

Received March 26, 2022, accepted April 8, 2022, date of publication April 21, 2022, date of current version April 28, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3169174

# A Novel Wireless Propagation Model Based on Bi-LSTM Algorithm

YU LU YANG, GUO CHUN WAN<sup>id</sup>, AND MEI SONG TONG<sup>id</sup>, (Senior Member, IEEE)

Department of Electronic Science and Technology, Tongji University, Shanghai 201804, China

Corresponding author: Mei Song Tong (mstong@tongji.edu.cn)

**ABSTRACT** Establishing accurate wireless propagation models is essential for high-quality communications. Aiming at the low accuracy and complexity of the traditional wireless propagation model, a novel accurate wireless propagation model is proposed based on the bi-directional long short-term memory (Bi-LSTM) algorithm of machine learning. The model uses machine learning technology driven by big data and can achieve high real-time performance with low complexity. Also, it can accurately predict the wireless signal coverage intensity in a new environment. To allow the model to accommodate the actual environment of target areas, the propagation model can be dynamically corrected by deep learning and training. The Bi-LSTM is used to describe the relationship between features themselves and the relationship between features and target values of reference signal receiving power (RSRP). The Bi-LSTM is also used to represent the relationship through a full-connection layer to obtain the results so that sufficient parameter space can be provided for the model. The propagation model parameters are searched and fitted through a full-connection optimization. After training and tuning, the model's predicted value of poor coverage recognition rate (PCRR) can reach 0.2371, while the predicted value of root mean squared error (RMSE) can be 10.4855, which demonstrates the better accuracy of the proposed model.

**INDEX TERMS** Bi-LSTM, deep learning, feature extraction, fully connected layer, wireless propagation.

## I. INTRODUCTION

The rapid development of 5G technology, which can meet more communication needs of users, is inseparable from the reasonable deployment of base stations and the wireless propagation model is the key to solve the problem of base station deployment [1]. This propagation model can estimate indicators, such as cell coverage and inter-cell network interference, by predicting the propagation characteristics of radio waves in the designated communication coverage area. The results can then help suppliers and wireless operators select reasonable base station sites. Efficient network estimation is of great significance in the entire planning process of wireless network and can help determine the accuracy of 5G network deployment [2]. Since the propagation process of radio waves is greatly affected by environmental factors, the existing wireless propagation models need to collect a large amount of engineering data for further corrections in practical applications [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Jon Atli Benediktsson<sup>id</sup>.

There have been many papers addressing the path loss (PL) models and many effective propagation models have been proposed. Phillips *et al.* analyzed 28 PL models by using a large amount of data from wireless networks in rural New Zealand to obtain a minimum root mean squared error (RMSE) of 12 dB in a relatively simple rural environment [4]. However, the current research on wireless propagation models mainly focused on the use of measured data to modify the traditional empirical model, and the most common correction method is the least square method [5]. Only a few scholars considered combining wireless signal propagation research with deep learning methods. For example, Haq *et al.* presented a comparative analysis of a variety of deep features with several sequential learning models to select the optimized hybrid architecture for energy consumption prediction [6]. A novel architecture called 'AB-Net': a one-step forecast of RE generation for short-term horizons by incorporating an autoencoder (AE) with bidirectional long short-term memory (BiLSTM) [7]. Also, a novel hybrid architecture called 'CL-Net' based on convolutional long short-term memory (ConvLSTM) and long short-term memory (LSTM) is proposed for multi-step SOH and power

consumption forecasting [8]. Zhang *et al.* presented an accurate and efficient propagation modeling technique, which is combining the vector parabolic equation (VPE) and waveguide mode theory [9]. Adeogun *et al.* used the MoM estimator to calibrate the random polarization propagation map model, eliminating the intermediate multipath parameter extraction. However, due to the need to derive a new expression for each model, the applicability is severely limited [10].

With the emergence and development of deep learning methods, it is possible to build a more adaptive wireless propagation model. Compared with the traditional method of computing time and space, the deep learning method is more efficient. For example, the training speed is greatly improved, and the training effect becomes better than the traditional method [11]. Liu *et al.* proposed a Node Classification Redundancy Decoding (NCRD) algorithm based on the channel reliability of the received sequence. Additional decoding can be obtained through list decoding method and neural network “learning” NCRD algorithm [12]. Zhang *et al.* introduced a new method of TR channel modeling based on Bp neural network. According to the basic principle of time reversal, the peak feature can be firstly obtained by the fitting method and then processed by using the principal component analysis technique. Finally, deconvolve the received signal to get the channel characteristics [13]. However, this method cannot achieve the global optimal solution in many situations, and can only achieve a partial optimal solution, which is prone to problems such as “local convergence.” Gao *et al.* proposed that Deep Learning (DL) technology has become the mainstream direction of wireless physical layer transmission research, introduced the RNN algorithm based on DL technology and summarized and explained its application results [14]. RNN is an improvement of BP, but it also has the problem of gradient explosion and gradient disappearance. Zhao *et al.* adopted the multi-layer neural network algorithm LSTM, and uses historical experience data to analyze and predict short-wave sky-wave communication circuits to obtain higher prediction accuracy than the theoretical calculation method [15]. However, the LSTM method only considers one-way propagation, and cannot predict the data law after the time node. On this basis, using Bi-LSTM to establish a prediction model has better prediction performance and generalization performance compared with LSTM. Al Barazanchi *et al.* proposed a new framework scheme for the path loss in wireless body area network and apply the new Framework Scheme on three case scenarios that help us to get parameters by measuring vital information about the human body [16]. In addition, the influence of artificial intelligence in healthcare and the gains it can provide in the face of the COVID-19 pandemic have been highlighted [17]. In a summary, this paper designs and selects effective features, and using Bi-LSTM to establish a prediction model, it is possible to predict the value of reference signal receiving power (RSRP) of a new specific geographic location quickly and accurately. The whole paper is divided into five chapters. The first part introduces the research status

of wireless propagation models, the second part introduces feature design and feature extraction, and the third part introduces how to select effective feature parameters. The fourth part introduces the establishment of the wireless propagation model based on Bilstm and the data result test, and the fifth part summarizes the full text.

## II. FEATURE DESIGN AND EXTRACTION

A lot of research has been done on wireless channel propagation models at home and abroad. In engineering practice, a large amount of data can be obtained through the collection of historical data. However, which data can better characterize the problem target and how to select the data to improve the effect of the machine learning model, need to be solved through feature engineering. In order to study wireless signal propagation, this paper designs a better and more accurate wireless propagation model. The wireless propagation model is the basis of cell planning for mobile communication networks. The establishment of the wireless propagation model provides an important basis for judging whether the cell planning is reasonable or not. In each cell, it is hoped that areas far away from the base station can be well covered. The average signal acceptance rate can be kept within the acceptable range of users.

The data used in this paper is the wireless signal propagation data set provided by Huawei. The data set contains the measured data of 4000 cells. According to the classical model Cost231-Hata model in the city and the data set information, consider the factors that may affect the average signal receiving power to design suitable features for the model [18]. In the traditional wireless propagation model, the predictive variable is the propagation path loss [19], but in practice, the more intuitive predictive variable is the average signal reception rate, which can be expressed as the RSRP. Therefore, the average signal reception rate in the data set is set as the model predictive variable, and the average signal reception rate is equal to the transmit power of the cell transmitter minus the propagation path loss. First, through the analysis of the Cost231-Hata model, the elements in the model are designed as features, that is, the carrier frequency, the effective height of the user’s antenna, and the link distance [20]. These characteristics will affect the propagation path loss, thereby affecting the signal received power. The next step is to analyze the data set and extract valid data features in the data set. Finally, the feature can be further extracted based on the geographic location relationship between the target grid and the transmitter [21].

### A. FEATURE DESIGN

According to the design characteristics, we use the Cost231-Hata model and its formula can be written as

$$PL = 46.5 + 33.91 \lg(f) - 13.82 \lg(h_b) - \alpha(h_m) + [44.9 - 6.55 \lg(h_b)] \lg(d) + C_m \quad (1)$$

where represents the propagation path loss, represents the carrier frequency, represents the effective height of the base

station antenna (m), represents the user antenna height correction term, represents the link distance (km), and represents the scene correction constant. It can be seen from the formula that the carrier frequency, the effective height of the base station antenna, and the link distance all affect the propagation loss. According to the following formula

$$RSRP = P_t - PL \tag{2}$$

where  $P_t$  is the transmit power of the cell transmitter, it can be seen that the change of propagation loss will directly affect the average signal received power RSRP. Therefore, factors related to propagation loss need to be taken into consideration. Carrier frequency  $f$ , base station antenna effective height  $h_b$ , and link distance  $d$  are designed as features respectively. At the same time, the transmitted power of the cell transmitter will also directly affect the received average signal power, so  $P_t$  will also be designed as a feature. Also, we need to analyze the data set and geometric design features. Since the link distance can be expressed by the following formula

$$d = \sqrt{(X_{cdl} - X)^2 + (Y_{cdl} - Y)^2} \tag{3}$$

where  $X_{cdl}$  represents the X coordinate of the grid position of the site to which the cell belongs, X represents the X coordinate of the grid position,  $Y_{cdl}$  represents the Y coordinate of the grid position of the site to which the cell belongs, and Y represents the Y coordinate of the grid position.

Tall buildings will hinder the normal propagation of signal lines, shield and interfere with the transmitted signal, which makes electromagnetic waves produce complex transmission, diffraction and other losses and cause loss. Therefore, it is necessary to consider the height of the building at the grid  $(C_x, C_y)$  and the building at the grid  $(X, Y)$  where the cell site is located. So this paper takes the two elements of Cell Building Height and Building Height as features, called  $h_{cb}$  and  $h_b$ .

At the same time, the propagation path will also be affected by the environmental geomorphology, that is, the type of features on the grid may affect the average signal received power. However, only the ground object type index is provided in the data set, so this paper expands these two ground object type elements into a 20-dimensional vector. The corresponding feature type is marked with 1, and the rest are marked with 0. The 20-dimensional vector of the ground object type of the grid  $(C_x, C_y)$  where the cell site is located and the ground object type of the grid  $(X, Y)$  are selected as the features of this paper after processing.

In addition to a simple analysis of the data set, design three features based on physical knowledge and geometrically considering the angle of antenna emission are needed.

To simplify the actual situation, this paper makes a description of the meaning of the engineering parameter data, as shown in the Fig. 1 and Fig. 2. According to the top view above shown in Fig. 2, the horizontal angle of the grid  $(X, Y)$  is important. The difference between the horizontal

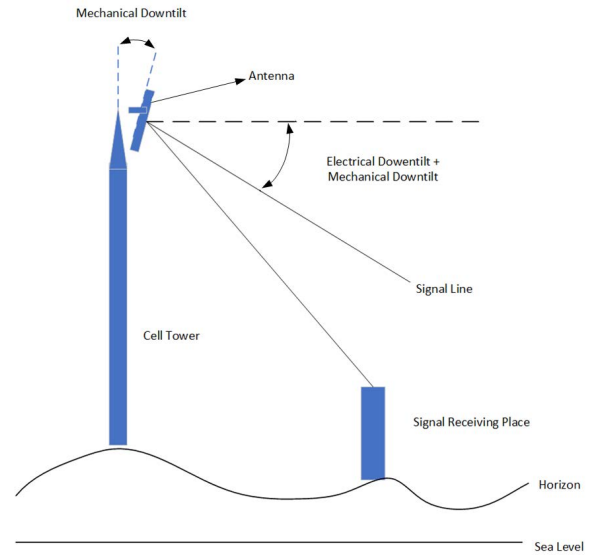


FIGURE 1. Side view of wireless signal propagation model.

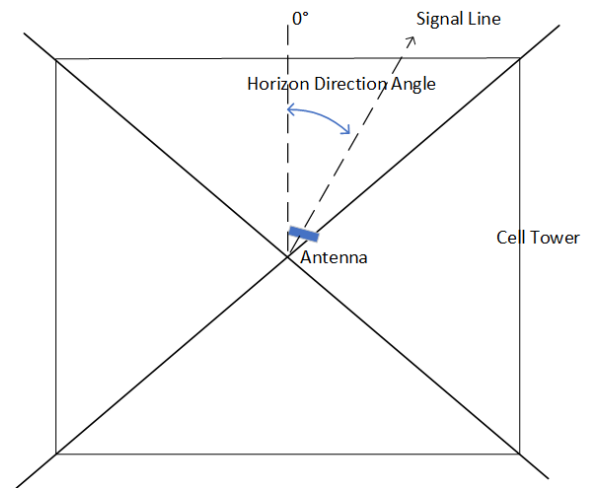


FIGURE 2. Top view of wireless propagation model.

angle of the grid and the grid where the cell is located can be used as a feature. And considering that the transmitter and the target grid are at different altitudes in actual situations, the difference between the vertical angle and the sum of the vertical electrical downtilt angle and the vertical mechanical downtilt angle of the cell transmitter is related to the transmission of the signal line, so the vertical angle difference is taken as a feature. What's more, it can also be seen intuitively that the linear distance between the measuring point and the transmitter can be used as a feature. So we can obtain the horizontal angle difference, vertical angle difference, and the linear distance between the measuring point and the transmitter, respectively. The corresponding formulas can be written as

$$\theta_x = \left| \arcsin \left( \frac{X_{cdl} - X}{d} \right) - \theta_A \right| \tag{4}$$

TABLE 1. Units for magnetic properties.

Feature name	Unit
$f$	MHZ
$h_d$	m
$d$	m
$P_t$	dBm
$h_{CB}$	m
$h_B$	n
$(C_X, C_Y)$	-
$(X, Y)$	-
$\theta_X$	Deg
$\theta_Y$	Deg
$l$	m

$$\theta_y = \left| \arctan \left( \frac{A_c + h_b - A}{d} \right) - (\theta_E + \theta_M) \right| \quad (5)$$

$$l = \sqrt{d^2 + (A_c + h_b - A)^2} \quad (6)$$

where  $\theta_A$  represents the horizontal direction angle of the cell transmitter,  $\theta_E$  represents the vertical electrical downtilt angle of the cell transmitter,  $\theta_M$  represents the vertical mechanical downtilt angle of the cell transmitter,  $A_c$  represents the altitude of the grid  $(C_X, C_Y)$  where the cell site is located, and  $A$  indicates the altitude of the grid  $(X, Y)$ , respectively.

**B. DESIGN OF ALL CHARACTERISTIC VALUES**

In summary, this paper has designed 11 features in total as shown in Table 1. They are carrier frequency  $f$ , base station antenna effective height  $h_b$ , link distance  $d$ , cell transmitter transmission power  $P_t$ , cell building height  $h_{bc}$ , building height  $h_b$ , ground object type of the grid  $(C_X, C_Y)$ , ground object type of the grid  $(X, Y)$ , horizontal angle difference  $\theta_x$ , vertical angle difference  $\theta_y$ , and linear distance  $l$  from the measurement point to the transmitter, respectively.

**III. FEATURE SELECTION**

Feature selection is to select a subset of relevant features from the designed features to calculate the correlation between two variables. An efficient machine learning model is sensitive to the distribution of feature items, so preprocessing of data is important for machine learning basis. This can not only alleviate the dimensionality catastrophe problem, but also remove irrelevant features to reduce the difficulty of the learning task [22]. In this paper, Statistical Product and Service Solutions (SPSS) is used to average and normalize the data to test the correlation between the two variables, that is, to test the correlation between the selected features and the standard value RSRP, and to calculate the correlation between the features and the target. Based on the above analysis, this paper will screen out suitable features through the following three steps.

**A. PREPROCESSING OF FEATURE SELECTION**

Obviously, the features involved are all discrete variables. Based on the RSRP measurement data set of the cell, the above features can be preliminarily filtered by the variance of the features [23]. Firstly, randomly select the given training

dataset of three cells for preliminary processing, and then the features  $d, l, \theta_X, \theta_Y, h_B$  are obtained on the basis of the engineering parameter data of each cell.

According to the training dataset, it can be found that the features of one cell remain unchanged, including  $f, h_B, P_t, h_{CB}$ , and Cell Clutter Index. The variances of these features are all 0, which means the features do not diverge. That is, the samples have basically no differences in these features. In sum, the above features can be initially eliminated. The feature index in the data set has no numerical meaning, so this paper expands it to a 20-dimensional space, and then uses the neural network connection to fit the impact of the ground object type on the target value [24]. When the topography belongs to a certain type, the corresponding position is 1, and the rest are 0. In this way, only the connection to the position in the model network will play an effect on the target value, and the connection can also reflect the effect of the target value. Therefore, this feature can be directly eliminated and obtains the following 5 divergent features:  $d, h_B, \theta_X, \theta_Y, l$ .

**B. PRINCIPAL COMPONENT ANALYSIS**

There are too many features designed based on the geometric location of the site and target grid and features based on the channel model, and some features have little impact on RSRP. Therefore, we can first reduce the dimension of the features to reduce the number of features by using Principal components analysis (PCA) [25]. PCA was first proposed by Hotelling in 1933. It is a multivariate statistical method that converts multiple original indexes into several comprehensive ones under the assumption of losing as little information as possible [26]. These comprehensive indexes generated by the transformation are usually called the principal components. Each one is a linear combination of the original variables. At the meantime, all these principal components are not related to each other, which makes them have some advantages over the original ones. It becomes easier and more efficient to grasp the main contradiction of the complex problem to reveal the regularity between internal variables of things because the components to consider are fewer and sufficient [27].

The general steps of using the PCA can be summarized as follows. First, the initial analysis variables are selected. Second, according to the characteristics of the initial ones, we determine whether adopting the covariance matrix or using the correlation matrix of these variables to further build comprehensive indexes. Third, both the eigenvalues and the corresponding standard eigenvectors of the determined matrix are calculated. Fourth, we can easily judge whether there is an obvious multicollinearity based on the previous results. If so, then we go back to the first step, otherwise we stop the process. Through the above steps, the number of principal components and their expressions can be identified. Finally, the principal components are used to analyze the problem.

In this paper, the software SPSS is used to realize PCA, and the components are selected based on the contribution rate of eigenvalues [28]. Table 2 shows the analysis of total

**TABLE 2. Explanation of total variance (cell 0).**

Initial eigenvalue			
Ingredient	Eigenvalues	Eigenvalues Percentage of variance of accidental eigenvalues	Accumulation%
$d$	2.563	42.721	42.721
$l$	1.024	17.075	59.796
$\theta_X$	0.974	16.236	76.033
$\theta_Y$	0.834	13.901	89.933
$h_B$	0.604	10.067	100.000

**TABLE 3. Explanation of total variance (cell 1).**

Initial eigenvalue			
Ingredient	Eigenvalues	Eigenvalues Percentage of variance of accidental eigenvalues	Accumulation%
$d$	2.781	46.353	46.353
$l$	1.052	17.540	63.893
$\theta_X$	0.944	15.725	79.619
$\theta_Y$	0.740	12.331	91.950
$h_B$	0.483	8.050	100.000

**TABLE 4. Explanation of total variance (cell 2).**

Initial eigenvalue			
Ingredient	Eigenvalues	Eigenvalues Percentage of variance of accidental eigenvalues	Accumulation%
$d$	2.945	49.081	49.081
$l$	1.021	17.023	66.104
$\theta_X$	0.982	16.367	82.471
$\theta_Y$	0.762	12.696	95.167
$h_B$	0.290	4.833	100.000

variance for Cell 0 and the eigenvalues of the first component and the second one are greater than 1 while those of the third and fourth ones are close to 1. Besides, the first four factors account for 89.933% of the total variance. Table 3 and Table 4 provide the analysis of total variance for Cell 1 and Cell 2, respectively, and the eigenvalues of the above mentioned components are similar while the accumulative variance contribution of these four factors collected from the two cells are 91.950% and 95.167%, respectively. According to these results, the main components of RSRP can be determined by  $d$ ,  $l$ ,  $\theta_X$ , and  $\theta_Y$ .

**TABLE 5. Feature name and its relevance to the target (cell 0).**

Initial eigenvalue		
Sort	Feature name	Relevance of this feature to the target
1	$d$	-0.263
2	$l$	-0.257
3	$\theta_X$	-0.136
4	$\theta_Y$	-0.091

**TABLE 6. Feature name and its relevance to the target (cell 1).**

Initial eigenvalue		
Sort	Feature name	Relevance of this feature to the target
1	$d$	-0.258
2	$l$	-0.240
3	$\theta_X$	-0.175
4	$\theta_Y$	-0.082

**TABLE 7. Feature name and its relevance to the target (cell 2).**

Initial eigenvalue		
Sort	Feature name	Relevance of this feature to the target
1	$d$	-0.235
2	$l$	-0.270
3	$\theta_X$	-0.159
4	$\theta_Y$	-0.070

**C. CORRELATION ANALYSIS OF FEATURES**

Correlation analysis is the process to measure the degree of linear correlation between things and express it with appropriate statistical indicators. It is a commonly used statistical method to study the closeness of variables. This paper again uses SPSS to analyze the bi-variate correlation between the target, RSRP, and the principal components,  $d$ ,  $l$ ,  $\theta_X$ ,  $\theta_Y$ , respectively. Based on the absolute value of the correlation coefficient, the order of these four features is as following tables. From three tables below, it can be found that the absolute value of the correlation coefficient between  $d$  and RSRP is the largest, and that value between  $\theta_Y$  and RSRP is the smallest. Overall, the order of the absolute value of the correlation coefficient between the features and the target is  $d > l > \theta_X > \theta_Y$ .

**IV. ESTABLISHMENT OF WIRELESS PROPAGATION MODEL**

For the prediction of RSRP, this model designs and selects multiple related features, aiming to find the implicit relationship between multiple features and RSRP, so as to realize

the prediction of RSRP value in unknown areas. The model mainly takes multiple selected features as input, and uses a layer of Bi-LSTM network to capture the connections between input features. Although Bi-LSTM is mainly used for context sequence coding in the field of natural language processing, it also has a good effect on deep modeling of multiple input features and characterizing their implicit relationships [29]. In this paper, the feature representation of Bi-LSTM passes through a fully connected layer to obtain the final model prediction result [30].

**A. MODEL ESTABLISHMENT-FEATURE FITTING MODEL BASED ON BI-LSTM NETWORK**

However, due to the large number of input features, and the relationship between the features and the target value is difficult to be explicitly deduced, so we can only use machine learning methods to create a reasonable model and set enough parameters to fit the implicit function between multiple input features and target values.

1) BI-LSTM LAYER

In Long Short Term Memory (LSTM), a memory unit including input gate, forget gate and output gate is used to replace the implicit function in RNN. The memory unit can make the choice of retaining and forgetting the output content, and then can retain the long sequence of calculation information. LSTM adds three gates to prevent the gradient from disappearing and improve memory for a long time. Input gate ( $i_t$ ), forget gate ( $f_t$ ), output gate ( $o_t$ ), internal memory ( $c_t$ ) and LSTM unit output ( $h_t$ ) at time step  $t$ . The input gate controls the input of the output information of the previous LSTM unit to the input of the LSTM unit information, and the LSTM unit retains the sequence past information. The calculation can be shown as

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{7}$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{8}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{9}$$

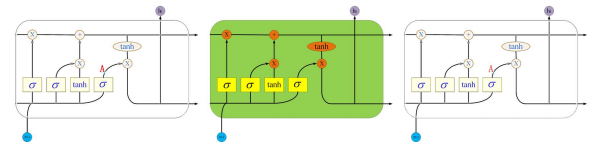
$$c_t = f_t c_{t-1} + i_t \tan h(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{10}$$

$$h_t = o_t \tan h(c_t) \tag{11}$$

where  $\sigma$  is the sigmoid activation function, and

$$f(z) = \frac{1}{1 + e^{-z}} \tag{12}$$

The one-way LSTM model uses the previous information to derive the follow-up information. In this paper considering the information laws before and after the prediction can improve the prediction accuracy, so Bi-LSTM neural network is adopted, and two LSTM neural networks are used. Bi-LSTM is a recurrent neural network [31]. It adds four gates to the basic recurrent neural network to control the input and output of the model. It can be used to realize the implicit association and dependency representation of long sequences. Since the cyclic structure of Bi-LSTM can repeatedly add and multiply features, this paper uses this feature



**FIGURE 3. Cyclic structure of Bi-LSTM that can repeatedly perform the addition and multiplication for the features of network.**

to make the network repeatedly process the input features while ensuring enough parameter space to fit the implicit relationship between input features and output RSRP. The principle of Bi-LSTM is shown in Fig. 3.

The input  $X_t$  and output  $H_t$  of the network are both a single-column vector, the input of an LSTM unit and the output of the previous LSTM unit are subjected to sigmoid and tanh operations, and then the sum operation is performed to form the output of the current LSTM unit and pass it to the next LSTM unit. Repeat the above calculations until the calculations of all units are completed, and the final output is given.

2) FULLY CONNECTED LAYER

The fully connected layer is a basic neural network structure, each node of which is connected to all nodes of the upper layer, and is used to integrate the features extracted from the upper layer. Due to its fully connected characteristics, generally the fully connected layer has the most parameters, so the fully connected layer can provide more parameters for the model and ensure the fitting effect of the model [32]. The principle of full connection is shown in Fig. 4. Firstly, the input data enters the fully connected layer from the input layer node, and then passes through the operation between the hidden layer and the parameters to reach the hidden layer node. The data enters the activation function after passing through the hidden layer nodes, and uses the activation function to activate to ensure the non-linear characteristics of the fully connected layer. After the activation function, a data offset is added to the node to achieve correction. In this way, after multiple hidden layers, activation functions and biases, the data is output to a fully connected layer [33]. In practice, the input and output of the fully connected layer can be adjusted by controlling the dimensions of input nodes and output nodes.

In this paper, the output of the two-way LSTM network is used as the input of the fully connected layer, and the hidden layer of the fully connected layer is set as one layer [34]. After a layer of full connection, the output of the two-way LSTM is reduced to one dimension, and the prediction of the model RSRP value is obtained [35]. The basic idea of this work has been presented in a conference [36] but we significantly expand it in this paper.

**B. MODEL INPUT DATA PREPROCESSING**

Data preprocessing is mainly to process the original data set into a data set that meets the model input requirements [37].

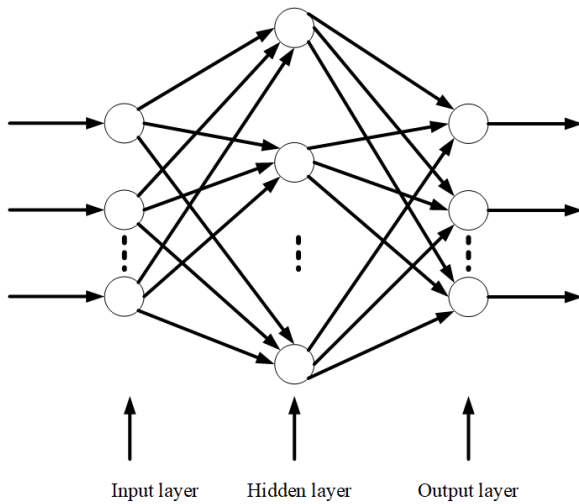


FIGURE 4. Fully connected layer schematic.

In practice, there are many ways of data preprocessing, such as principal component analysis, dimensionality reduction, and data discretization analysis. In this paper, the data is first divided into training set and test set. Because of the large amount of data, the split ratio is set to 0.02, that is, 2% of the data is selected for the test set and 98% of the data is used for training. Considering the characteristics of the data set and the degree of noise of the data set, the data set is mainly processed in the following aspects.

### 1) SCREENING OF RAW DATA

There are some missing values in the data set. If they are not processed, the model will fall into infinite variables after entering the model. Therefore, we need to deal with the missing values reasonably. Considering that the training set needs to provide the model with accurate and true features, we directly delete the missing values in the training set, which can avoid the problem of inaccurate data set features caused by blindly filling in missing values. In the same time, it is inconvenient for us to delete some data when testing data in the test set, so the missing value in the test set can be replaced with the average value of the column where the missing value is located.

### 2) DATA NORMALIZATION PROCESSING

The magnitude of the data in the data set varies greatly, ranging from 1 to tens of thousands. If the data is directly input to the model, the speed of model training will be reduced, resulting in unbalanced parameter distribution and reducing the accuracy of the model. Therefore, this paper normalizes the data, using the maximum and minimum normalization: Assume that the maximum value of a certain feature of the data set is  $X_{max}$ , and the minimum value is  $X_{min}$ , the processing formula for the data in this feature after the maximum and

minimum normalization is shown as

$$X_1 = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (13)$$

where  $X_1$  is the normalized data. In this way, we can compress all the data to the interval (0, 1) to ensure that all the data are at the same magnitude. In the same time the normalized data distribution still maintains the original relative size and dispersion.

### 3) TREATMENT OF DATA IMBALANCE

This paper sets the value of the weak coverage decision threshold as  $-103$  dBm. If the predicted or measured RSRP value is less than  $-103$  dBm, it is regarded as weak coverage and marked as 1. If the value is greater than or equal to  $-103$  dBm, it is regarded as non-weak coverage and marked as 0. Under the requirement that the evaluation index PCRR is greater than or equal to 20%, the positive sample in this data set is defined as the data strip with the RSRP value less than  $-103$  dBm, and the negative sample is defined as the data strip with the RSRP greater than  $-103$  dBm. Through the calculation, it is found that the ratio of positive and negative sample data is about 1 : 5, that is, the difference between the positive and negative samples of the data set is large, which will cause the imbalance of model training, and the model may be more inclined to find the parameters that make the prediction results of a large proportion of the data set correct. Therefore, in order to avoid that the features fitted by the trained model tend to account for a large proportion of negative data, the negative data is deleted. However, considering the proportional distribution of the test data should still be focused, so only a part of the negative data is removed, so that the ratio of positive and negative data after removal is approximately 1 : 2.

### C. MODEL TRAINING

Model training is mainly to repeatedly input data into the model, and then use the neural network to fit the model, so that the model can correctly predict the target value. Model verification is to observe the effect of model training by viewing the target value predicted by the model in real time.

#### 1) DEFINITION OF MODEL LOSS

In this model, the two parts of the model loss are respectively measured by the corresponding indicators of PCRR and RMSE, and the product of the two parts of the loss is used as the index of fitting. The loss used to measure PCRR is defined by the following formula

$$\text{loss}_p = \text{count} \left[ \frac{(-103 - Y_p)(-103 - Y) - |(-103 - Y_p)(-103 - Y)|}{n} \right] \quad (14)$$

where  $Y$  is the true value of RSRP,  $Y_p$  is the predicted value of RSRP,  $n$  is the number of samples used for prediction, and  $-103$  is the standard for measuring positive and negative data. In this loss, this paper uses the absolute value method to detect whether the prediction result meets the standard, and

then count the non-zero format, that is, the prediction result that does not meet the requirements of PCRR. The purpose of the final fitting is to prefer that  $Y$  and  $Y_p$  are both on the same side of  $-103$ , which also meets the requirements of PCRR.

$$\text{loss}_R = \sqrt{\frac{\sum (Y_p - Y)^2}{n}} \tag{15}$$

$$\text{loss} = \text{loss}_p \cdot \text{loss}_R \tag{16}$$

The loss used to measure the RMSE indicator is defined by the formula as shown in (15), which represents the mean square error of the predicted RSRP value  $Y_p$  and the real RSRP value  $Y$ . Our goal is to make as small as possible. And the final error  $\text{loss}$  used for model fitting is defined by the formula as shown in (16). We make  $\text{loss}$  as small as possible to achieve both  $\text{loss}_p$  and  $\text{loss}_R$  as small as possible, so that the model prediction results obtained will tend to be larger in PCRR and smaller in RMSE.

### 2) PARAMETER SETTING

In training, the parameter setting mainly includes the dimensions of the input data, the number of hidden layer nodes of the Bi-LSTM, the number of units, the input and output dimensions of the full connection, the learning rate and the optimization algorithm. The dimensionality of the input data depends on the design of our feature engineering. According to the design of the feature project, the dimension of the input data is 49, including 9 one-dimensional features: carrier frequency  $f$ , base station antenna effective height  $h_b$ , link distance  $d$ , cell transmitter transmission power  $P_t$ ,  $h_{CB}$ ,  $h_B$ , horizontal angle Difference  $\theta_X$ , vertical angle difference  $\theta_Y$ , the linear distance from the measuring point to the transmitter  $l$ , and 20-dimensional features which includes the feature type of the grid (Cell $_X$ , Cell $_Y$ ) and  $(X, Y)$ . In order to smoothly input the data into the two-way LSTM, this paper copies the input data 15 times to match the input of the two-way LSTM.

The optimization algorithm used in this paper is the Adam optimizer, which mainly contains the significant advantages of simple implementation, efficient calculation, low memory requirements, parameter updates are not affected by gradient scaling, parameters have good interpretability and usually no adjustment or very little fine-tuning, etc. After the tuning of the actual running program, the hidden layer dimension of the two-way LSTM is finally set to 49, the number of units is 15, the input dimension of the fully connected layer is 49 and the output dimension is 1 (the predicted RSRP value), the learning rate is 0.02, and the learning rate is set to be reduced to 0.8 of the original learning rate.

### 3) THE RESULT OF MODEL TRAINING

After the model training is completed, upload the relevant code and configuration files to HUAWEI CLOUD, and use the ModelArts model deployment tool provided by HUAWEI CLOUD to successfully perform deployment predictions.

Predicted value of RSRP/dBm

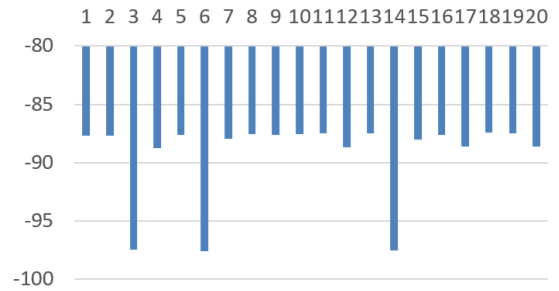


FIGURE 5. Graph of  $\text{loss}_p$  with iteration.

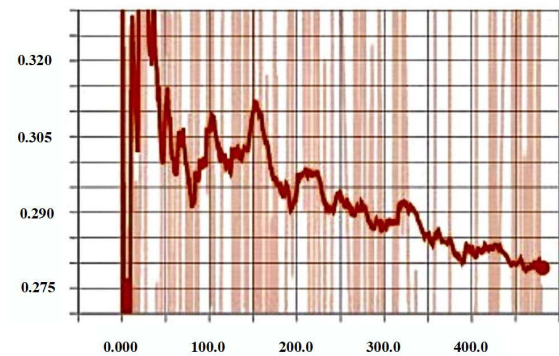


FIGURE 6. Graph of  $\text{loss}_p$  with iteration.

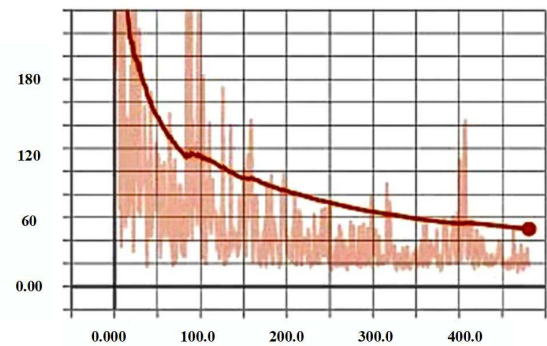


FIGURE 7. Graph of  $\text{loss}_R$  with iteration.

Take the test data in the test set as an example, the first 20 groups of prediction information are shown in the Fig. 5.

In practice, the three losses of the model, are greatly reduced through continuous iteration and parameter update, as shown in the following Fig. 6, Fig. 7, and Fig. 8. The results of the final validation set show that as the three losses continue to decrease, the effect of the model continues to improve, and finally reaches the optimal value of the model. The PCRR value of the model prediction result is 0.2371, which meets the requirement greater than 0.2 and the RMSE



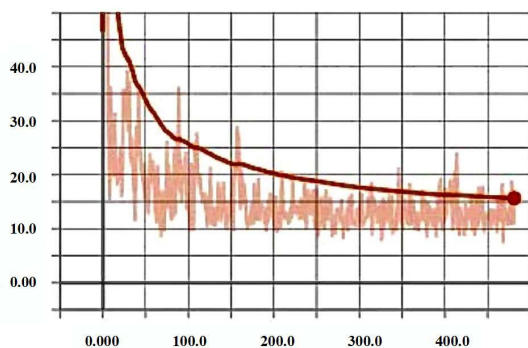


FIGURE 8. Graph of loss with iteration.



FIGURE 9. Visual map of RSRP predicted value.

value is 10.4855. The wireless intelligent propagation model based on deep learning designed in this section can overcome the problem that traditional wireless propagation models need to be corrected in real time, becoming more intelligent. It can be achieved only by using collected historical data combined with deep learning methods.

In addition, for the data given in the training dataset, this paper uses the ArcGIS software ArcSence to visualize the RSRP predicted value, as shown in the Fig. 9. By visualizing the color change in the map, you can intuitively observe the geographic shape of the target grid. Combined with the RSRP predicted value we can clearly visualize the signal power distribution, different colors represent different values, so as to clearly identify the signal strength and coverage area and make it more convenient to browse the test results.

## V. CONCLUSION

With the birth and development of the fifth-generation mobile communication system, people’s requirements for communication quality are getting higher and higher. The quality of network planning and optimization is an important guarantee for the construction of a high-quality network, and the propagation model occupies a vital position in the network planning. It is the basis of the network coverage planning of the mobile communication network. The accuracy of the propagation model is related to whether the community planning is reasonable and whether the operator can meet the needs of users with relatively economical investment. With the rise of artificial intelligence technology in recent

years, an intelligent wireless communication model and optimization of traditional theoretical models have gradually become the focus of research. The accuracy of the intelligent wireless propagation model has a great impact on the quality of network planning, especially for 5G networks with multi-service characteristics. The accuracy of the propagation model is directly related to the speed and accuracy of signal reception. The main contribution of this paper is to combine the deep learning method with the wireless propagation model to establish an intelligent wireless propagation model based on deep learning, which can quickly predict the average signal reception rate in a specific environment. By understanding the current research status of domestic wireless propagation models, combining traditional wireless propagation models and practical experience, 11 eigenvalues are designed from the aspects of model and formula, data set and geometry, and then principal component analysis is used to screen the design features and select, and use the feature with higher comprehensive index for modeling. In this paper, we use Bi-LSTM feature to get the final model prediction result through a fully connected layer. The RMSE values of the training set and the test set are relatively low, which shows that the prediction results of this model are better. In general, the AI-based wireless propagation model designed in this paper has the following advantages:

### A. GOOD GENERALIZATION

The model is trained from a large amount of data, and the fitted features are universal and suitable for prediction in most scenarios.

### B. USING THE BI-LSTM MODEL

This model can capture the implicit relationship between the input features, and then fit the target value RSRP through the relationship between the captured input features, so that we can better optimize the parameters and the accuracy of the model.

### C. FULLY CONNECTED LAYER IS USED

A large number of pairs of parameters are provided for the model and the parameter space is expanded. A sufficient number of parameters enable the model to fit and predict better results. In summary, the model designed in this paper can fit the direct relationship between the selected features and RSRP well, and can predict the RSRP value and achieve better results, the PCRR value of the model prediction result is 0.2371, and the RMSE value is 10.4855. In addition, the model uses Bi-LSTM and a fully connected network, and the fitting effect Better, high precision. But there are also shortcomings and shortcomings. Due to the large parameter space of this model, the process of seeking parameters is relatively long, training is time-consuming and resource consumption is large. At present, 5G networks are under construction in full swing, and research on wireless propagation models at home and abroad is getting deeper and deeper. Therefore, the selection of features, interference factors in the process

of wireless propagation, and model construction all need to be deepened and expanded.

## REFERENCES

- [1] M. A. M. Albreem, "5G wireless communication systems: Vision and challenges," in *Proc. Int. Conf. Comput., Commun., Control Technol. (I4CT)*, Apr. 2015, pp. 493–497.
- [2] G. Yang and J. Chen, "Research on the propagation model of 5G mobile communication system," *Mobile Commun.*, vol. 42, no. 10, pp. 28–33, Jan. 2018.
- [3] K. Li and Z. Wang, "Research on radio wave propagation model of wireless communication," *Wireless Commun. Technol.*, vol. 11, pp. 10–12, Jan. 2008.
- [4] C. Phillips and S. Raynel, "The efficacy of path loss models for fixed rural wireless links," *Passive Act. Meas.*, vol. 6579, pp. 42–51, Mar. 2011.
- [5] G. Guo, G. Sun, S. Bin, and F. Shao, "Design and analysis of field telemedicine information communication protocol based on wireless sensor network," *IEEE Access*, vol. 7, pp. 50630–50635, 2019.
- [6] I. U. Haq, A. Ullah, S. U. Khan, N. Khan, M. Y. Lee, S. Rho, and S. W. Baik, "Sequential learning-based energy consumption prediction model for residential and commercial sectors," *Mathematics*, vol. 9, no. 6, p. 605, Mar. 2021.
- [7] N. Khan, F. U. M. Ullah, I. U. Haq, S. U. Khan, M. Y. Lee, and S. W. Baik, "A novel deep learning assisted framework for renewable energy generation forecasting," *Mathematics*, vol. 9, no. 19, p. 2456, Oct. 2021.
- [8] N. Khan, I. U. Haq, F. U. M. Ullah, S. U. Khan, and M. Y. Lee, "ConvLSTM-based hybrid architecture for Batteries' state of health and power consumption forecasting," *Mathematics*, vol. 9, no. 24, p. 3326, Dec. 2021.
- [9] Z. E. Khaled, W. Ajib, and H. Mcheick, "An accurate empirical path loss model for heterogeneous fixed wireless networks below 5.8 GHz frequencies," *IEEE Access*, vol. 8, pp. 182755–182775, 2020.
- [10] R. Adeogun, T. Pedersen, C. Gustafson, and F. Tufvesson, "Polarimetric wireless indoor channel modeling based on propagation graph," *IEEE Trans. Antennas Propag.*, vol. 67, no. 10, pp. 6585–6595, Oct. 2019.
- [11] R. Adeogun, "Calibration of stochastic radio propagation models using machine learning," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 12, pp. 2538–2542, Dec. 2019.
- [12] B. Liu, Y. Xie, and J. Yuan, "A deep learning assisted node-classified redundant decoding algorithm for BCH codes," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5338–5349, Sep. 2020.
- [13] X. Zhang, N. Sood, and C. D. Sarris, "Fast radio-wave propagation modeling in tunnels with a hybrid vector parabolic equation/waveguide mode theory method," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 6540–6551, Dec. 2018.
- [14] J. H. Gao, "Application and implementation of deep learning in the physical layer of wireless transmission," *Wireless Commun.*, vol. 10, no. 1, pp. 1–12, 2020.
- [15] S. Fan, N. Xiao, and S. Dong, "A novel model to predict significant wave height based on long short-term memory network," *Ocean Eng.*, vol. 205, Jun. 2020, Art. no. 107298.
- [16] I. Al-Barazanchi, Y. Niu, H. R. Abdulshaheed, W. Hashim, A. A. Alkahtani, E. Daghghi, Z. A. Jaaz, S. A. Shawkat, and H. T. Rauf, "Proposed new framework scheme for path loss in wireless body area network," *Iraqi J. Comput. Sci. Math.*, vol. 3, no. 1, pp. 11–21, 2022.
- [17] K. Aggarwal, M. M. Mijwil, A. H. Al-Mistarehi, S. Alomari, M. Gök, A. M. Z. Alaabdin, and S. H. Abdulrhaman, "Has the future started? The current growth of artificial intelligence, machine learning, and deep learning," *Iraqi J. Comput. Sci. Math.*, vol. 3, no. 1, pp. 115–123, 2022.
- [18] R. V. Akhpahev and A. V. Andreev, "COST 231 Hata adaptation model for urban conditions in LTE networks," in *Proc. 17th Int. Conf. Young Spec. Micro/Nanotechnol. Electron Devices (EDM)*, Jun. 2016, pp. 64–66.
- [19] L. Yuan, X.-S. Yang, and B.-Z. Wang, "Prediction of time reversal channel with neural network optimized by empirical knowledge based genetic algorithm," *Acta Phys. Sinica*, vol. 68, no. 17, 2019, Art. no. 170503.
- [20] D. J. Jabbar and A. H. Al-Nakkash, "Field measurement and empirical models for radio signal propagation prediction in baghdad," in *Proc. 4th Scientific Int. Conf. Najaf (SICN)*, Apr. 2019, pp. 31–36.
- [21] N. Sood, S. Baroudi, X. Zhang, J. Liebeherr, and C. D. Sarris, "Integrating physics-based wireless propagation models and network protocol design for train communication systems," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 6635–6645, Dec. 2018.
- [22] M. L. Jakobsen, T. Pedersen, and B. H. Fleury, "Analysis of stochastic radio channels with temporal birth-death dynamics: A marked spatial point process perspective," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3761–3775, Jul. 2014.
- [23] Y. Liu, J. Cao, X. Diao, and X. Zhou, "A review on the stability of feature selection," *J. Softw.*, vol. 29, no. 9, pp. 2559–2579, 2018.
- [24] Y. Li, W. Dai, and W. Zhang, "Bearing fault feature selection method based on weighted multidimensional feature fusion," *IEEE Access*, vol. 8, pp. 19008–19025, 2020.
- [25] Y. Zhu, "Introduction to multivariate statistical analysis and its application in economic research," *China Bus. Forum*, vol. 3, pp. 234–235, Jan. 2019.
- [26] X. Deng, X. Tian, S. Chen, and C. J. Harris, "Nonlinear process fault diagnosis based on serial principal component analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 560–572, Mar. 2018.
- [27] T. Gu, "Detection of small floating targets on the sea surface based on multi-features and principal component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 5, pp. 809–813, May 2020.
- [28] Z.-C. Liu, Z.-H. Ling, and L.-R. Dai, "Statistical parametric speech synthesis using generalized distillation framework," *IEEE Signal Process. Lett.*, vol. 25, no. 5, pp. 695–699, May 2018.
- [29] M. E. Morochó-Cayamcela, M. Maier, and W. Lim, "Breaking wireless propagation environmental uncertainty with deep learning," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5075–5087, Aug. 2020.
- [30] F. Guo, R. He, and J. Dang, "Implicit discourse relation recognition via a BiLSTM-CNN architecture with dynamic chunk-based max pooling," *IEEE Access*, vol. 7, pp. 169281–169292, 2019.
- [31] B. Ganguly, A. Ghosal, A. Das, D. Das, D. Chatterjee, and D. Rakshit, "Automated detection and classification of arrhythmia from ECG signals using feature-induced long short-term memory network," *IEEE Sensors Lett.*, vol. 4, no. 8, pp. 1–4, Aug. 2020.
- [32] F. Zhou, L. Jin, and J. Dong, "Review of convolutional neural network research," *Chin. J. Comput.*, vol. 40, no. 6, pp. 1229–1251, 2017.
- [33] M. Lan, J. Wang, Y. Wu, Z.-Y. Niu, and H. Wang, "Multi-task attention-based neural networks for implicit discourse relationship representation and identification," in *Proc. EMNLP*, Sep. 2017, pp. 1299–1308.
- [34] D. Yu, L. Xingyu, D. Shuzhan, and Y. Lei, "Group emotion recognition based on global and local features," *IEEE Access*, vol. 7, pp. 111617–111624, 2019.
- [35] F. Zhang and Q. Zhou, "Ensemble detection model for profile injection attacks in collaborative recommender systems based on BP neural network," *IET Inf. Secur.*, vol. 9, no. 1, pp. 24–31, Jan. 2015.
- [36] Y. L. Yang, G. Chun Wan, and M. S. Tong, "A wireless propagation model based on artificial intelligence technology," in *Proc. IEEE Int. Conf. Comput. Electromagn. (ICCEM)*, Aug. 2020, pp. 273–274.
- [37] L. Liu, Y. Xu, and Z. Ma, "Overview of data preprocessing techniques in data mining," *J. Gansu Sci.*, vol. 1, pp. 117–119, Jan. 2003.



**YU LU YANG** received the B.S. degree in electrical engineering and intelligent control from Shanghai Maritime University, Shanghai, China, in July 2019. She is currently pursuing the M.S. degree in microelectronics science and engineering with the Department of Electronic Science and Technology, Tongji University, Shanghai. Her research interests include wireless communication, RFID technology, and structural health monitoring.



**GUO CHUN WAN** received the B.S. degree in electrical engineering from Nanchang Normal University, Nanchang, China, in 1991, and the M.S. and Ph.D. degrees in electrical engineering from Tongji University, Shanghai, China, in 2005 and 2011, respectively. He was a Lecturer with Nanchang Normal University, from 1991 to 2002, and an Associate Professor with the Jiangxi Normal University of Science and Technology, Nanchang, from 2002 to 2006.

Since 2006, he has been an Associate Professor with Tongji University. His research interests include wireless communications, information and signal processing, and intelligence of vehicles and transportations. He was a recipient of the Award for Advance of Science and Technology of Shanghai's Municipal Government, in 2016.



**MEI SONG TONG** (Senior Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, and the Ph.D. degree in electrical engineering from Arizona State University, Tempe, AZ, USA.

He is currently a Distinguished Professor and the Head of the Department of Electronic Science and Technology and the Vice Dean of the College of Microelectronics, Tongji University, Shanghai, China. He has also held an adjunct professorship at the University of Illinois at Urbana-Champaign, Urbana, IL, USA, and an honorary professorship at The University of Hong Kong, Hong Kong. He has published more than 500 papers in refereed journals and conference proceedings and coauthored six books or book chapters. His research interests include electromagnetic field theory, antenna theory and design, simulation and design of RF/microwave circuits and devices, interconnect and packaging analysis, inverse electromagnetic scattering for imaging, and computational electromagnetics. He is a fellow of the Electromagnetics Academy and the Japan Society for the Promotion of Science (JSPS) and a Full Member (Commission B) of the USNC/URSI. He has been the Chair of Shanghai Chapter (since 2014) and the Chair of SIGHT Committee (in 2018) in IEEE Antennas and Propagation Society. He also frequently served as the

session organizer/chair, the technical program committee member/chair, and the general chair for some prestigious international conferences. He was a recipient of the Visiting Professorship Award from Kyoto University, Japan, in 2012, and from The University of Hong Kong, Hong Kong, in 2013. He advised and coauthored 12 papers that received the Best Student Paper Award from different international conferences. He was a recipient of the Travel Fellowship Award of USNC/URSI for the 31th General Assembly and Scientific Symposium (GASS), in 2014; the Advance Award of Science and Technology of Shanghai Municipal Government, in 2015; the Fellowship Award of JSPS, in 2016; the Innovation Award of Universities' Achievements of Ministry of Education of China, in 2017; the Innovation Achievement Award of Industry-Academia-Research Collaboration of China, in 2019; and the "Jinqiao" Award of Technology Market Association of China, in 2020. He has served as an Associate Editor or Guest Editor for several well-known international journals, including *IEEE Antennas and Propagation Magazine*, *IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION*, *IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY*, *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, *Progress in Electromagnetics Research*, and *Journal of Electromagnetic Waves and Applications*. In 2018, he was selected as a Distinguished Lecturer (DL) of IEEE Antennas and Propagation Society, from 2019 to 2021.

...