virtual standard" by establishing the electrical relationship between the group of transformers, and finally calculate the error value of the operating transformers according to the error calculation formula.

The device for voltage transformer online monitoring is installed in the main control room of the substation screen cabinet, and the parallel connection with the energy meter, the collection of voltage transformer metering winding output voltage signals, only through the collection of data to complete the internal monitoring of the voltage transformer error calculation. The installation structure of the device for voltage transformer online monitoring is shown in Fig. 3.

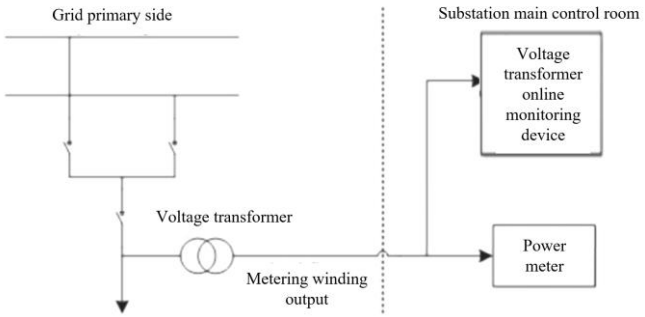


Fig. 3. Installation structure of voltage transformer online monitoring device.

Voltage transformer metering winding is generally 0.2 level, that is, the normal state of operation of the voltage transformer error should meet the following range: ratio difference $\leq \pm 0.2$ %, phase difference $\leq \pm 10'$. Reference transformer blackout offline error detection, with the accuracy level of the physical standard should be theoretically higher than the inspected transformer 2 accuracy level, that is, the standard should meet the 0.05 level, that is, to meet the standard error: ratio difference $\leq \pm 0.05$ %, phase difference $\leq \pm 2'$. As a voltage transformer online monitoring device, the overall sampling accuracy also refers to the 0.05 level requirements.

In addition, there are usually several sets of voltage transformers in the substation, so the device for voltage transformer online monitoring should be able to collect at least 4 sets of 12 voltage transformer signals simultaneously. Considering the requirements of data analysis and calculation, the synchronization of all signals must reach the level of 1 μs. The structure principle of the device for voltage transformer online monitoring is shown in Fig. 4.

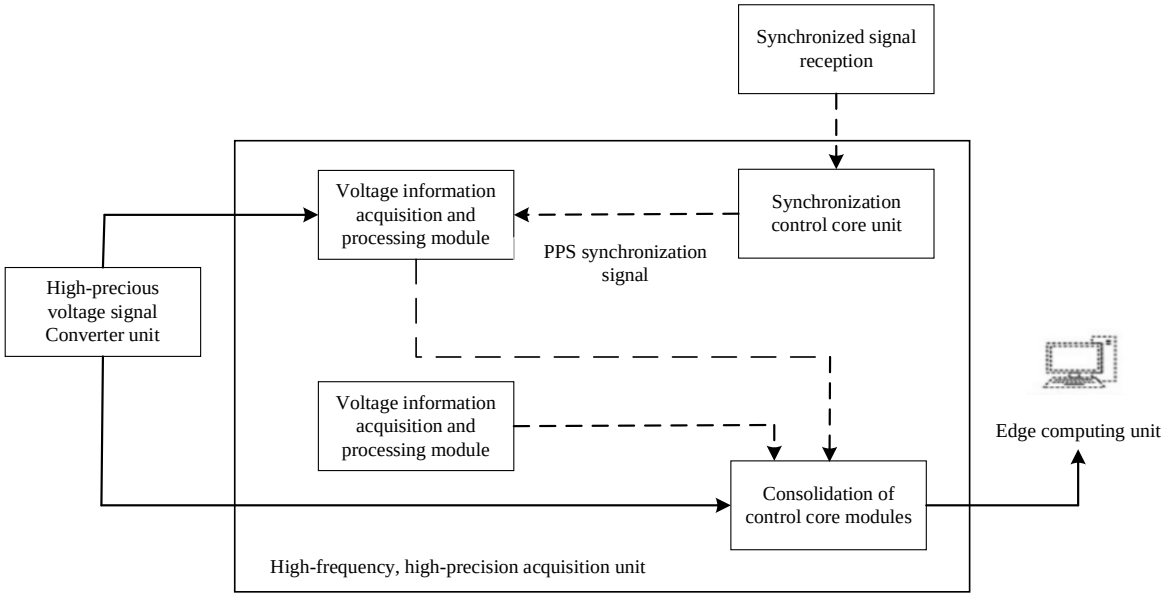


Fig. 4. Voltage transformer online monitoring device structure principle.

The appearance of the device for voltage transformer online monitoring is shown in Fig. 5.

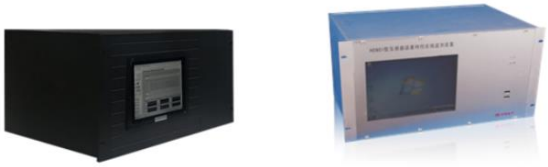


Fig. 5. Exterior photo of the voltage transformer online monitoring device.

3. METHODS

A. *Model design of the multilayer perceptron machine optimized by the sparrow search algorithm*

Structure of the multilayer perceptron machine

A multilayer perceptron (MLP) is a feed-forward neural network consisting of two layers of neurons that receive input signals from the input layer and pass them to the output layer to obtain the result. Fig. 6 shows a simple model of a functional neural structure. Each neuron obtains the cumulative value by accumulating one or more weighted input values, and then uses an activation function to perform a nonlinear transformation before passing the value to the next neuron.

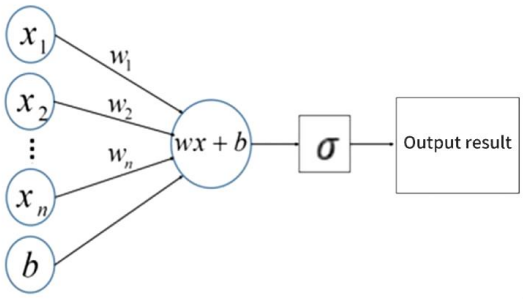


Fig. 6. Structural diagram of functional neurons.

1. Inputs: $x_1, x_2, x_3, \ldots, x_n$ are the inputs to the neuron with vector expressions:

$$X = [x_1, x_2, x_3, \ldots, x_n]^T \tag{5}$$

2. Weighted summation: $w_1, w_2, w_3, \ldots, w_n$ are the weights of the neurons and their vector expressions are:

$$W = [w_1, w_2, w_3, \ldots, w_n] \tag{6}$$

b is the offset of the neuron node. The neuron node first performs a weighted summation of the input vectors to obtain the implicit representation value:

$$z = \sum_{i=1}^{n} w_i x_i + b \tag{7}$$

3. Output: a function is applied to the weighted summed values to improve the representation of the neuron nodes.

$$o = \sigma(z) = \sigma(WX + b) \tag{8}$$

Since the perceptron model has only one layer of functional neurons, a better performance can only be achieved for simple problems. In order to make the model handle the problem in nonlinear situations, the method of increasing the number of hidden layers in the perceptron model is applied, which also increases the model's ability to strengthen the model's fitting ability, and finally a MLP model is obtained. It consists of an input layer, several hidden layers and an output layer with a characteristic multilayer structure. The layers are fully interconnected, and the model plane is shown in Fig. 7. Its main features are (1) full connection of neuron nodes between neighboring layers, (2) no connection of neuron nodes within the same layer, and (3) each connection has a weight.

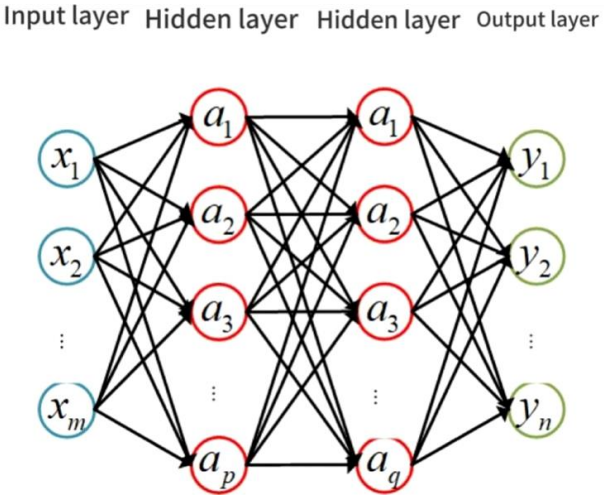


Fig. 7. Multilayer perceptual machine model.

The model can be updated using the error back propagation (BP) algorithm to update the weights and thresholds of all the individual neurons, and the final trained neural networks can be made to produce outputs that are closer to the true values. The training dataset is represented as $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, then for one of the training samples (x_k, y_k) the output of the MLP is $(\hat{y}_1 \ldots \hat{y}_n)$, and the error function of the dataset is defined as follows:

$$E = \frac{1}{n} \sum_{k=1}^{n} (y_k - \hat{y}_k)^2 \tag{9}$$

Then the weights W and the bias B in the MLP are updated iteratively according to the update law.

$$W^{(l)} = W^{(l)} - \alpha \frac{\partial E}{\partial W^{(l)}} \tag{10}$$

$$B^{(l)} = B^{(l)} - \alpha \frac{\partial E}{\partial B^{(l)}} \tag{11}$$

Steps of the SSA to optimize the initial weights of a multilayer perceptual machine

The MLP model is a model that simulates the neural connections in the human brain and can establish a kind of non-linear mapping connection between the input variables and the target variables through training and learning. At present, people often rely on experience to set the initial weight threshold of the multilayer perceptual machine or

choose a randomly generated initial weight threshold, but such setting of the initial weight threshold often affects the final convergence speed of the neural network training, takes more time and the training results may fall into the local optimum. Therefore, this paper introduces the SSA, which is used for global optimization of initial weight thresholds in multilayer perceptual machines and can improve the accuracy and convergence speed of multilayer perceptual machine training.

The steps of the SSA to optimize a MLP are as follows:

1. Determine the four-layer neural network.
2. Search the space using the SSA by encoding the initial weight threshold of the multilayer perceptron.
3. Evaluate the weights by using the absolute error of the MLP network training as the adaptation value weights, and finally use the optimal solution of the bird population as the initial weight threshold for neural network training.
4. Calculate the MLP training error and update the weights and thresholds according to the fitness function.
5. If the end condition is satisfied, output the optimal initial weight threshold, otherwise return to (2).

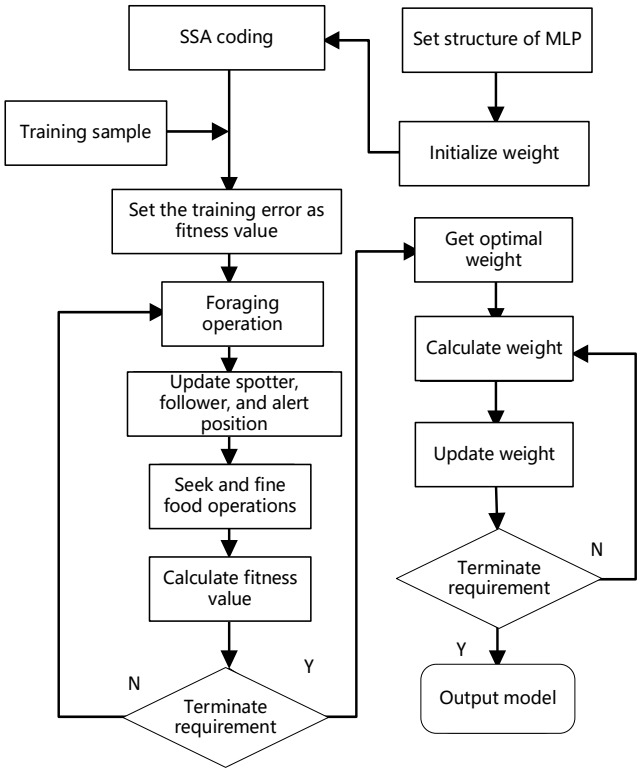


Fig. 8. SSA optimized MLP algorithm diagram.

The flowchart of the SSA-optimized multilayer perceptron is shown in Fig. 8. After optimizing the sparrow search algorithm, the sparrow individual with the optimal fitness is output to the built multilayer perceptron model as the initial weight threshold of the multilayer perceptron, i.e., the multilayer perceptron is given the optimal initial weight threshold, then the model is trained, and after the model continuously iterates to update the weight threshold, the trained multilayer perceptron model with the optimal weight threshold is finally obtained.

B. Distributed sparrow search optimization algorithm for MLP machines

SSA is an evolutionary algorithm. In the SSA execution process, the need to calculate a large number of individual sparrows, and after experiments found that, with the sparrow population settings becoming larger, the number of iterative search increases, thus, the greater the time complexity of implementing the SSA will be. When using the SSA to solve certain problems, one pass will not get the optimum results one expects. In general, it is necessary to find the optimal solution in several experiments, which increases the computational cost of the experiment. As people's data requirements increase, the data model becomes larger to a certain extent, the time complexity of the algorithm increases exponentially. The SSA in the stand-alone version of the implementation of the shortcomings of the algorithm is exposed. At present it is necessary to study how to solve this problem.

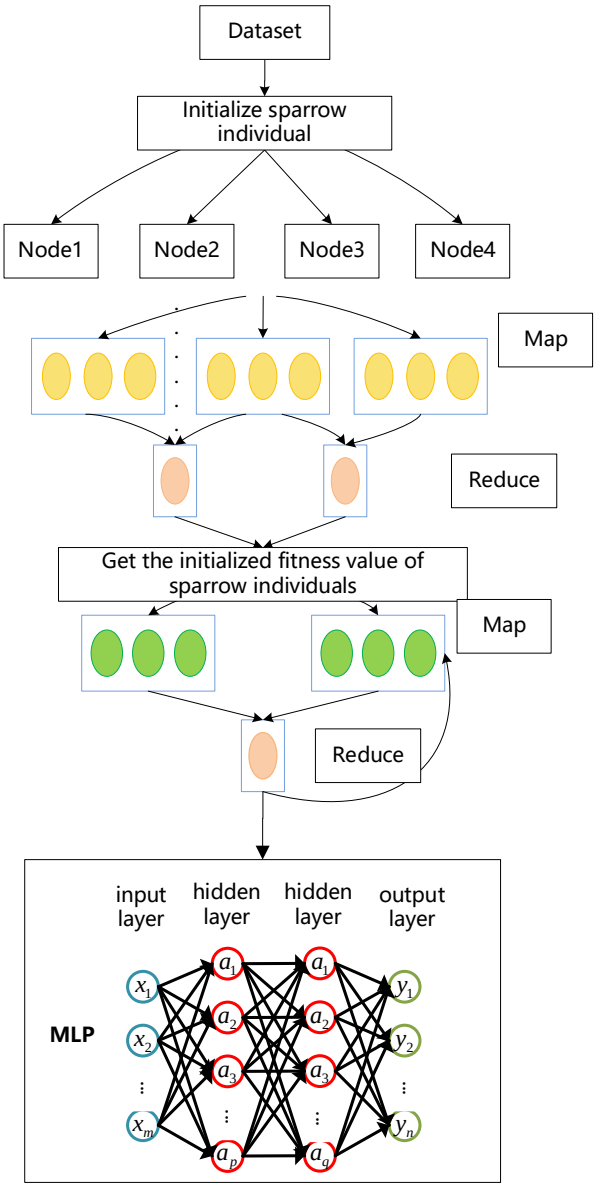


Fig. 9. Architecture of the distributed improved sparrow search optimized MLP algorithm.

Therefore, according to the advantages of the parallel computing framework MapReduce, this section combines the SSA and the parallel computing framework MapReduce to solve the shortcomings of the SSA with a large amount of time with a large amount of data and a large number of iterations, so that more experimental results can be obtained simultaneously.

The distributed improved sparrow search optimized MLP algorithm architecture diagram is shown in Fig. 9. The input for the Map function consists of two parts: the dataset stored in the HDFS and the location of the initialized population. First, this dataset is partitioned and each partition is stored in rows and entered as key-value pairs. In each Map operation, different features are selected according to the individuals and the accuracy obtained after K-fold cross-training of the samples is calculated as the fitness value of ESSA, iteratively until all samples in the partition have been processed. The update calculates the position of the finder and the position of the follower. To improve the efficiency of the reducer operation, the intermediate results can be sorted and partitioned according to the keys on the mapper, and then the interval number and updated position are sent to Reduce. The pseudo-code of the Map function is shown below:

Function Map (key: PopID, value: Pop_Sparrow)

```
Initialization:
Tfid=2
1:while t<max_iter do
2:     fitness←Fvalue(Pop_Sparrow, train_d,train_l)
3:     number←trainfloat.getNumPartitions()
4:     c1=2*math.exp(-(4*t/max_iter)**2)
5:     //updata finder sparrrow and follower sparrow
6:     poplist←trnasferFun(Pop_Sparrow,Tfid)
7:     new_key ← (partitionIndex,poplist)
8 :    new_value ← fitness
9:     Emit (new_key,new_value)
10:end while
```

In the Reduce phase, the populations of each partition are merged and individuals that exceed the boundary values are checked. The change in the optimal value and the change in the mean value of the fitness values are recorded for each generation. The results are written to HDFS in the form of interval numbers and fitness values, ready for the next round of MapReduce jobs. The pseudocode of the Reduce function is shown below:

Function Reduce (key: partitionIndex, poplist, value: fitness)

```
Initialization:
Poptemp=poplist[0][1]
1:for i<len(poplist) do
2:     pop_temp=np.vstack(poplist,pop_temp)
3:     // Individuals beyond boundaries become boundary
values
4:     popfitness← cal_obj_value(pop_temp)
5:end for
6:     //update Fit_food and Fit_best
7:     Fit_best=max(fitness )
8:     key ← (partitionIndex,poplist)
9:     new_value ← Fit_best
10:emit(key,new_value)
```

4. RESULTS

A. Dataset

We conducted experiments with simulated monitoring data from transformers at a power plant. We extracted the "amplitude" according to 15 test points: "Main transformer group I phase A", "Main transformer group I phase B", "Main transformer group I phase C": "Main transformer group II phase A", "Main transformer group II phase B", "Main transformer group II phase C", "Main transformer group III phase A", "Main transformer group III phase B", "Main transformer group III phase C", "Line I phase A", "Line I phase B", "Line I phase C", "Line II phase A", "Line II phase B", "Line II phase C", and based on the "frequency", the "magnitude" was extracted from 15 test points: "line II phase B", "line II phase C". B phase", "line II C phase" 15 test points to extract the "amplitude", based on the "frequency" consistent, so a total of 16 feature dimensions, and then according to the 15 labeled data, a total of 350000 records. Then the resulting data is disorganized and divided into training and testing sets according to the ratio of 7:3. Examples of transformer monitoring data are shown in Table 1.

Table 1. Example of transformer monitoring data.

	500 kV MTG I		500 kV MTG II		500 kV Line I	
	Measured value [V] / [°]	Ratio difference [%] / Phase difference [']	Measured value [V] / [°]	Ratio difference [%] / Phase difference [']	Measured value [V] / [°]	Ratio difference [%] / Phase difference [']
A Amplitude	59.8644	-0.00367488	59.859	-0.00902039	59.866	0.00267271
A Phase	177.752	-0.36	177.747	-0.3	177.739	-0.78
B Amplitude	59.8714	0.00100222	59.8435	-0.04659988	59.881	0.01603437
B Phase	57.767	0.06	57.778	0.66	57.76	-0.42
C Amplitude	59.992	-0.01383705	59.9864	-0.00933458	59.9994	0.01233498
C Phase	297.776	0.18	297.777	0.06	297.769	-0.42
A Amplitude	59.8659	-0.00333923	59.8609	-0.00835200	59.8683	0.00400896
A Phase	292.528	-0.3	292.523	-0.3	292.515	-0.78
B Amplitude	59.867	0.00033402	59.8395	-0.04593516	59.8768	0.01636962
B Phase	172.531	0.06	172.543	0.72	172.526	-0.3
C Amplitude	59.9839	-0.01399967	59.9782	-0.00950255	59.9914	0.01250336
C Phase	52.559	0.12	52.56	0.06	52.553	-0.36
A Amplitude	59.894	-0.00367297	59.8893	-0.00784720	59.8961	0.00350619
A Phase	66.858	-0.36	66.854	-0.24	66.845	-0.78

B. Acceleration ratio and scalability ratio experiments

We first conducted experiments with the distributed aspect. In the experiments, the metrics parallel scaling and acceleration of computation were used to evaluate the performance of the proposed distributed SSA algorithm based on the distribution. In particular, the acceleration ratio was introduced in the experiments to be able to verify how the cluster size affects the task execution time in the acceleration ratio metric, i.e., the ratio of the task execution time of the cluster to that of a server, denoted as:

$$speedup(\mathrm{i}) = t_1 \,/\, t_i \tag{12}$$

t_1 denotes the runtime on one node and t_i denotes the total runtime on the i^{th} node. In this experiment, 100 experiments are performed to test the acceleration of the algorithm. The experimental environment remains unchanged, and the number of computational nodes is increased from 1 to 7 in order to compare and analyze the acceleration ratios for different data sizes; D1 to D5 denote the five different sizes of datasets corresponding to 70000, 140000, 210000, 280000, and 350000 on-line monitoring data of mutual inductors, respectively.

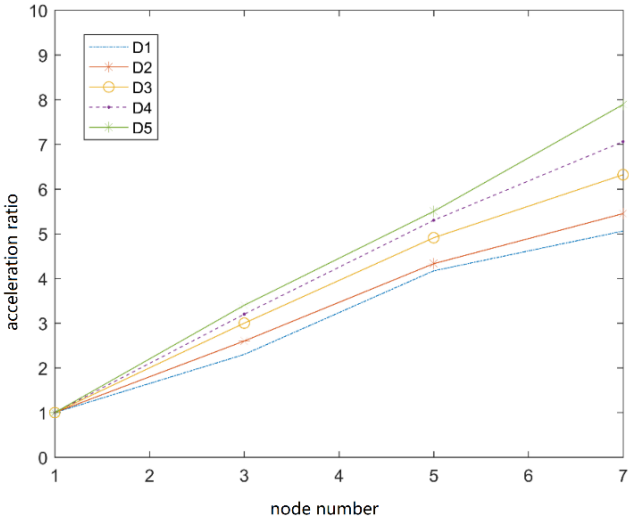


Fig. 10. Graph of acceleration ratio results for different data sizes.

As shown in Fig. 10, the acceleration ratio for all datasets does not change if the cluster nodes are set to 1. As the number of cluster nodes increases, the acceleration ratio scales well with the cluster size if the cluster nodes are smaller than 5. For a data size of D5, increasing the number of nodes effectively improves the acceleration ratio of the algorithm. If the number of nodes is increased to more than 5 and the data size is small, the acceleration then increases slowly due to the increased overhead caused by the greater communication between the servers.

Scale-up ratio is a speed metric that increases with the size of the dataset to effectively utilize the cluster nodes and evaluate the capability of the parallel algorithm. Where t_{si} is the runtime of the dataset of execution size S on the i^{th} node and t_{s1} is the running time of the dataset of execution size S on the 1^{st} node.

$$speedup(S, \mathrm{i}) = t_{S1} \,/\, t_{Si} \tag{13}$$

In Fig. 11, the scalability of the performance of the different algorithms was investigated by varying the number of nodes in each run to measure the runtime and scalability. To observe the effect of dataset size on performance scalability, the algorithms were analyzed for efficiency using variations of larger datasets with 7 nodes, and the runtime of all algorithms decreased almost linearly as the number of nodes in the cluster increased.

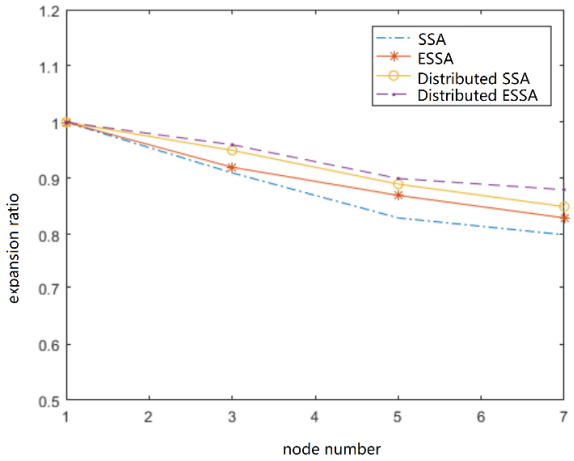


Fig. 11. Scalability ratio results for the D5 dataset.

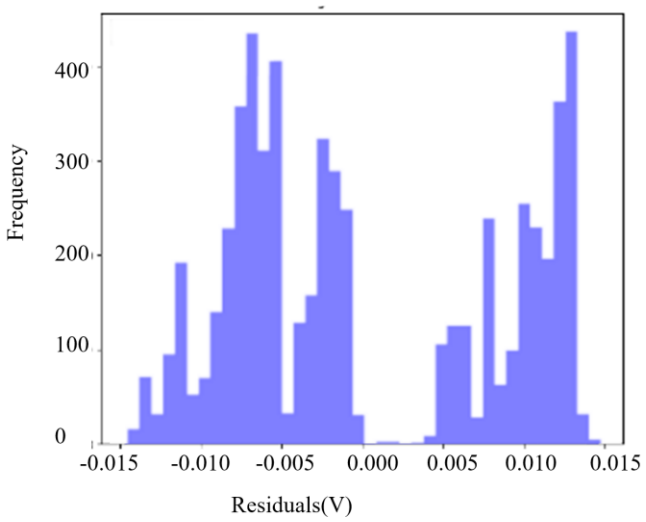
C. Transformer error prediction experiment

First, we compare the mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R-Square) of the models built with the particle swarm optimization multilayer perceptron (PSO-MLP), genetic algorithm optimization multilayer perceptron (GA-MLP), and SSA optimization multilayer perceptron (SSA-MLP), and the results are shown in Table 2.

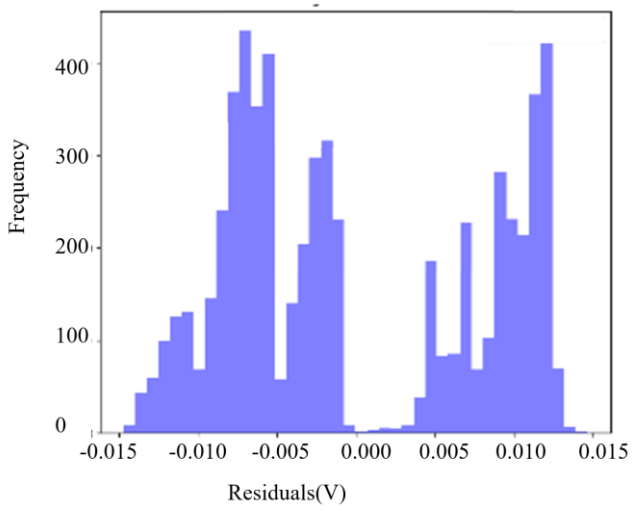
Table 2. Comparison and contrast algorithm evaluation results.

Methodologies	MAE [V]	RMSE [V]	R^2
PSO-MLP	0.0600	0.0040	0.88407
GA-MLP	0.1443	0.0040	0.814891
SSA-MLP	0.0167	0.0054	0.815631

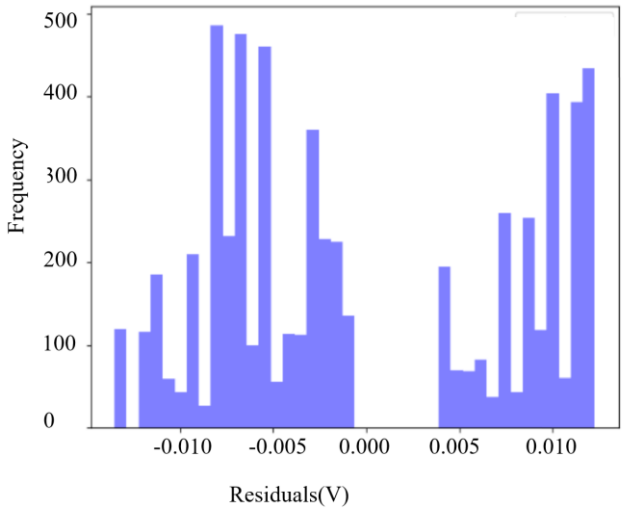
The comparison results of the residual histograms of the three algorithms are shown in Fig. 12. The results of the comparison of the three algorithms are shown in Fig. 13. The Q-Q plot is a type of scatter plot corresponding to the Q-Q plot of the normal distribution, i.e., the scatter plot with the quantile of the standard normal distribution as the horizontal coordinate, and the sample value as the vertical coordinate. To determine if the sample data is approximately normally distributed using the Q-Q plot, we only need to see if the points on the Q-Q plot are approximately close to a straight line and the slope of this straight line is the standard deviation, the intercept is the mean. The Q-Q plot can also be used to obtain rough information about the skewness and kurtosis of the sample. The prediction results of the three methods are plotted in Fig. 14, where the blue line shows the actual values and the orange line shows the predicted values.



(a) PSO-MLP

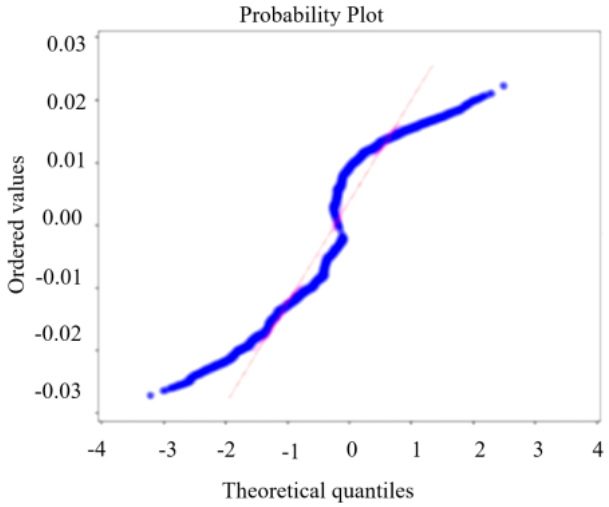


(b) GA-MLP

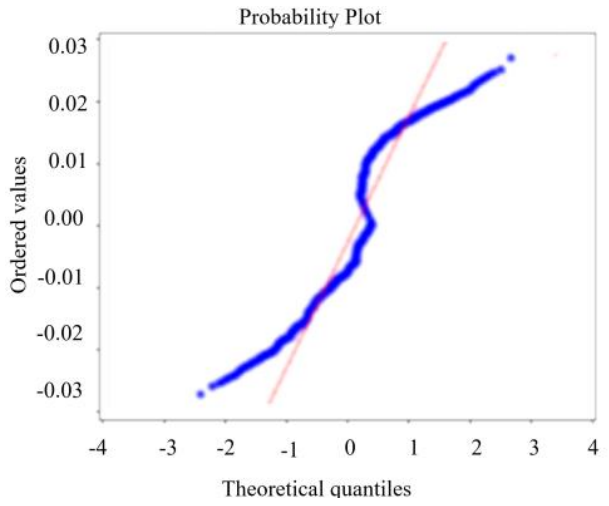


(c) SSA-MLP

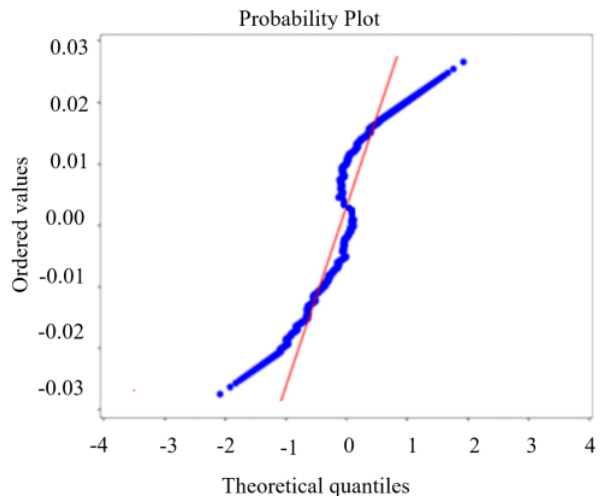
Fig. 12. Residual histogram comparison results for the three algorithms.



(a) PSO-MLP

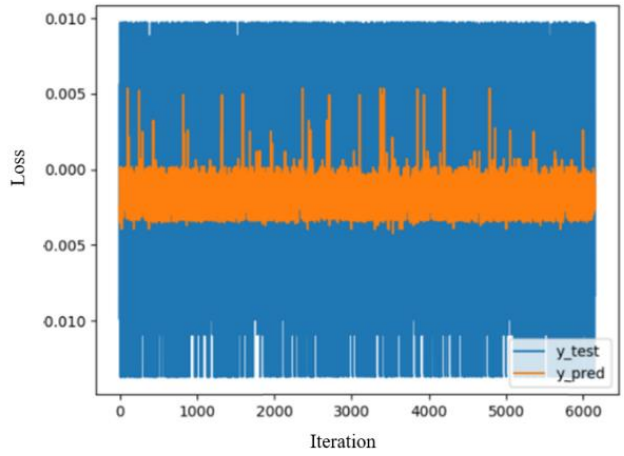


(b) GA-MLP

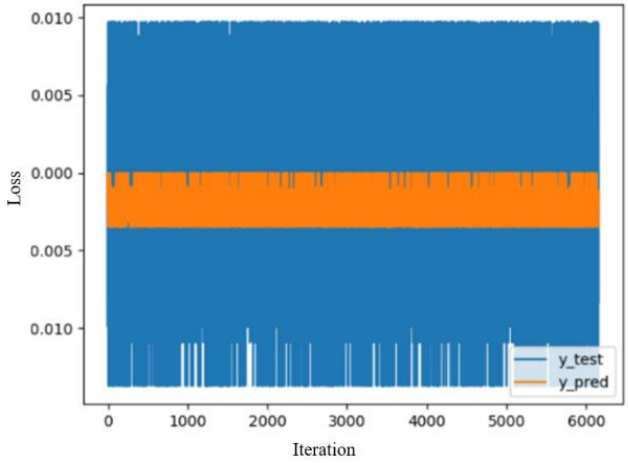


(c) SSA-MLP

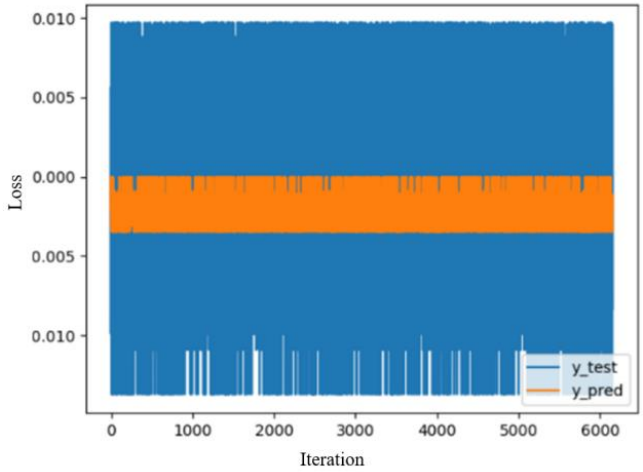
Fig. 13. Comparison results of Q-Q plots of the three algorithms.



(a) PSO-MLP



(b) GA-MLP



(c) SSA-MLP

Fig. 14. Comparison results of the prediction effect graphs of the three algorithms.

The complete error of each point is shown in Table 3. The experimental results show that the prediction errors of all the proposed methods are small.

Table 3. Complete errors at each point.

Point ID	Average absolute error [%]
MTG I A	$3.4839*10^{-4}$
MTG I B	$3.2072*10^{-5}$
MTG I C	$1.3902*10^{-4}$
MTG II A	$8.5602*10^{-6}$
MTG II B	$5.2492*10^{-6}$
MTG II C	$4.1639*10^{-5}$
MTG III A	$9.7242*10^{-6}$
MTG III B	$7.4469*10^{-5}$
MTG III C	$2.0402*10^{-6}$
“Line I A”	$5.3032*10^{-2}$
“Line I B”	$1.1202*10^{-3}$
“Line I C”	$8.5352*10^{-5}$
“Line II A”	$1.4852*10^{-5}$
“Line II B”	$2.0302*10^{-4}$
“Line II C”	$4.2032*10^{-5}$

5. Conclusion

This article mainly studies the online error prediction and parallelization method of voltage transformers. First, a MLP model optimized based on the SSA is proposed. The weight initialization process is optimized using the SSA to improve the prediction accuracy of the multilayer perceptron. Then, this article studied and analyzed the implementation principle and process of the parallel SSA based on Hadoop, and provided a detailed explanation of the overall process of the SSA and the specific design and implementation process of MapReduce. Based on this, a distributed sparrow search optimization MLP algorithm was developed and verified. The proposed algorithm helps to improve the measurement accuracy, prevent equipment failure, optimize the power grid operation strategies, and improve the stability and economy of the power grid. In the future, we may explore the application of deep learning algorithms for CVT measurement error prediction, such as using neural network models to perform nonlinear fitting and prediction on error data to improve the accuracy and reliability of prediction results.

References

[1] Wang, C., Bian, W., Wang, R., Chen, H., Ye, Z., Yan, L. (2020). Association rules mining in parallel conditional tree based on grid computing inspired partition algorithm. *International Journal of Web and Grid Services*, 16 (3), 321-339. https://doi.org/10.1504/IJWGS.2020.109475

[2] Wang, C., Xing, S., Gao, R., Yan, L., Xiong, N., Wang, R. (2023). Disentangled dynamic deviation transformer networks for multivariate time series anomaly detection. *Sensors*, 23 (3), 1104. https://doi.org/10.3390/s23031104

[3] Gao, R., He, W., Yan, L., Liu, D., Yu, Y., Ye, Z. (2024). Hybrid graph transformer networks for multivariate time series anomaly detection. *The Journal of Supercomputing*, 80 (1), 642-669. https://doi.org/10.1007/s11227-023-05503-w

[4] Crotti, G., Giordano, D., Sardi, A. (2011). Development and use of a medium voltage RC divider for on-site calibration. In *2011 IEEE International Workshop on Applied Measurements for Power Systems (AMPS)*. IEEE, 53-57. https://doi.org/10.1109/AMPS.2011.6090347

[5] Zhang, H., Zhang, D., Zhang, X. (2023). The role of output-based emission trading system in the decarbonization of China's power sector. *Renewable and Sustainable Energy Reviews*, 173, 113080. https://doi.org/10.1016/j.rser.2022.113080

[6] Zhang, H., Zhao, F., Hao, H., Liu, Z. (2023). Comparative analysis of life cycle greenhouse gas emission of passenger cars: A case study in China. *Energy*, 265, 126282. https://doi.org/10.1016/j.energy.2022.126282

[7] Li, H., Ma, C., Zhang, C., Chen, Q., He, C., Jin, Y. (2024). A knowledge-based cooperative co-evolutionary algorithm for non-contact voltage measurement. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 8 (2), 1142-1155. https://doi.org/10.1109/TETCI.2023.3300526

[8] Zang, X., Cao, Z., Xiao, M., Yang, X. (2023). Research on online monitoring of grid current transformer based on transformers and BiGRU. *Journal of Computational Methods in Sciences and Engineering*, 23 (3). https://doi.org/10.3233/JCM-226763

[9] Valtierra-Rodriguez, M. (2019). Fractal dimension and data mining for detection of short-circuited turns in transformers from vibration signals. *Measurement Science and Technology*, 31 (2), 025902. https://doi.org/10.1088/1361-6501/ab48ac

[10] Phadke, A. G., Thorp, J. S. (2008). *Synchronized Phasor Measurements and Their Applications*. Springer. https://doi.org/10.1007/978-0-387-76537-2

[11] Pal, A., Chatterjee, P., Thorp, J. S., Centeno, V. A. (2016). Online calibration of voltage transformers using synchrophasor measurements. *IEEE Transactions on Power Delivery*, 31 (1), 370-380. https://doi.org/10.1109/TPWRD.2015.2494058

Received February 22, 2024
Accepted November 27, 2024