IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 13, 2022, accepted June 4, 2022, date of publication June 13, 2022, date of current version June 20, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3182376

Wireless Channel State Prediction Method Based on Improved Adaptive and Parameter-Free Recurrent Neural Structure

QINGLI LIU, NA CAO, MENGQIAN LI, ZHENYA ZHANG, AND GUOQIANG YANG
Communication and Network Laboratory, Dalian University, Dalian 116622, China

Corresponding author: Qingli Liu (lql0808@sina.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61931004.

ABSTRACT Aiming at the problem of large error of channel state prediction caused by channel time-varying characteristics in wireless communication system, propose a wireless channel state prediction method based on improved adaptive and parameter-free recurrent neural structure(APF-RNS). The method though Aquila Optimizer for find the number of hidden layer units and the optimal value of learning rate of the neural network, using the optimal parameters to construct adaptive without recursive neural network. In this way, the convergence speed and fitting effect of the objective function of the neural network can be improved, and the problems of large prediction error and poor generalization ability of the neural network in the prediction process can be avoided, thereby improving the prediction accuracy of the channel state information. The simulation results show that compared with Genetic Algorithm, Particle Swarm Optimization and Sparrow Search Algorithm, the improved APF-RNS has better performance in the optimization ability and convergence speed. Meanwhile, it also has a significant improvement in prediction accuracy compared to the APF-RNS.

INDEX TERMS Adaptive and parameter-free recurrent neural structure, channel state information prediction, aquila optimizer, wireless communication.

I. INTRODUCTION

In the propagation environment of wireless communication system, there are multipath propagation phenomena such as reflection, diffraction and scattering of signals, and doppler extension is inevitably generated by the relative movement of transmitter and receiver, which makes the wireless channel present frequency selectivity and time-varying characteristics. The above factors make it difficult to obtain accurate Channel State Information (CSI). However, accurate acquisition of CSI is crucial for high spectral efficiency at the transmitter and receiver. To solve the above problems, the initial method periodically uses known pilot symbols to estimate the channel in real time, but incurs pilot overhead. In addition, after channel estimation at the receiver, in order to enable the transmitter to obtain channel status information, CSI feedback is required in Frequency Division Duplex (FDD). CSI feedback consumes a large amount of reverse link resources. More importantly, introduce a feedback delay or pilots are sent in the opposite direction to estimate the CSI of the reverse link [1]. Channel reciprocity needs to be assumed in Time Division Duplex (TDD). Due to time delays in channel estimation, signal processing and feedback, the CSI available at the transmitter may be outdated before it is actually used [2]. Especially in high mobility environments, the channel conditions may have changed after the feedback delay. In this case, accurate CSI is more difficult to obtain.

Aiming at the problem that it is difficult to accurately obtain channel state information in wireless communication systems, the modulation technology of filter bank multi-carrier is proposed in literature [3], which improves the utilization rate of spectrum, but passively compensates performance loss at the cost of scarce wireless resources. In contrast, the channel prediction method can directly improve the accuracy of CSI without consuming additional wireless resources, so it has received extensive attention from researchers. Literature [4] proposes a method which based on sinusoidal modeling forecasting. This method is used to overcome the computational complexity with time-varying parameters by using a spectral estimation method to

The associate editor coordinating the review of this manuscript and approving it for publication was Lin Lin.

determine the relevant parameters of each sinusoid. However, due to the cumbersome estimation process, the estimated parameters will soon expire with the fluctuation of the fading channel, so iterative re-estimation is required, and the computational complexity is high. In literature [5], an autoregressive (AR) Model was proposed to predict AR Model parameters based on CSI historical observations by using maximum entropy method, so as to predict channel. However, in AR channel prediction, large-scale matrix inversion is complex and computationally intensive, and the error generated by coefficient estimation will reduce the accuracy of channel prediction.

Most of the propagation channels in the real world are unstable, it do not conform to the above hypothesized model, and the errors generated by coefficient estimation bring great challenges to channel prediction.With the widespread application of massive Multi-Input Multi-Output (MIMO) systems, the number of antennas increases, which makes accurate modeling of large-matrix channels more and more difficult. Deep learning uses a large amount of training data to establish the underlying relationship between input and output.This makes it possible to generate a mathematical model that is not easily described by mathematical formulas but is very efficient. For example, literature [6] has proposed a channel prediction model based on Recurrent Neural Network (RNN), which uses the entire vectorized channel matrix as input to predict channel state through Neural Network, proving the feasibility of deep learning. However, since RNN has only short-term memory, there is the problem of gradient disappearance, which reduces the speed of prediction. Literature [7] proposed a satellite channel predictor composed of Long Short-Term Memory network (LSTM) to eliminate the negative impact of outdated CSI on low Earth orbit (LEO) satellite system. LSTM network combines short-term memory with long-term memory through sophisticated gate control. It solves the problem of gradient disappearing to some extent, but has the disadvantages of low efficiency and not being able to predict in real time. Literature [10] proposed a small and efficient Adaptive and Parameter-Free Recurrent Neural Structure (APF-RNS), which compared the prediction results of APF-RNS with the AR model. The results showed, APF-RNS has greatly improved the prediction accuracy. Although the structure has achieved good results, it is easy to fall into local optimization and over-fitting problems. In view of the above problems, this paper proposes an improved adaptive recursive neural network without parameters, which solves the problems of large prediction error and poor generalization ability of neural network in the prediction process, and enables it to have more accurate and fast prediction results in the real complex wireless channel prediction. Aquila Optimizer (AO) [9] is first used to generate a group of candidate solutions within a certain range, each corresponding to APF-RNS learning rate and the number of hidden layer units, and use the training set to select the position with the minimum fitness function value in the candidate set. Then, the position of each individual in the eagle population is transformed according to the fitness function, and the optimal parameters are obtained to complete the optimization process. Finally, the channel state information is predicted by constructing APF-RNS network model using optimal parameters.

The main contributions of this paper are as follows: (1) An improved APF-RNS wireless channel state prediction method is proposed to solve the problem of channel state prediction error caused by channel time-varying characteristics in wireless communication system. (2) Use AO algorithm to obtain the optimal parameters of neural network, and use the optimal parameters to construct APF-RNS network, so as to solve the problems of large prediction error and poor generalization ability of neural network in the prediction process. (3) The on-line training method used in channel prediction can reduce the pilot cost required in the communication link. (4) The method does not need to know channel knowledge in the prediction process, such as long-term statistics or channel parameters, so it can be extended to any propagation environment.

This paper is divided into the following parts: Section II introduces the wireless communication system model and APF-RNS network structure. In Section III, a prediction method of channel state information based on AO improved APF-RNS is proposed. In the IV section, the simulation analysis is carried out, and compared with GA algorithm, PSO algorithm, SSA algorithm, verify the proposed algorithm in the optimization ability and convergence speed superiority. Then, the feasibility and accuracy of the proposed method are verified by multiple data sets. Finally, Section V is comprised of the conclusion.

II. SYSTEM MODEL

A. PROBLEM FORMULATION

For a Single-Input Single-Output (SISO) system, the relationship between the original signal and the corresponding received signal is shown in Formula (1).

$$y[t] = h[t]\,x[t] + z[t] \tag{1}$$

where $x[t]$ and $y[t]$ respectively represent the signal sent at the transmitter side and the corresponding received signal at the receiver side, $h[t]$ is the complex channel state information, and $z[t]$ is the additive White Gaussian noise at time t.

MIMO systems are similar to single-input single-output systems. A typical MIMO system model is shown in Fig.1. As can be seen from the figure, this system is equipped with N_t transmitting antennas at the transmitting end and N_r receiving antennas at the receiving end. Suppose the transmitting signal is $x[t] = \left[x_1[t], x_2[t], \cdots, x_{N_t}[t]\right]^T$, and the receiving signal is $y[t] = \left[y_1[t], y_2[t], \cdots, y_{N_r}[t]\right]^T$, the channel matrix $h[t]$ is shown in Formula (2).

$$h[t] = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N_t} \\ h_{21} & h_{22} & \cdots & h_{2N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r1} & h_{N_r2} & \cdots & h_{N_rN_t} \end{bmatrix}_{N_r \times N_t} \tag{2}$$

where $h[t]_{ij}$ $(i = 1, 2, \ldots, N_t; j = 1, 2, \ldots, N_r)$ represents the channel impact response generated by the combination of the i^{th} root transmitting antenna and the j^{th} root receiving antenna. The signal transmission requires a medium, and the channel is this medium in the communication system.

In order to adapt to the input of neural network, the channel matrix $h[t]$ vector needs to be transformed into the vector of $1 \times N_r N_t$, as shown in Formula (3).

$$h[t] = [h_{11}[t], h_{12}[t], \ldots, h_{NrNt}[t]] \tag{3}$$

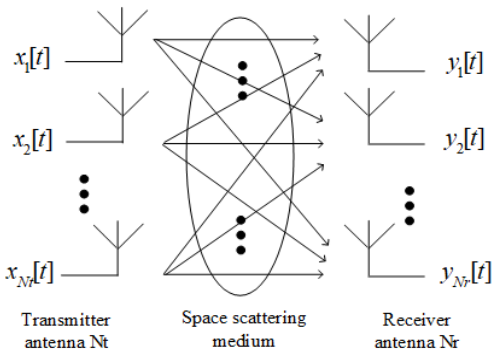

FIGURE 1. MIMO system model diagram.

In order to obtain the channel state information, sent the known pilot symbol $p[t]$, and calculate the measured value of CSI from the received signal, as shown in Formula (4).

$$\hat{h}[t] = \frac{y[t]}{p[t]} = h[t] + \frac{z[t]}{p[t]} \tag{4}$$

In this paper, channel prediction based on pilot symbol and data symbol is applied. In a frequency division duplex system, the receiver estimates the channel state information according to the pilot symbol and uses the interpolated channel state information for data detection. The channel state information estimation at the data location can then be refined from the decoded data. Finally, the transmitter uses the feedback channel state information to predict the channel for transmitter precoding. In a time division duplex system, the receiver can use the reverse link signal for CSI estimation and predict the channel used for forward link precoding.

Assuming that the measured CSI is known at the first P time steps, the system predicts the CSI at the next R time steps. Better performance can be achieved by predicting CSI difference between two adjacent symbols instead of directly predicting CSI. CSI difference between two adjacent symbols is shown in Formula (5).

$$\hat{h}^d[t+1] = \hat{h}[t+1] - \hat{h}[t] \tag{5}$$

Then, the problem becomes to predict the next R differences $\tilde{h}^d[t]$, $P+1 \le t \le P+R$ based on P-1 known differences $\hat{h}^d[t]$, $2 \le t \le P$. The final predicted CSI is calculated by the system as shown in Formula (6).

$$\tilde{h}[t] = \begin{cases} \hat{h}[t-1] + \tilde{h}^d[t] & t = P+1 \\ \tilde{h}[t-1] + \tilde{h}^d[t] & P+2 \le t \le P+R \end{cases} \tag{6}$$

where $\tilde{h}^d[t]$ and $\tilde{h}[t]$ are the CSI difference predicted at t time and the final CSI result predicted.

B. APF-RNS NETWORK STRUCTURE

Channel prediction is to predict the future of several channel state information based on the recent history, where channel sample sequences are correlated. LSTM is used to construct an online APF-RNS network structure for wireless channel prediction, and its structure is shown in Fig.2.

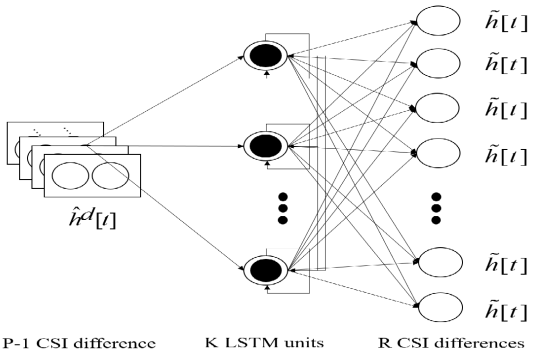

FIGURE 2. APF-RNS network structure diagram.

APF-RNS predicts R unknown CSI differentials in the future according to the known P-1 CSI differentials. Because the training process of neural network has a certain time cost, and the wireless channel is time-varying, the network designed in this paper is simple and efficient, with only one hidden layer. The input layer is P-1 known CSI difference value, namely $\hat{h}^d[t]$, $2 \le t \le P$, and then input into the hidden layer. The hidden layer is K LSTM units. LSTM extracts useful information through gate structure, and there are synapses between K LSTM, forming a recursive network, which enables the neural network to share information in the training stage. Finally, the output layer is a fully connected layer that uses linear activation functions to generate $2 \times R$ real numbers. The output content is converted into R complex numbers after data processing, namely $\tilde{h}^d[t]$, $P+1 \le t \le P+R$, to obtain the final predicted CSI.

Although APF-RNS can make channel prediction according to the nonlinear characteristics of wireless channels, due to the slow convergence rate of objective function and poor generalization ability in the prediction process, the prediction error is large in the case of large CSI variation range. Since the learning rate of APF-RNS determines whether and when the objective function converges to the local minimum, the number of units in the hidden layer will affect the fitting effect and directly determine the generalization ability of the obtained prediction model [10]. APF-RNS directly sets the number of hidden layer units, which greatly increases the complexity of the network structure if the number of hidden layer units is too large, and makes the learning speed of the network slow. If the number of hidden layer units is too small, the network does not have the necessary learning ability and information processing ability. The learning rate of APF-RNS

is the default value when Adam optimizer is used, which makes the training speed of neural network slow and the model performance poor. Therefore, in order to improve the accuracy of APF-RNS prediction, this paper obtained the optimal values of APF-RNS learning rate R and hidden layer element number K according to the unique exploration and development behavior of AO algorithm, and used the optimal values to construct APF-RNS, which reduced the complexity of the network and improved the performance of the network.

III. IMPROVED APF-RNS PREDICTION MODEL BASED ON AO

A. AO IMPROVED APF-RNS ALGORITHM

AO algorithm is a kind of latest swarm intelligence optimization algorithm. Firstly, the decision value X of N is determined by Equation (7).

$$X_{ij} = \text{rand} \times \left(UB_j - LB_j\right) + LB_j, \quad i = 1, 2, \ldots, N; \quad j = 1, 2, \ldots, Dim \tag{7}$$

where *rand* is a random number between [0, 1], UB_j and LB_j are the upper and lower bounds of the number of hidden layer units K and the learning rate R respectively, and *Dim* is 2. A set of candidate solutions X is generated through the above formula.

When $t \leq \left(\frac{2}{3}\right) \times T$, it comes to the exploration phase, there are two approaches in this phase.

The exploration behavior of the first Aquila is shown in Equation (8).

$$X_1(t+1) = X_{\text{best}}(t) \times \left(1 - \frac{t}{T}\right) + \left(X_M(t) - X_{\text{bcst}}(t) \times \text{rand}\right) \tag{8}$$

where $X_1(t+1)$ is the solution of hidden layer unit number and learning rate in iteration t + 1, $X_{\text{best}}(t)$ is the optimal solution in iteration t, $1 - \frac{t}{T}$ is used for search control in the exploration process, t and T respectively represent the current iteration number and maximum iteration number, rand is a random number between [0, 1], and $X_M(t)$ represents the mean value of all solutions in iteration t.

The exploration behavior of the second Aquila is shown in Equation (9).

$$X_2(t+1) = X_{\text{best}}(t) \times \text{Levy}(D) + X_R(t) + (y - x) * \text{rand} \tag{9}$$

where $X_2(t+1)$ is the solution of the $t + 1^{th}$ iteration obtained by the second exploration method, Levy(D) is the Levy flight distribution function, D is the dimensional space, $X_R(t)$ is the solution of random individuals within the range of [1, N] at the t^{th} iteration. y and x are used to present the spiral shape in the search. The Levy flight is calculated using Equation (10).

$$\text{Levy}(D) = s \times \frac{u \times \sigma}{|v|^{\frac{1}{p}}} \tag{10}$$

In Equation (10), s is fixed at 0.01, u and v is a random value between [0,1], σ as shown in Equation (11).

$$\sigma = \left(\frac{\Gamma(1+\beta) \times \sin e\left(\frac{xp}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2\left(\frac{\beta-1}{2}\right)}\right) \tag{11}$$

where β is fixed at 1.5.

When $t > \left(\frac{2}{3}\right) \times T$, it will enter the development stage, and there are two different methods in the development stage.

The development behavior of the first Aquila is shown in Equation (12).

$$X_3(t+1) = (X_{best}(t) - X_M(t)) \times \alpha - \text{rand} + ((UB - LB) \times \text{rand} + LB) \times \delta \tag{12}$$

where $X_3(t+1)$ is the solution calculated by the development method in the iteration of $t + 1$, α and δ are exploration adjustment parameters, fixed at 0.1.

The development behavior of the second Aquila is shown in Equation (13).

$$X_4(t+1) = QF \times X_{best}(t) - (G_1 \times X(t) \times \text{rand}) - G_2 \times \text{Levy}(D) + \text{rand} \times G_1 \tag{13}$$

where $X_4(t+1)$ is the solution calculated by using the second development method in the iteration of $t + 1$. QF is the quality function used to equalize the search strategy. G_1 represents the movement of the Aquila to track the prey during its escape. The value of G_2 decreases from 2 to 0. $X(t)$ is the solution of the t^{th} iteration.

The minimum value $X_{\text{best}}(t)$ of fitness function is obtained after T iterations, which represents the optimal value of the number of hidden layer units and learning rate. The fitness function value $FinessX(t)$ is calculated by mean square error, as shown in Equation (14).

$$FinessX(t) = \frac{1}{N_a} \sum_{t_a=1}^{N_a} (\hat{h}[t_a] - \tilde{h}[t_a])^2 \tag{14}$$

where N_a is the number of training sets, $\hat{h}[t_a]$ is the CSI measured value, and $\tilde{h}[t_a]$ is the CSI predicted value.

The detailed steps of AO to improve APF-RNS algorithm are shown in Algorithm1.

B. AO IMPROVED APF-RNS ALGORITHM FLOW

The variation range of parameters is determined according to historical data and prediction requirements. Then use the AO algorithm to find the optimal number of hidden layer units K and learning rate R within the given range. The flow chart of AO improving APF-RNS is shown in Fig.3. The detailed algorithm steps are as follows:

Step 1: Select an appropriate APF-RNS network structure, and determine the upper and lower bounds of the number of hidden layer units and learning rate according to the historical data.

Algorithm 1 AO Improved ANF-RNS Algorithm

Initialize parameters in the AO algorithm, such as population size N, iteration times T, and upper bounds UB_j and lower bounds LB_j of the problem, etc.
Initialize population X.
while $t < T$ **do**
  Calculate the fitness function of population X.
  The optimal individual was obtained according to the fitness function value.
  Calculate x, y, G_1, G_2, $Levy(D)$, and so on.
  for $i = 1, 2, \ldots, N$ **do**
    Calculate the mean value $X_M(t)$ of all individuals in population X.
    if $t \leq \left(\frac{2}{3}\right) * T$ **then**
      if $rand \leq 0.5$ **then**
        Use Equation (8) to calculate $X(t+1)$.
      else
        Use Equation (9) to calculate $X(t+1)$.
      end if
    else
      if $rand \leq 0.5$ **then**
        Use Equation (12) to calculate $X(t+1)$.
      else
        Use Equation (13) to calculate $X(t+1)$.
      end if
    end if
    if $FinessX(t+1) \leq FinessX(i)$ **then**
      $X(i) = X(t+1)$;
      if $FinessX(t+1) \leq FinessX_{\text{best}}(t)$ **then**
        $X_{\text{best}}(t) = X(t+1)$;
      end if
    end if
  end for
end while
return $X_{\text{best}}(t)$;

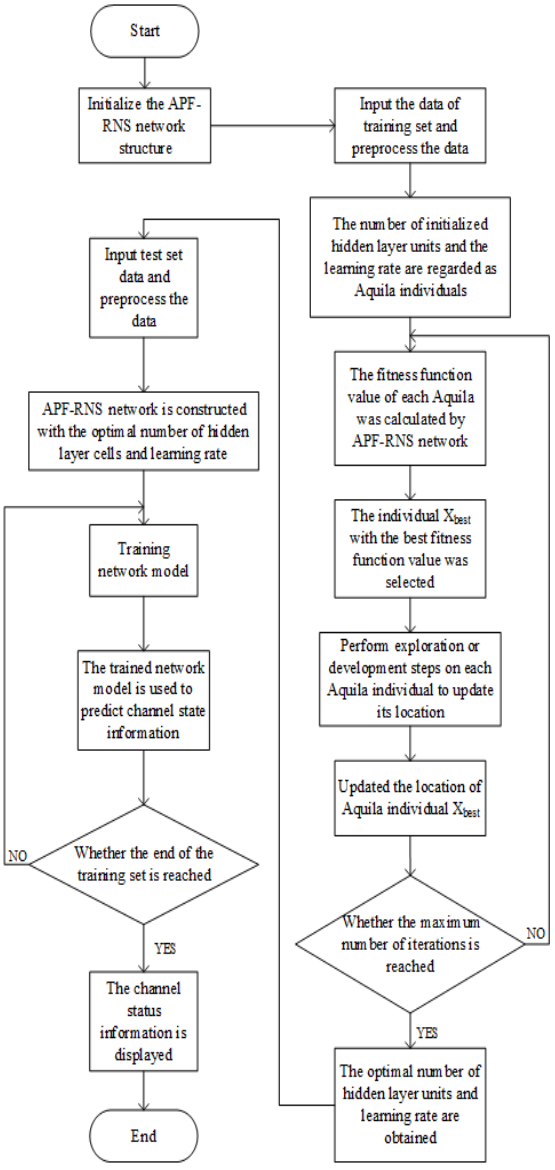

FIGURE 3. Flow chart of APF-RNS algorithm optimized by AO algorithm.

Step 2: Input the training set data and preprocess the data;

Step 3: The number of initialization hidden layer units and learning rate were taken as the Aquila individuals, and set the population N. At the same time, the upper limit T of iteration times and relevant parameters of AO algorithm are determined;

Step 4: Take the mean square error of APF-RNS network training set as the fitness function value of AO algorithm;

Step 5: Calculate the fitness value of each Aquila, and rank the fitness function value, select the Aquila with the highest target value X_{best}.

Step 6: Perform exploration and development steps for each individual to update its location, and update the location of Aquila individual X_{best};

Step 7: Judge whether the AO algorithm reaches the upper limit of iteration times. If so, retain the final X_{best} Aquila individual, that is the optimal number of hidden layer units and learning rate; Otherwise, repeat Step 5 to Step 6.

Step 8: Read and preprocess the test set data, then construct the APF-RNS network with the number of hidden layer units and learning rate corresponding to X_{best} Aquila.

Step 9: Train the network model and use the trained network model to predict the channel state information;

Step 10: Judge whether reach the end of the training set data, if so, output all the predicted channel state information; Otherwise, repeat Step 9.

IV. ANALYSIS OF SIMULATION RESULTS

A. SIMULATION PARAMETER SETTING

The simulation verification of the proposed method uses channel measurements in two wireless environments. The first is the measurement data of the wireless system in

the industrial environment from National Institute of Standards and Technology(NIST) [11], in which the frequency is 2.245GHz, the polarization of the receiver and transmitter antenna is Omni-directional and V Pol, and the antenna gain of the receiver is −3.8dBi. The antenna gain of the transmitter is 2.9dBi, the transmit power is 1.5watts, the attenuation is 50dB, and the sample rate is 80MHz. The second data set adopts the indoor wireless channel measurements at 2.4GHz provided by Mohamed AlHajri and Nazar Ali et al in the Machine Learning library of University of California, Irvine (UCI) [12]. The measurement system for the dataset consists of the ZVB14 Vector Network Analyzer (VNA), low loss RF cables, and omnidirectional antennas at the transmitter and receiver ends. The parameter Settings of APF-RNS network model improved by AO of these two data sets are shown in Table 1.

TABLE 1. Parameter table of AO-APF-RNS.

AO-APF-RNS	Value
Aquila populations N	20
Maximum iteration T	50
Hidden layer units K	[1,50]
Learning rate R	[0.001,0.5]
Known CSI length P	20
Predicted CSI length R	10
Batch size	100
Dropout	0.2

B. AO IMPROVED APF-RNS ALGORITHM PERFORMANCE ANALYSIS

In order to verify the convergence speed and optimization speed of AO improved APF-RNS algorithm(AO-APF-RNS), The convergence curve of fitness function was used to compare the APF-RNS algorithm improved by Genetic Algorithm (GA-APF-RNS), APF-RNS algorithm improved by Particle Swarm Optimization (PSO-APF-RNS), and APF-RNS algorithm improved by Sparrow Search Algorithm (SSA-APF-RNS). The simulation results are shown in Fig.4 and Fig.5. The optimization results are shown in Table 2 and Table 3.

It can be seen from Fig.4 that GA-APF-RNS algorithm has poor overall optimization ability when the four algorithms conduct optimization of APF-RNS network parameters. PSO-APF-RNS algorithm can decrease continuously in the process of optimization, but the convergence speed is slow. The SSA-APF-RNS algorithm has strong initial optimization ability. Compared with the other three algorithms, AO-APF-RNS algorithm has better searching ability (lower mean square error) and faster convergence rate. It can be seen from Fig.5 that the AO-APF-RNS algorithm is obviously superior to the other three algorithms, because four updating ideas are adopted according to the hunting behavior of the Aquila during position updating, thus increasing the exploration and development ability of the Aquila in the solution space and improving the execution efficiency of the algorithm.

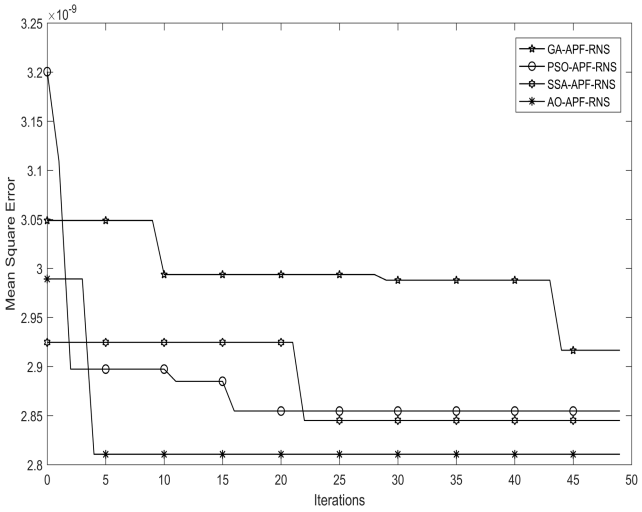

FIGURE 4. Comparison of fitness functions of the four algorithms.

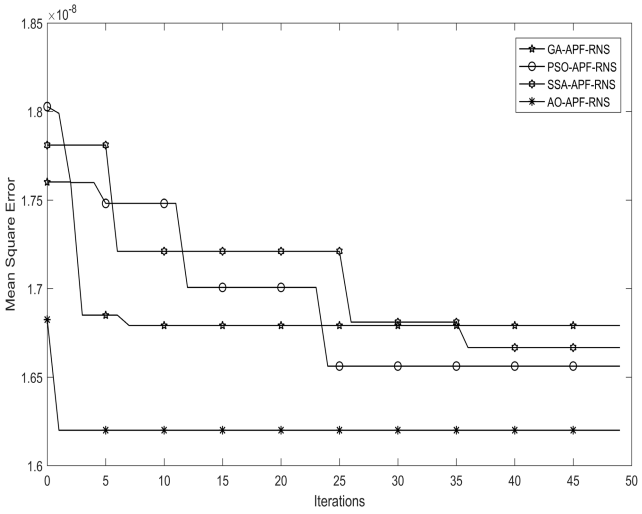

FIGURE 5. Comparison of fitness functions of four algorithms with SNR = 15dB.

TABLE 2. Optimization results.

Algorithm	Hidden layer units	Learning rate
GA	20	0.23349
PSO	29	0.15277
SSA	19	0.13912
AO	**8**	**0.19935**

It can be seen from Table 2 that the AO-APF-RNS algorithm can obtain a small number of hidden layer units while obtaining the lowest fitness function value. The learning rate is higher than that of PSO-APF-RNS algorithm and SSA-APF-RNS algorithm because there is no additive white gaussian noise in the data set and the channel state is not complex. There is no need for too small a learning rate to increase the convergence complexity of the network. Table 3 shows the optimization results with the additive white gaussian noise and a signal-to-noise ratio of 15dB. It can be

TABLE 3. Optimization results.

Algorithm	Hidden layer units	Learning rate
GA	22	0.20952
PSO	19	0.12276
SSA	23	0.04241
AO	**12**	**0.03547**

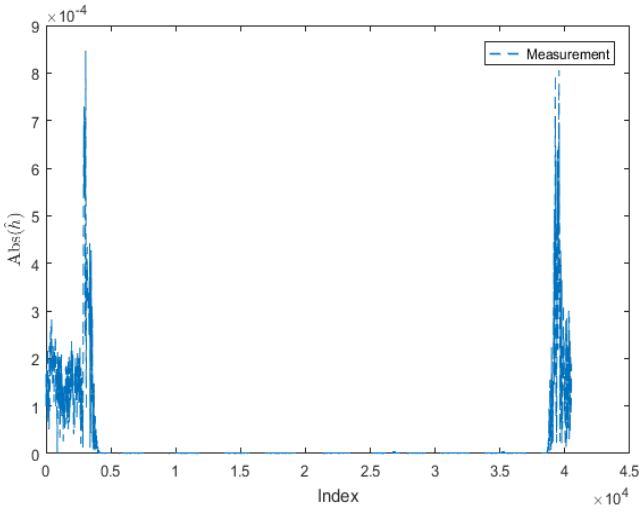

FIGURE 6. Measurement data - Use the first measurement data in NIST AAPlantD1_2GHz_TX1_hpol_run4.

seen from this that the AO-APF-RNS algorithm can obtain a smaller number of hidden layer units while obtaining the lowest fitness function value. The low learning rate is because the channel state is complex at this time, and a small learning rate is required to prevent the loss function from directly crossing the global optimal point, so that the gradient oscillates back and forth near the minimum value. In conclusion, AO-APF-RNS can obtain a smaller number of hidden layer units and a suitable learning rate while obtaining the lowest fitness function value, thereby reducing the complexity of the APF-RNS network and the time spent on prediction.

C. ANALYSIS OF CHANNEL PREDICTION RESULTS

In order to analyze the predictive ability of the AO-APF-RNS algorithm, a piece of data with a length of 40500 timestamps in NIST is used, where the data exists in the form of complex numbers. The data is shown in Fig.6. The horizontal axis is the number of CSI measurement values, and the vertical axis is the result of modulo CSI measurement values. As can be seen from the figure, the channel data changed dramatically in the first 4000 timestamps, and then remained small for a long time (30,000 timestamps), after which the CSI began to change again, with a change process of about 2000 timestamps. In order to verify the performance of AO-APF-RNS algorithm in complex and diversified real channels, set the first 12,000 data sets as the test set, and the last 28500 data sets are used as the training set.

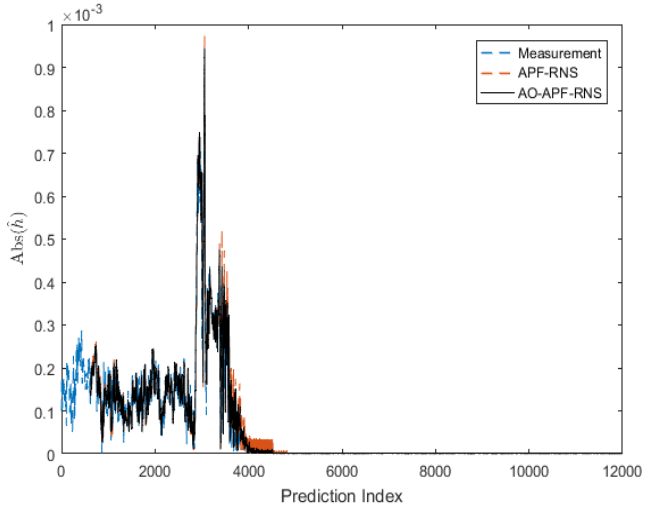

FIGURE 7. Performance comparison - using the first measurements in NIST AAPlantD1_2GHz_TX1_hpol_run4.

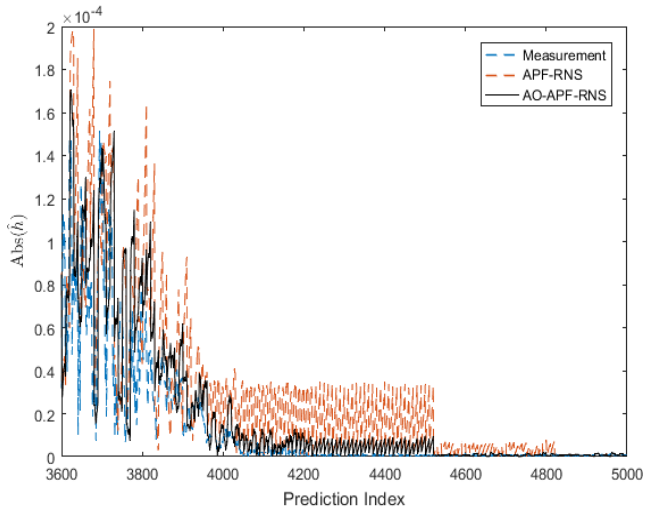

FIGURE 8. Magnification of prediction results - using the first measurements from NIST AAPlantD1_2GHz_TX1_hpol_run4.

The comparison of AO-APF-RNS algorithm in predicting CSI performance is shown in Fig.7.

It can be seen from Fig.7 that AO-APF-RNS is slightly better than APF-RNS in prediction as the channel state keeps changing. But details aren't clear because the numbers are so volatile. In order to analyze the prediction performance of the proposed algorithm, the time stamp data from 3600 to 5000 in the prediction results are amplified as shown in Fig.8. In Fig.8, the predicted results of AO-APF-RNS are more consistent with the reality in the face of the rapidly declining channel state. This is because the predicted results between time stamps 4200 and 4500 use a network of data trained between time stamps 3600 and 3900, when CSI fluctuates greatly. When training the network using data from timestamps 3900 to 4200, the predicted results improved between timestamps 4500 and 4800. During the whole period,

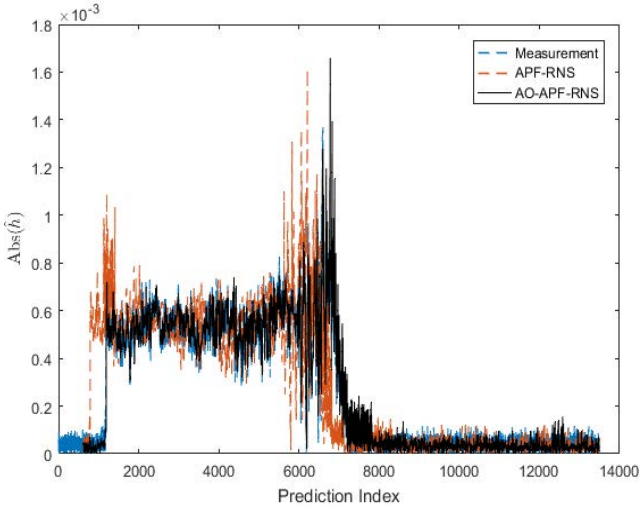

FIGURE 9. Performance comparison - Second measurement AAPlantD1_2GHz_TX1_vpol_run3 using SNR = 15dB in NIST.

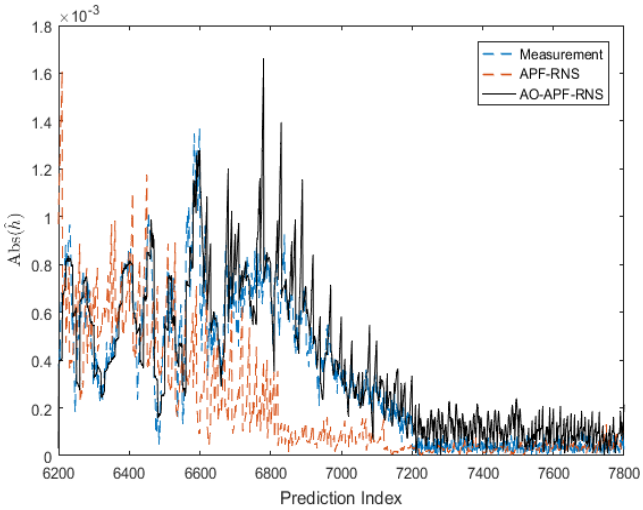

FIGURE 10. Magnified prediction results - using the second measurement AAPlantD1_2GHz_TX1_vpol_run3 from NIST with SNR = 15dB.

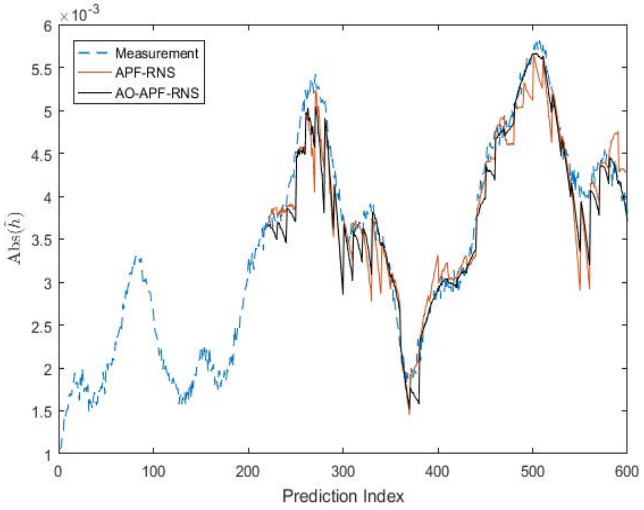

FIGURE 11. Performance comparison - using the second measurement Loc_1211_Lab_139 from the UCI machine learning library.

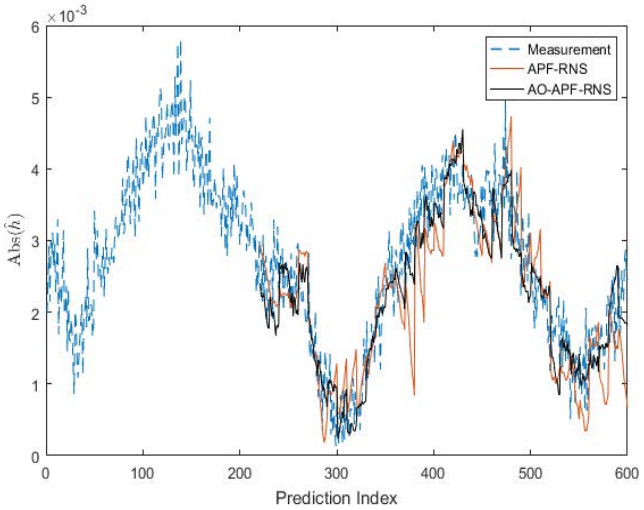

FIGURE 12. Performance comparison - using the fourth measurement Loc_0209_Lab_139 from the UCI machine learning library SNR = 15dB.

The AO algorithm enhances the convergence ability of the network by obtaining the optimal learning rate, and the optimal number of hidden layer units can reduce the complexity of the network and improve the accuracy of prediction.

In order to further verify the performance of AO-APF-RNS algorithm in more complex channels, additive white gaussian noise was added to another data with a length of 44400 timestamp to simulate the channel estimation error in real scenarios, and the signal-to-noise ratio was 15dB. The first 13,500 data were used as the test set and the last 30900 as the training set. CSI prediction results of AO-APF-RNS algorithm are shown in Fig.9.

Fig.9 shows that the AO-APF-RNS algorithm still shows better prediction performance compared with APF-RNS after adding additive Gaussian white noise. In order to analyze the details of the predicted results, the data between time stamps 6200 and 7800 are enlarged, as shown in Fig.10. It can be seen that the prediction effect of APF-RNS is not good during the period of time stamp 6200 to 7200. After 7200, the measured value of CSI approaches 0 and APF-RNS begins to fit. However, certain errors occur during this period and APF-RNS cannot cope with complex and changeable channel state information. However, the channel state information predicted by AO-APF-RNS algorithm is close to the measured value.

In order to verify the generalization ability of AO-APF-RNS, data from UCI Machine Learning Repository were used to verify the generalization ability. The length of the data is 600, the length of the training set is 220, and the length of the test set is 380. The comparison of AO-APF-RNS algorithm in predicting CSI performance is shown in Fig.11, Fig.12 and Fig.13.

As can be seen from Fig.11, the AO-APF-RNS algorithm is slightly superior to APF-RNS on the whole, and

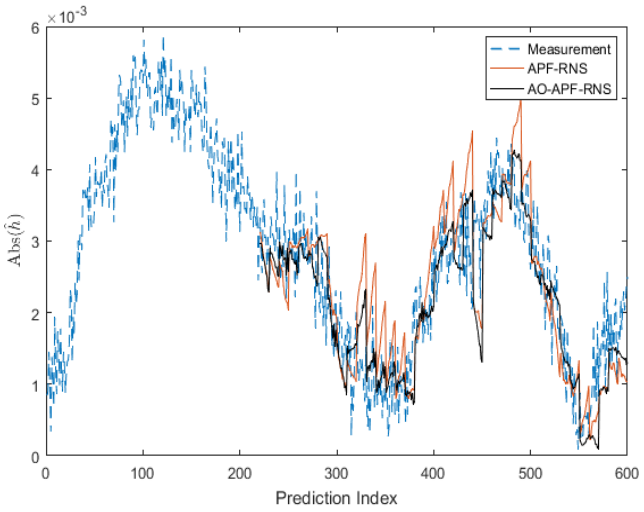

FIGURE 13. Performance comparison - using the sixth measurement Loc_0109_Lab_139 with SNR = 15dB in the UCI machine learning library.

the prediction curve is relatively smooth, which conforms to the real channel state change trend. Fig.12 and Fig.13 show the data after adding additive White Gaussian noise. Compared with APF-RNS, AO-APF-RNS algorithm is closer to the real measured value, which proves that the proposed algorithm can predict CSI under different channels, has certain generalization ability, and improves the prediction accuracy.

V. CONCLUSION

Aiming at the problem of large prediction error of wireless communication system, this paper proposes a wireless channel state prediction method based on improved adaptive and parameter-free recurrent neural structure. The population of Aquila was set as the learning rate and the number of hidden layer units of APF-RNS network, and take the mean square error of APF-RNS network as the optimization target of AO algorithm. Finally, the optimal hyperparameters are obtained to construct the network model to predict the channel state. Channel knowledge, such as long-term statistics or channel parameters, is not required in the prediction process. Therefore, it can be generalized to any propagation environment. The on-line training method can also reduce the pilot cost in the communication link. Simulation results show that compared with APF-RNS algorithm, the proposed algorithm has higher CSI prediction accuracy, and compared with PSO-APF-RNS algorithm, GA-APF-RNS algorithm and SSA-APF-RNS algorithm, AO-APF-RNS algorithm has better optimization ability and faster convergence speed. To sum up, the algorithm proposed in this paper performs well in all aspects.

REFERENCES

[1] S. Gao, P. Dong, Z. Pan, and G. Y. Li, "Deep multi-stage CSI acquisition for reconfigurable intelligent surface aided MIMO systems," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 2024–2028, Mar. 2021.

[2] W. Jiang, H. Dieter Schotten, and J.-Y. Xiang, "Neural network–based wireless channel prediction," in *Machine Learning for Future Wireless Communications*. Hoboken, NJ, USA: Wiley, 2020, pp. 303–325.

[3] W. Jiang, T. Kaiser, F. Luo, and C. Zhang, "From OFDM to FBMC: Principles and comparisons," in *Signal Processing for 5G: Algorithms and Implementations*. Hoboken, NJ, USA: Wiley, 2016, pp. 48–66.

[4] M. Chen, T. Ekman, and M. Viberg, "New approaches for channel prediction based on sinusoidal modeling," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, pp. 1–13, Dec. 2006.

[5] L. Liu, H. Feng, T. Yang, and B. Hu, "MIMO-OFDM wireless channel prediction by exploiting spatial-temporal correlation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 310–319, Jan. 2013.

[6] W. Jiang and H. D. Schotten, "Neural network-based fading channel prediction: A comprehensive overview," *IEEE Access*, vol. 7, pp. 118112–118124, 2019.

[7] Y. Zhang, Y. Wu, A. Liu, X. Xia, T. Pan, and X. Liu, "Deep learning-based channel prediction for LEO satellite massive MIMO communication system," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1835–1839, Aug. 2021.

[8] Y. Zhu, X. Dong, and T. Lu, "An adaptive and parameter-free recurrent neural structure for wireless channel prediction," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 8086–8096, Nov. 2019.

[9] L. Abualigah, D. Yousri, M. A. Elaziz, A. A. Ewees, M. A. A. Al-Qaness, and A. H. Gandomi, "Aquila optimizer: A novel meta-heuristic optimization algorithm," *Comput. Ind. Eng.*, vol. 157, Jul. 2021, Art. no. 107250.

[10] C. L. Giles, S. Lawrence, and A. C. Tsoi, "Noisy time series prediction using recurrent neural networks and grammatical inference," *Mach. Learn.*, vol. 44, nos. 1–2, pp. 161–183, 2001.

[11] (Jan. 2017). National Institute of Standards and Technology. *Networked Control Systems Group—Measurement Data Files*. Accessed: Mar. 18, 2022. [Online]. Available: https://www.nist.gov/ctl/smart-connected-systems-division/networked-control-systems-group/measurement-data-files

[12] N. A. M. AlHajri and R. Shubair. (Nov. 2018). *2.4 GHz Indoor Channel Measurements Data Set*. Accessed: Mar. 18, 2022. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/2.4+GHZ+Indoor+Channel+Measurements

[13] W. Lu, L. Ma, H. Chen, X. Jiang, and M. Gong, "A clinical prediction model in health time series data based on long short-term memory network optimized by fruit fly optimization algorithm," *IEEE Access*, vol. 8, pp. 136014–136023, 2020.

[14] T. W. Y. Jialun and H. Gaoming, "Optimization of LSTM radar beam scanning behavior prediction based on fruit fly algorithm," *Fire Control Command Control*, vol. 45, no. 3, p. 6, 2020.

[15] T. Zemen, C. F. Mecklenbrauker, F. Kaltenberger, and B. H. Fleury, "Minimum-energy band-limited predictor with dynamic subspace selection for time-variant flat-fading channels," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4534–4548, Sep. 2007.

[16] B. Wang, X. Gu, L. Ma, and S. Yan, "Temperature error correction based on BP neural network in meteorological wireless sensor network," *Int. J. Sensor Netw.*, vol. 23, no. 4, pp. 265–278, Jan. 2017.

[17] L. Jing, "Massive MIMO system channel modeling and prediction technology research," Ph.D. dissertation, School Telecommun. Eng., Xidian Univ., Xi'an, China, 2020.

[18] Z. Guangwei, "Research on wireless channel prediction algorithm based on quantum singular value estimation method," Ph.D. dissertation, School Electron. Inf., Hangzhou Dianzi Univ., Hangzhou, China, 2020.

[19] C. Zhongkang, "Research on channel prediction and signal detection of high-speed railway mobile communication based on deep learning," Ph.D. dissertation, School Commun. Inf. Eng., Nanjing Univ. Posts Telecommun., Nanjing, China, 2020.

[20] T. Zheng, T. Ling, Z. Yao, W. Wang, X. Li, and S. Jin, "Large scale MIMO channel prediction method based on improved Kalman filter," *Radio Commun. Technol.*, vol. 47, no. 4, pp. 459–465, 2021.

[21] Y. Hu, C. Yang, and T. Liu, "Wireless channel prediction: The federal learing with centralized communication overhead," *J. Signal Process.*, vol. 37, no. 10, pp. 1930–1940, 2021.

[22] L. Yang, Z. Wang, B. Hu, and Q. Nie, "A novel time-varying channel prediction method based on deep learning," *Telecommun. Sci.*, vol. 37, no. 1, pp. 39–47, 2021.

QINGLI LIU received the Ph.D. degree from Northeast University, China, in 2012. He was a Visiting Scholar at the Department of Electronic Engineering, Utah State University, from September 2015 to March 2016. Since 2019, he has been a Professor with the Information Engineering College, Dalian University. His current research interests include satellite networks, wireless communication, and unmanned aerial vehicle systems.

NA CAO received the B.S. degree in network engineering from Zaozhuang University, Zaozhuang, China, in 2020. She is currently pursuing the master's degree with the Communication and Network Laboratory, Dalian University. Her current research interest includes wireless communication.

MENGQIAN LI received the B.S. degree in network engineering from Taiyuan University, Taiyuan, China, in 2019. She is currently pursuing the master's degree with the Communication and Network Laboratory, Dalian University. Her current research interest includes wireless communication.

ZHENYA ZHANG received the B.S. degree in the internet of things from the Henan University of Science and Technology, Luoyang, China, in 2019. He is currently pursuing the master's degree with the Communication and Network Laboratory, Dalian University. His current research interest includes wireless communication.

GUOQIANG YANG received the B.S. degree in computer science and technology from Tongling University, Tongling, China, in 2020. He is currently pursuing the master's degree with the Communication and Network Laboratory, Dalian University. His current research interest incudes wireless communication.