JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 40, 71-88 (2024)
DOI: 10.6688/JISE.202401_40(1).0005

Build IPSO-ABiLSTM Model for Network Security Situation Prediction*

YA-XING WU[1] AND DONG-MEI ZHAO[2,3,4,+]
[1]*Hebei Institute of Mechanical and Electrical Technology*
Xingtai, Hebei, 054002 P.R. China
[2]*College of Computer and Cyber Security, Hebei Normal University*
[3]*Hebei Provincial Key Laboratory of Network and Information Security*
[4]*Hebei Provincial Engineering Research Center for Supply Chain Big Data Analytics and Data Security*
Shijiazhuang, Hebei, 050024 P.R. China
E-mail: zhaodongmei666@126.com

There are security risks in interaction and communication using wireless mobile networks, and network security situation prediction technology is to predict the next development trend with the previous and current network status, which can grasp the wireless mobile networks security status in time and make decisions in advance to avoid attacks. This paper proposes an Improved Particle Swarm Optimization Attention Bi-directional Long Short-Term Memory (IPSO-ABiLSTM) model for network security situation prediction. First, we construct the real situation values of the raw UNSW-NB15 dataset from the perspective of the impact of the attack on the situation indicator system, the sliding window method was introduced to reconstruct the situation values of the data set obtained by computing the data used for prediction. Secondly, the traditional PSO algorithm has the shortage of unbalanced search speed and tends to get local optimal solutions. The IPSO algorithm in this paper makes the global and local search ability of the algorithm more balanced and converges faster. Finally, the IPSO-ABiLSTM model is used to implement the situation prediction in different sliding window sizes. The experimental results show that the IPSO-ABiLSTM of this paper fits up to 0.9922, which verifies the effectiveness of the model proposed in this paper in the network situation prediction problem.

***Keywords*:** network security, situation prediction, attention, bi-directional long short-term memory (BiLSTM), improved particle swarm optimization (IPSO)

1. INTRODUCTION

Internet of Things (IoT), Internet of Vehicle (IoV) and 5G communication network and other wireless mobile networks technology is the research hotspot of this era, smart cameras, cars, drones and other devices are connected to wireless mobile networks, he made the world more intelligent. However, there are significant security risks for users using these devices to interact and communicate via wireless mobile networks [1], and if these devices are attacked and control is lost, it could lead to irreparable damage. Traditional security protection facilities include firewalls, intrusion detection systems and vulnerability scanning facilities, the main function of the firewall is to carry out access control, the main function of the intrusion detection system is to deal with the attack, vulnerability scanning is mainly used to find the existence of security risks, the function of these facilities is too single, cannot grasp the security situation in a timely manner,

Received October 10, 2022; revised October 31 & December 8, 2022; accepted December 13, 2022.
Communicated by Xiaohong Jiang.
*The preliminary version was presented at the Networking and Network Applications (NaNA) in 2022.

belongs to the passive protection system. Network security situation awareness is an active protection technology that requires comprehensive analysis of network security elements, assessment of the current network security situation and prediction of the trend of network security situation in the following period [2].

In this paper, we will conduct an in-depth study of network security situation prediction in situation awareness. By predicting the next network situation through the previous and current network conditions, this will allow dynamic mastery of wireless mobile networks conditions to make decisions in advance to avoid losses to users and have a positive impact on the protection of future mobile network security.

At present, many scholars have conducted research on the related technologies of network security situation prediction. The network security situation prediction analyzed by the time series method is to build a series of host hidden Markov models and make full use of multi-source heterogeneous information to dig deeper. The relationship between the network security situation of the preceding and subsequent time periods is predicted, and the network security situation of the next moment is predicted. At the same time, the security indicators of the hosts in the network are integrated to calculate the situation value of the next moment, which can well reflect the security situation of the network. However, the construction of the model and the calculation of the security situation in this paper are relatively complicated and difficult to implement [3]. Convolutional neural networks are also used in network security situation prediction. In order to enhance the learning ability of convolutional neural networks and reduce the training time of convolutional neural networks, some researchers have proposed a convolutional neural network based on a compound convolutional structure. The situation value is generated according to the hierarchical evaluation method. The experimental results show that the proposed method can predict the network security status very well, but the effect of the convolutional neural network on the time series prediction problem needs to be improved [4]. The diversification of devices in the Internet era makes the network environment more complex, and the attack detection is more difficult. Therefore, an integrated learning method is proposed to improve the accuracy. The experimental results show that the integrated classifier can effectively improve the accuracy and reduce the false alarm rate [5]. The Gravitational Search Algorithm (GSA) has the characteristics of strong global optimization ability and easy implementation, and the hyperparameters of the Support Vector Machine (SVM) are optimized through the GSA, and a GSA-SVM network security situation prediction model is constructed. The SVM with optimal parameters improves the network security situation prediction Accuracy, but the ability of SVM to predict time series is slightly insufficient, and the accuracy of situation prediction needs to be improved [6]. Combine the Cross-layer Particle Swarm Optimization with Adaptive Mutation (AMCPSO) algorithm and DS evidence theory to evaluate the current network conditions, and then introduce Fuzzy C-Means (FCM), Algorithms such as Hybrid Hierarchy Genetic Algorithm (HHGA) optimize the parameters and structure of traditional Radial Basis Function (RBF) neural network. Experimental results show that this method effectively improves the accuracy of situation assessment and prediction. The shortcoming is that there is not too much narrative on the generation of the situation value, and the reflection of the network security situation needs to be improved. In response to this shortcoming, the research team has recently carried out the establishment of the situation value and situation indicators, and the research will be reflected in this paper [7]. Using improved particle swarm optimization to

optimize the weights and thresholds of the Extreme Learning Machine (ELM) network improves the prediction accuracy and speed of the ELM, but it does not take into accounts the influence of other parameters in the ELM structure on the results [8]. Some researchers have proposed a prediction method based on particle swarm optimization algorithm and Long Short Term Memory (LSTM), which optimizes related hyperparameters of LSTM through particle swarm optimization, improve the prediction accuracy of LSTM [9].

Network security situation prediction is a time-series prediction problem, and there is a correlation between the before and after data. As we can see from the current research state, the first problem is that the traditional models used in the research method are mostly limited in prediction capability. And the prediction does not take into accounts the before and after information correlation between data, and it happens that there is a logical correlation between network attacks, which seriously inhibits its prediction effect, and the second problem is that the traditional calculation of network security situation value is complex and difficult to achieve. In this paper, we use Bidirectional Long Short-Term Memory (BiLSTM), which has been widely used in recent years for temporal problems, to solve the problem of model prediction capability. Adding Attention mechanism to BiLSTM can pay more attention to the impact of different input features on the output, and focused selective learning of the input can further improve the prediction effect of the network model by fusing BiLSTM and Attention as ABiLSTM model, ABiLSTM will fully consider the important features that have influence on the prediction results before and after the data to improve the prediction effect, but for the network model, the selection of hyperparameters is a tricky problem, the common parameter settings are empirical rule and algorithm, empirical rule is time-consuming and not reliable enough, this paper selects Particle Swarm Algorithm (PSO) to perform parameter optimization, in order to avoid the PSO local optimal solution In order to avoid the problem of local optimal solution of PSO, it is improved to speed up the convergence speed of the algorithm. In order to avoid the problem of local optimal solution of PSO, it is improved to speed up the convergence of the algorithm. Finally, IPSO-ABiLSTM network security situation prediction model is constructed in this paper, and the network structure of ABiLSTM is optimized by IPSO algorithm to further improve the prediction ability of ABiLSTM. Compared with several other non-optimized or traditional models, the predicted situation values and the real situation values in this paper have a better fit and smaller error values.

The first section of this paper introduces the background, status, motivation and innovation of the study. Section 2 focuses on related works, BiLSTM, Attention and IPSO. Section 3 includes the framework of IPSO-ABiLSTM model, the method of calculating the situation values and the process of situation prediction. Section 4 includes the selection of experimental parameters, the analysis and comparison of the results. Section 5 provides conclusions and suggestions for future work.

2. THEORETICAL IPSO-ABILSTM MODEL

This section introduces the theory related to the IPSO-ABiLSTM based network security situation prediction model. The fusion of BiLSTM and Attention has good application prospects in the time-series prediction problem, and then the IPSO algorithm is used to optimize its network parameters, and the final model prediction capability is effectively improved to achieve the prediction of network situation changes in the next period.

2.1 BiLSTM

In 1997, Schmidhuber *et al.* proposed a variant cyclic neural network Long Short-Term Memory network [10], which introduced a gating mechanism to simply and effectively solve the gradient explosion or disappearance problem of traditional recurrent neural networks by controlling the information transfer between each cell through the gating mechanism. The three "gates" are Forget Gate, Input Gate and Output Gate, denoted by f_t, i_t, and o_t, and their internal structure is shown in Fig. 1.

In Fig. 1, x_t represents the input information at the current time, h_{t-1} and h_t represents the cell output value at the previous time and the current time, c_{t-1} and c_t represents the memory unit at the previous time and the current time respectively, σ representing the sigmoid activation function, and tanh representing the tangent function.

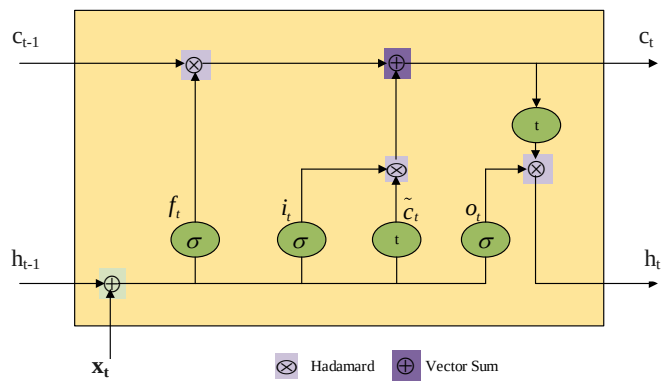

Fig. 1. LSTM basic structure.

The input gate is used to control the extent to which our cells need to store information at the current moment.

$$c_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

The forget gate is used to control how much information we discard from the memory unit at the previous moment.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{3}$$

Through the input gate and output gate, the new cell memory unit value at the current moment can be calculated.

$$c_t = f_t \otimes c_{t-1} + i_t \otimes c_t \tag{4}$$

The output gate is used to control how much information the cell memory unit value has at the current moment as the cell output value.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \otimes \tanh(c_t) \tag{6}$$

In Eqs. (1)-(6), W and b represent the weight matrix and the bias term respectively, represent the sigmoid activation function, tanh represent the tangent function, and σ represent the matrix Hadamard product.

As shown in Fig. 1, x_t represents the input information at the current moment, h_{t-1} and h_t represent the cell output values at the previous moment and the current moment, respectively, c_{t-1} and c_t represents the memory cell at the previous moment and the current moment, σ represents the sigmoid activation function and t represents the tangent function.

LSTM is a one-way extraction of sequence information, but for the problem of network security situation prediction, the current situation of the network is not only related to the situation at the previous moment, but may also be related to the situation in the future. In order to improve the prediction effect, we introduce a bidirectional long short-term memory network (BiLSTM) to predict the network security situation. BiLSTM is composed of two LSTM layers superimposed forward and backward [11], and its structure is shown in Fig. 2.

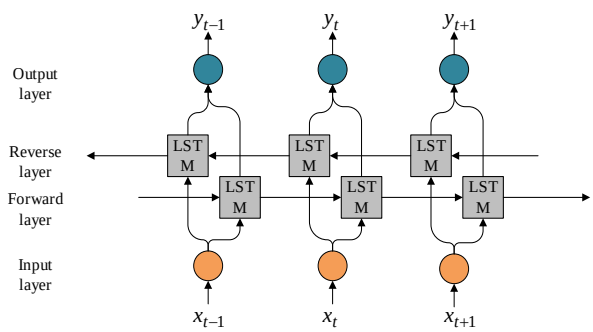

Fig. 2. BiLSTM basic structure.

The Forward layer in the above figure performs LSTM calculation in the positive order of moments and saves the result, while the Reverse layer performs LSTM calculation in the reverse order of moments and saves the result, and the calculation process of Forward and Reverse is shown below.

$$h_t^{\rightarrow} = LSTM^{\rightarrow}(h_{t-1}, x_t) \tag{7}$$

$$h_t^{\leftarrow} = LSTM^{\leftarrow}(h_{t+1}, x_t) \tag{8}$$

where $LSTM^{\rightarrow}$ and $LSTM^{\leftarrow}$ represent LSTM calculations, and $h_t^{\rightarrow}$ and $h_t^{\leftarrow}$ represent the results of forward and backward calculations.

The output of BiLSTM is obtained by summing the results of both calculations, and the summation process is shown below.

$$y_t = w^{\rightarrow} h_t^{\rightarrow} + w^{\leftarrow} h_t^{\leftarrow} + b \tag{9}$$

where w and b represent the corresponding weights and biases, and y_t represents the output value of the BiLSTM cell.

2.2 BiLSTM Fusion Attention

BiLSTM can already achieve good results in extracting sequence information, but in real network conditions, the importance of different features is also very different. BiLSTM alone cannot identify the importance of features in the sequence.

The Attention mechanism is inspired by the working mechanism of our human brain. In the process of cognition of the things around us, people always give priority to what they want to see, and ignore some things that they don't need. The Attention mechanism has been applied in many fields. Attention mechanism in the image field, computer vision, and natural language field, and achieved good results [12-14]. The network structure of the ABiLSTM model in this paper is shown in Fig. 3.

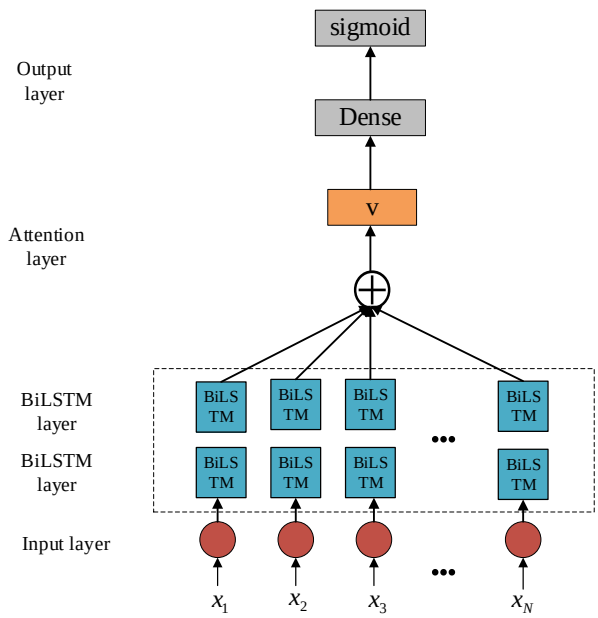

Fig. 3. ABiLSTM model network structure.

In Fig. 3, the ABiLSTM model consists of an input layer, an implicit layer composed of two stacked BiLSTM layers, an Attention layer and an output layer, and the stacked BiLSTM can extract the timing features more effectively. The specific implementation of the ABiLSTM model used in this paper:

(1) Pass the BiLSTM hidden layer output h_t through the nonlinear activation function to obtain the correlation coefficient u_t output at other times, as shown in Eq. (10).

$$u_t = \tanh(W_w h_t + b) \tag{10}$$

Among them, W_w represents the weight matrix, and b represents the offset.

(2) Assign importance weights to each output of the hidden layer to obtain the weight coefficient a_t, as shown in Eq. (11).

$$a_t = \frac{\exp(u_t)}{\sum_{j=1}^{t} u_j} \tag{11}$$

(3) Calculate the product of the weight coefficient at each moment and the output of the hidden layer to obtain the output vector *v* of the Attention layer, as shown in Eq. (12).

$$v = \sum a_t h_t \tag{12}$$

(4) Finally, the prediction result is obtained through the sigmoid function.

For the ABiLSTM network, the choice of parameters in its structure is very important to the effect of the model, such as the number of hidden layers, weights, number of hidden layer units, learning rate and other parameters. Many researchers based on experience or trial and error Determining these parameters makes the robustness and accuracy of the model unreliable. Therefore, this paper selects a particle swarm algorithm that is simple in principle, low in complexity, fast in convergence, and suitable for dealing with realvalued problems to optimize the structural parameters of the ABiLSTM network.

2.3 IPSO

The particle swarm algorithm is a bionic swarm optimization algorithm proposed by Dr. Eberhart and Dr. Kennedy [16] in 1995. The algorithm is derived from the study of the regular predation behavior of bird swarms. The basic idea of the particle swarm algorithm is to treat each solution of the problem as a D-dimensional massless particle, and each particle has a fitness value determined by the fitness function. In the search space, each particle is optimized according to the individual. The position and the global optimal position are used to update its own speed and position. Through iterative search, the optimal position of the entire particle swarm is obtained [17].

In each iteration, the particles in the group determine their search direction and distance through speed. The basic particle group speed and position update formula are as follows:

$$V_{id}^{k+1} = wV_{id}^{k} + c_1 r_1 (pbest_{id}^{k} - X_{id}^{k}) + c_2 r_2 (gbest_{gd}^{k} - X_{id}^{k}), \tag{13}$$

$$X_{id}^{k+1} = X_{id}^{k} + V_{id}^{k+1}. \tag{14}$$

Among them, *k* represents the current iteration number, *w* represents the inertia weight factor, that is, the ability of the particle to inherit the speed of the previous iteration, c_1 and c_2 represents the acceleration factor, which is used to adjust the influence of the individual optimal solution and the global optimal solution on the speed of each iteration. The sum is a random number between [0, 1]. V_{id}^{k} and X_{id}^{k} respectively represent the speed and position of the *d*-dimensional space of the *i*th particle in the kth iteration, $pbest_{id}^{k}$ and $gbest_{id}^{k}$ respectively represent the individual optimal position and the global optimal position of the *d*th dimensional space of the *i*th particle in the *k*th iteration.

In the particle swarm algorithm, the inertia weight factor and acceleration factor are very important to the efficiency and results of the PSO algorithm. When the inertia weight factor and acceleration factor are large, the global optimization ability is better, and if the inertia weight factor and acceleration factor are small, the local optimization. The optimi-

zation ability is better. Because the inertia weight factor and acceleration factor in the traditional particle swarm algorithm are fixed, this limits the local and global optimization capabilities of the algorithm, and it is easy to cause the algorithm to fall into a local minimum. Aiming at the limitations of the algorithm, this paper improves the inertia weight factor and acceleration factor to make the speed change from linear to nonlinear.

The improvements to the inertia weight factor w are as follows,

$$w = -\pi * arcsin(0.01 * (t - \max\_iter)). \tag{15}$$

The values of acceleration factors c_1 and c_2 are as follows,

$$c_1 = c_{1\max} - (c_{1\max} - c_{1\min}) * ((t)/(\max\_iter)) ** 2, \tag{16}$$

$$c_2 = c_{2\max} - (c_{2\max} - c_{2\min}) * ((t)/(\max\_iter)) ** 2. \tag{17}$$

Where t represents the current number of iterations, max_*iter* represents the maximum number of iterations, $c_{1\max}$ and $c_{1\min}$ represents the maximum and minimum values of respectively, $c_{2\max}$ and $c_{2\min}$ represents the maximum and minimum values of respecttively.

3. NETWORK SECURITY SITUATION PREDICTION

The network security situation value is a quantitative representation of the network security situation in a certain range of values. This part firstly describes the framework of IPSO-ABiLSTM model, and then establishes a network security situation indicator system based on the impact of attacks. The indicator system fully considers the intrinsic correlation of the main influencing elements in the network, and because network attacks are increasingly complex, diversified and frequent, different types of attacks have different impacts on the whole network, so it is necessary to improve as much as possible the detection accuracy in order to perceive the network condition more accurately. According to the network security situation index system, we quantify the situation values that can visually represent the network security situation, and finally give the process of implementing the situation prediction.

3.1 IPSO-ABiLSTM Model Framework

The specific framework of ABiLSTM consists of four layers, each of which has the following roles.

(1) BiLSTM layer: It contains two BiLSTM layers in total, and by stacking BiLSTM layers can make full use of their combined before-and-after capabilities to enhance the model learning.
(2) Attention layer: extracts the key information related to the output from the results of BiLSTM layer.
(3) Dense layer: set the last layer as a fully connected layer, transform the dimensionality of the output, and get the prediction result by the activation function.

The ABiLSTM model framework used in this paper is shown in Fig. 4.

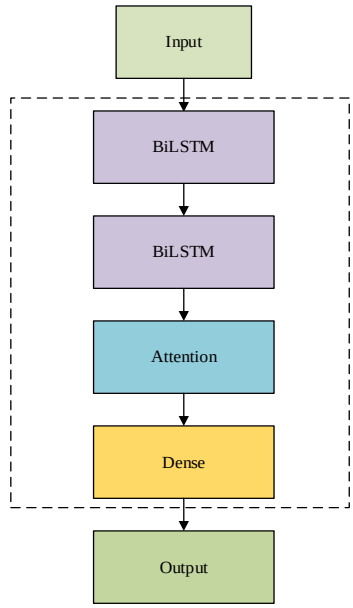

Fig. 4. IPSO-ABiLSTM framework.

3.2 Network Security Situation Value

The attack traffic characteristics and methods collected by the commonly used KDD cup99 and NSL-KDD [18, 19] datasets can no longer represent the network conditions of the current era. The novel UNSW-NB15 dataset [20, 21] does not contain the situation value in the UNSW-NB15 dataset, so we adopt the above calculation method to generate the situation value representing the security degree of the network. There is no real situation value in the UNSW-NB15 dataset, so this paper constructs the situation indicator system from the perspective of attack impact and generates the real situation value of the UNSW-NB15 dataset to analyze the network security situation. The situation factors include the number of attacks and the threat of attacks. The attack number factor is the number of attack samples in a period of time, denoted by N. Attack threat factors are different types of attacks on the network security threat value represented by X_i. The threat factors of each attack type are shown in Table 1. The situation value of t time period is

$$SA(t) = f(N, X_i) = \sum_{i=1}^{N} X_i. \tag{18}$$

According to the timing of the attack samples, every 3000 samples are divided into a time period. After the calculation is completed according to Eq. (18), all time periods are mapped to [0, 1]. The final training set is composed of 58 time periods, and the test set consists of 27 time periods. Since the collection of the data set itself is time-sequential, and the situation value generated according to the impact of the attack has a strong representativeness, the true situation value calculation method of the data set used in this paper is feasible.

Table 1. Attack threat factors of UNSW-NB15 dataset.

Category	Threat factor	Category	Threat factor
Normal	1	Generic	6
Analysis	2	Shellcode	7
Reconnaiss	3	Worms	8
Fuzzers	4	Exploits	9
Dos	5	Backdoor	10

3.3 Network Security Situation Prediction Process

The process of the IPSO-ABiLSTM prediction model is shown in Fig. 5. The specific steps of IPSO-ABiLSTM model are as follows.

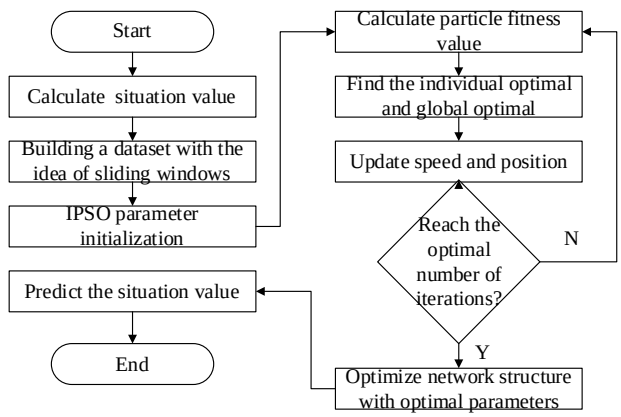

Fig. 5. IPSO-ABiLSTM forecasting process.

(1) Construct training set samples and test set samples according to the size of the sliding window.

In this paper, the rule of using sliding window is to assume that the window size is $m + 1$, then the first m periods are the input samples, and the m+1th period is the output result, sliding backward in order. The structure of the divided dataset is listed in Table 2.

Table 2. Dataset structure for prediction.

Serial number	Input sample	Output sample
1	$(x_1, x_2, \ldots, x_m)$	x_{m+1}
2	$(x_1, x_2, \ldots, x_m)$	x_{m+2}
…	…	…
$n - m$	$(x_{n-m}, x_{n-m+1}, \ldots, x_{n+1})$	x_n

(2) Initialize the relevant parameters in IPSO: the search dimension D, the number of particles pN, the maximum and minimum values of the acceleration factor c_1 and c_2 the maximum number of iterations max_*iter*, the initial position X_i^0 and the initial velocity V_i^0 of the particles, the inertia weight factor w and the learning factor r_1 and r_2 are automatically generated in the iteration.

(3) Set the range of values for each dimension in the particles to be optimized, which are the learning rate in the ABiLSTM model, the number of model iterations, the number of cells in the first hidden layer, the number of cells in the second hidden layer of the LSTM, and the random seed.
(4) Set the fitness function of the particle swarm algorithm, randomly generate the initial positions of the particle swarm, calculate the initial fitness value of each particle, and obtain the individual optimal solution $pbest_{id}^{k}$ and the global optimal solution $gbest_{id}^{k}$ at the beginning.
(5) Calculate the fitness value of each particle, update the individual optimal solution $pbest_{id}^{k}$ and the global optimal solution $gbest_{id}^{k}$, and calculate the velocity of the particle and update the position of the particle according to Eqs. (15)-(17).
(6) If the maximum number of iterations is reached, proceed to Step 7. Otherwise, return to Step 5 and continue iteration.
(7) Assign the obtained optimal parameters to the ABiLSTM model to obtain the state prediction results.

4. SIMULATION ANYLYSIS

This part compares the effects through simulation experiments and with representative models, and also verifies the effectiveness of the network security situation awareness prediction method in this paper. The parameters of the simulation experimental environment in this paper are shown in Table 3.

Table 3. Environment parameter.

Environment	Parameter
CPU	Intel Core
OS	Windows 7 (64-bit)
RAM	8GB
TOOL	Python 3.6.0
IDE	PyCharm 2020.2.3
Framework	Tensorflwo 2.0
Repository	Matplotlib, Sklearn

4.1 Selection of Simulation Parameters

(A) Selection of particle swarm related parameters

The relevant parameters of the basic PSO are as follows: the number of particles N = 10, the dimension of each particle (the number of parameters for optimization) = 5, the acceleration factor c_1 = 2.0, c_2 = 2.0, the learning factor r_1 = 0.8, c_2 = 0.3, Inertia weight w = 0.8, max_*iter* = 50.

Relevant parameters of IPSO proposed in this paper are as follows: the number of particles N = 10, the dimension of each particle (number of parameters for optimization) = 5, the acceleration factor $c_{1\max}$ = 2.1, $c_{1\min}$ = 0.8, $c_{2\max}$ = 2.1, $c_{2\min}$ = 0.8 The learning factor r_1 and r_2 is a random number between [0, 1), inertia weight and acceleration factor c_1 and c_2 are automatically generated in each iteration according to Eqs. (9)-(11), and the maximum number of max_*iter* = 50.

To prevent the particles from searching aimlessly in the search space, which affects the convergence speed, bounds are set on the search parameters, where the learning rate is taken between [0.001, 0.1], the number of BiLSTM training sessions is between [100, 500], and the number of cells in the two BiLSTM hidden layers is between [1, 100] The random seeds are taken between [1, 42].

(B) Model optimal parameters

First, when the window size is 2, the optimal parameters of BiLSTM are selected as follows: iteration number is 155, learning rate is 0.00778224879, the number of hidden layer units in the first layer is 18, the number of hidden layer units in the second layer is 82, and the number of random seeds is 4. In addition, for a window size of 3, the optimal parameters of BiLSTM are selected as follows: iteration number is 299, learning rate is 0.00890800961, the number of hidden layer units in the first layer is 76, the number of hidden layer units in the second layer is 93, and the number of random seeds is 8. Third, at a window size of 4, the optimal parameters of BiLSTM are selected as follows: iteration number is 132, learning rate is 0.00687720314, the number of hidden layer units in the first layer is 34, the number of hidden layer units in the second layer is 85, and the number of random seeds is 6.

4.2 Experimental Analysis

In order to verify the predictive ability of each model used in this paper, two typical regression evaluation indicators are selected to evaluate each model, which are divided into Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE), and the coefficient of determination of goodness of fit (R^2). The calculation formulas of the two evaluation indicators are as follows,

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\%, \tag{16}$$

$$SMAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{y}_i - y_i\right|}{(\left|\hat{y}_i\right| + \left|y_i\right|)/2} \times 100\%, \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}. \tag{18}$$

Among them, y_i represents the true situation value, $\hat{y}$ represents the predicted situation value, N represents the number of samples, and $\bar{y}$ represents the average value of the true situation value. The smaller the average percentage error, the better the model performance. but it has a drawback that if the true value is too small, the small error in the predicted value may lead to a large result, in order to compensate for this drawback, we added the SMAPE indicator. The goodness of fit determination coefficient is between [0, 1], and the closer to 1, the better the model fit.

(A) PSO and IPSO performance comparison analysis

The performance comparison of PSO and IPSO is necessary, the number of training iterations, learning rate, number of BiLSTM neurons in the first layer, number of BiLSTM neurons in the second layer, and random seeds in the BiLSTM are used as the target optimization parameters in the particle swarm, where the fitness function value is the loss function loss value of the BiLSTM. To verify that the IPSO algorithm in this paper has been improved compared to PSO, the optimization search process of IPSO-BiLSTM and PSO-LSTM models were compared at a window size of 2. The comparison results are shown in Fig. 6.

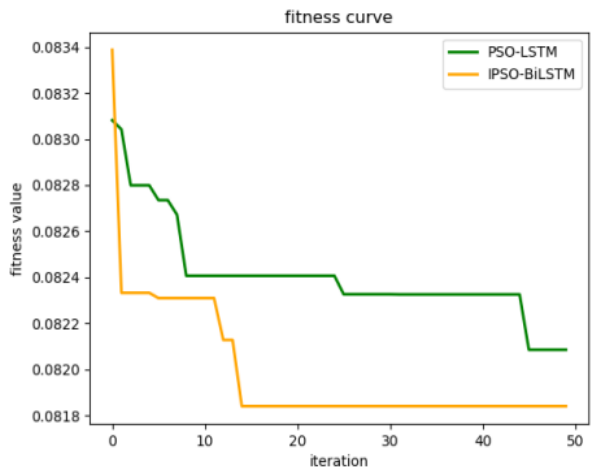

Fig. 6. Comparison of the change of fitness function.

Observing the comparison of the fitness values in Fig. 6 proves that IPSO algorithm has better global search ability with larger inertia weight and acceleration factor in the early stage, and its local search ability is better in the later stage as the acceleration factor and inertia weight decrease, IPSO algorithm balances the global search ability and local search ability, and can find the optimal solution faster, and its effect is better than PSO algorithm.

(B) Situation prediction performance analysis

In order to verify the prediction effect of each model, this paper conducts a comparative experiment on the prediction performance of each model when the window size is 2, 3, and the comparative experiment results are shown in Figs. 7-9. In this paper, the meaning of a window of 3 is to use the situation value of the first two time periods to pre-dict the situation value of the next time period.

It can be seen from Figs. 7-9 that when the window is 2, the IPSO-ABiLSTM proposed in this paper almost completely fits the real situation value, while the other three models all have a certain degree of fitting deviation. The window size is 3 and 4. In the first 3 time periods, the IPSO-ABiLSTM prediction effect proposed in this paper is not ideal, but it is almost completely fitted in the subsequent time periods. Overall, the fit of IPSO-ABiLSTM is still better than the other 3 models.

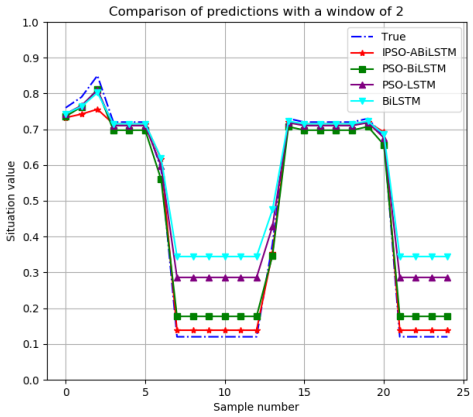

Fig. 7. Comparison results of window value 2.

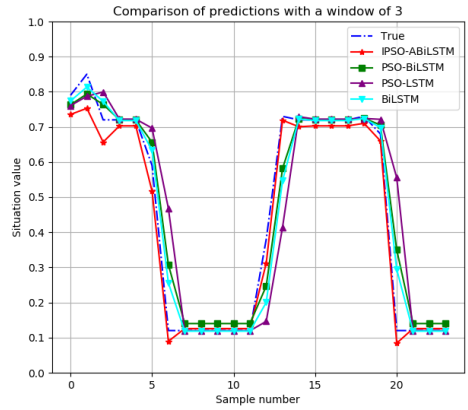

Fig. 8. Comparison results of window value 3.

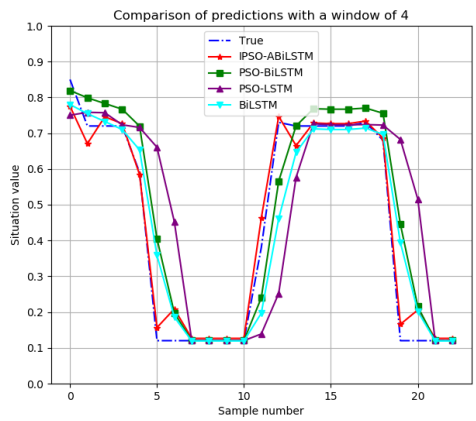

Fig. 9. Comparison results of window value 4.

Table 4. Assessment outcomes of evaluation indexs of each model.

Size	Index	IPSO-ABiLSTM	PSO-BiLSTM	PSO-LSTM	BiLSTM
2	MAPE	1.4038	1.4261	1.5621	1.6316
	SMAPE	0.1709	0.1761	0.1824	0.1843
	R^2	0.9922	0.9807	0.8719	0.7645
3	MAPE	1.3712	1.4590	1.4680	1.4245
	SMAPE	0.1732	0.1838	0.1839	0.1845
	R^2	0.9849	0.9327	0.7678	0.9425
4	MAPE	1.4541	1.5728	1.5617	1.4328
	SMAPE	0.1749	0.1843	0.1862	0.1849
	R^2	0.9809	0.8528	0.3910	0.8666

The evaluation indicators of each model in different windows are shown in Table 4. Our analysis of Table 4 can draw the following conclusions:

(1) When the window value is 2, the MAPE value of IPSO-ABiLSTM model proposed in this paper is 0.0223, 0.1583 and 0.2278 lower than PSO-BiLSTM, PSO-LSTM and BiLSTM, respectively, and the SMAPE value is 0.0052, 0.0115 and 0.0134 lower than

the other three, respectively, and the fit coefficient. The performance of this model is better than the other three models when the window value is 2.

(2) When the window value is 3, the MAPE value of IPSO-ABiLSTM model proposed in this paper is 0.0878, 0.0968 and 0.0533 lower than PSO-BiLSTM, PSO-LSTM and BiLSTM, respectively, and the SMAPE value is 0.0106, 0.0107 and 0.0113 lower than the other three models, respectively, and the fit. The performance of this model is better than the other three models when the window value is 3.

(3) When the window value is 4, the MAPE value of IPSO-ABiLSTM model proposed in this paper is 0.1187 and 0.1076 lower than PSO-BiLSTM and PSO-LSTM respectively, and the SMAPE values are 0.0094, 0.0113 and 0.0100 lower than the other three models, and the fit coefficient R2 is higher than the other three models respectively. Combining the two indicators, the performance of this model is better than the other three models when the window value is 4; (1) When the window value is 2, the MAPE value of the IPSO-ABiLSTM model proposed in this paper is 0.0223, 0.1583 and 0.2278 lower than that of PSO-BiLSTM, PSO-LSTM and BiLSTM, respectively, and the fitting coefficient R^2 is compared with the other three models. 0.0115, 0.1203 and 0.2277 higher respectively. The performance of the model in this paper is better than the other three models when the window value is 2.

(4) For prediction problems, different window sizes often have an impact on the prediction results. This paper also conducts comparative experiments on more window values, and finally obtains that, as far as the method in this paper is concerned, when the window value is smaller, the prediction effect of each model is often the better. Through the lateral analysis of (1)-(3), when the sliding window size is the same, the IPSO-ABiLSTM proposed in this paper has a better prediction effect than PSO-LSTM, PSO-BiLSTM and BiLSTM. At the same time, the fitting coefficient R^2 of each model is compared when the window value is 2, 3 and 4. As shown in Fig. 10, the model in this paper can achieve the best fitting effect when the window value is 2, and the fitting coefficient can be It reaches 0.9922, which is almost a complete fit, which proves the effectiveness of the model proposed in this paper in the problem of network security situation prediction.

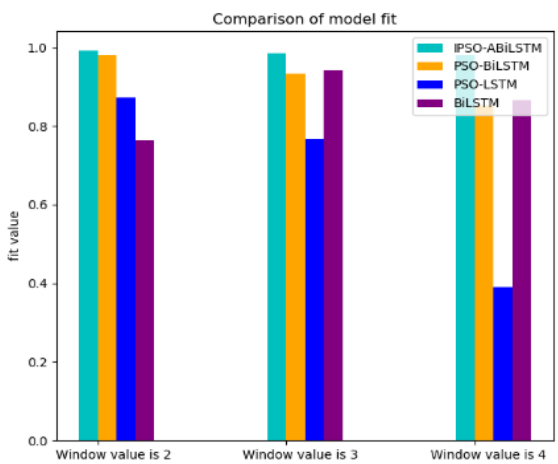

Fig. 10. Comparison of model fitting degree.

5. CONCLUSION

To address the security issues facing wireless mobile networks, this paper proposed a network security situation prediction model based on IPSO-ABiLSTM. Since there is no real situation value in the used dataset, this paper adopted a method to calculate the situation value from the attack-based impact for situation prediction. In the model construction, the PSO algorithm is improved by introducing nonlinearly varying inertia weights and acceleration factors to address the shortcomings of slow convergence of the PSO algorithm and its tendency to fall into local minima. Meanwhile, in order to learn the correlation between sequences more deeply, the Attention fusion BiLSTM network is introduced for situation prediction, and the improved PSO algorithm is added to optimize the ABiLSTM to improve the prediction ability of the model. The experimental results show that IPSO-ABiLSTM has obvious advantages over other models mentioned in the paper in network security situation prediction.

The next step of the job: The experiments made in this paper are simulation experiments, and the theoretical and experimental analyses show that they are convincing. In the next period of time, we will try to improve the existing experimental conditions, build a real wireless mobile networks environment to verify the prediction effect of the proposed method and model.

ACKNOWLEDGMENTS

We would like to thank: National Natural Science Foundation of China under Grant (No. 61672206), Central Government Guides Local Science and Technology Development Fund Project Under Grant (No. 216Z0701G), the Key Research and Development Program of Hebei Under Grant (No. 20310701D).

REFERENCES

1. G. Kambourakis, F. G. Marmol, and G. J. Wang, "Security and privacy in wireless and mobile networks," *Future Internet*, Vol. 10, 2018, p. 18.
2. M. R. Endsley, "Design and evaluation for situation awareness enhancement," in *Proceedings of the Human Factors Society Annual Meeting*, Vol. 32, 1988, pp. 97-101.
3. Z. C. Wen, Z. G. Chen, and J. Tang, "Network security situation prediction based on time series analysis," *Journal of South China University of Technology* (*Natural Science Edition*), Vol. 44, 2016, pp. 137-143+150.
4. R. C. Zhang, Y. C. Zhang, J. Liu, and Y. D. Fan, "Application of improved convolutional neural network network security situation prediction method," *Computer Engineering and Applications*, Vol. 55, 2019, pp. 86-93.
5. C. Luo, Z. Tan, G. Min, J. Gan, W. Shi, and Z. Tian, "A novel web attack detection system for Internet of Things via ensemble classification," *IEEE Transactions on Industrial Informatics*, Vol. 17, 2021, pp. 5810-5818.
6. Y. X. Chen, X. C. Yin, and R. Tan, "A network security situation prediction model based on GSA-SVM," *Journal of Air Force Engineering University* (*Natural Science Edition*), Vol. 19, 2018, pp. 78-83.

7. H. B. Wang, D. M. Zhao, and X. X. Li, "Research on network security situation assessment and forecasting technology," *Journal of Web Engineering*, Vol. 19, 2020, pp. 1239-1266.
8. Y. Q. Tang, C. H. Li, and Y. F. Song, "Network security situation prediction based on im-proved particle swarm optimization and extreme learning machine," *Computer Appli-cations*, Vol. 41, 2021, pp. 768-773.
9. K. Z. Liu, J. P. Gou, Z Luo, K Wang, X. W. Xu, and Y. J. Zhao, "Prediction method of dissolved gas concentration in transformer oil based on particle swarm optimization-long short-term memory network model," *Power Grid Technology*, Vol. 44, 2020, pp. 2778-2785.
10. X. L. Ma, Z. M. Tao, Y. H. Wang, H. Y. Yu, and Y. P. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C Emerging Technologies*, Vol. 54, 2015, pp. 187-197.
11. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, Vol. 9, 1997, pp. 1735-1780.
12. B. Ameen, D. Hussain, J. Ali, and F. Hassan, "Fall event detection using the mean absolute deviated local ternary patterns and BiLSTM," *Applied Acoustics*, Vol. 192, 2022.
13. K. Xu, *et al.*, "Show,attend and tell:Neural image caption generation with visual attention," in *Proceedings of the 32nd International Conference on Machine Learning*, Vol. abs, 2015, pp. 2048-2057.
14. M. Carrasco and A. Barbot, "Spatial attention alters visual appearance," *Current Opinion in Psychology*, Vol. 29, 2019, pp. 56-64.
15. L. Shi, Y. Wang, Y. Cheng, and R. B. Wei, "A review of research on attention mechanism in natural language processing," *Data Analysis and Knowledge Discovery*, Vol. 4, 2020, pp. 1-14.
16. J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of IEEE International Conference on Neural Networks*, 1995, pp. 1942-1948.
17. D. Zhao and J. Liu, "Study on network security situation awareness based on particle swarm optimization algorithm," *Computers & Industrial Engineering*, Vol. 125, 2018, pp. 764-775.
18. M. Tavallaee, E. Bagheri, W. Lu, and A. G. Ali, "A detailed analysis of the KDD CUP99 data set," in *Proceedings of IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 2009, pp. 53-58.
19. A. Kumarshrivas and A. K. Dewangan, "An ensemble model for classification of attacks with feature selection based on KDD99 and NSL-KDD data set," *International Journal of Computer Application*, Vol. 99, 2014, pp. 8-13.
20. M. S. Al-Daweri, K. A. Zainol Ariffin, S. Abdullah, and M. F. E. Senan, "An analysis of the KDD99 and UNSW-NB15 datasets for the intrusion detection system," *Symmetry*, Vol. 12, 2020, p. 1666.
21. S. Choudhary and N. Kesswani, "Analysis of KDD-cup'99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT," *Procedia Computer Science*, Vol. 167, 2020, pp. 1561-1573.

Ya-Xing Wu (吴亚星) is a teacher at the Hebei Institute of Mechanical and Electrical Technology. His main research interests include network and information security technology.

Dong-Mei Zhao (赵冬梅) received her Ph.D. degree from Xidian University. She is a Professor at Hebei Normal University and a senior member of the Chinese Computer Federation. Her main research interests include network information security and computer application.