ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

Optimization of Physical Layer Security for Channel Estimation based on Deep Learning

Yuzhou Ning

School of Anhui University of Finance and Economics, Anhui, 233000, China

*3357236055@qq.com

Abstract

Wireless channel estimation is the basis of communication security. In the communication system, due to the time-varying channel, there is a certain error between the channel state information(CSI) of the channel estimation and the real CSI, In this paper, considering the condition that the CSI of eavesdropping channel is unidentified, the legitimate channel is estimated by pilot frequency. the correlation between the main channel and eavesdropping channel is used to estimate the CSI of eavesdropping channel, introducing Deep Neural Network(DNN) to extract channel features. At the same time, improving the algorithm by adding the pre-training process to avoid the error caused by random initialization. After obtaining the channel characteristics, the channel security capacity is calculated to evaluate the performance of the physical layer security. The simulation results show that the scheme can improve the accuracy of channel estimation, enhance the channel secrecy capacity.

Keywords

Time-varying Channel Estimation; Deep Learning Algorithm; Physical Layer Security.

1. Introduction

Nowadays, the demand of wireless communication for quality of service and data transmission rate is growing. The number of devices in the sixth generation mobile communication (6G) system will reach hundreds of billions, and research on key technologies such as large-scale multiple input multiple output (MIMO), ultra dense network (UDN), millimeter wave (mmWave) communication has emerged. In this trend, the security of communication has attracted more and more attention. Privacy and security issues in wireless channels play an increasingly important role in wireless networks. How to improve the security performance of communication systems has become a research hotspot. As a supplement to the upper encryption technology, the physical layer security technology realizes information confidentiality and identity authentication by exploring the randomness of the physical layer transmission media. The factors of measuring physical layer security are studied in Ref[1]. Researching physical layer security by NOMA technology in Ref[2]. Analysing physical layer security in OFDM system model in Ref[3]. Researching physical layer security through the differences among channels in Ref[4].The research in the field of physical layer security can be divided into two major categories, namely, the methods to improve the security capacity of the physical layer from the theoretical point of view and the system strategy to realize the secure communication of the physical layer from the specific technology. Introducing deep learning into physical layer security in Ref[5-7].

The basis for evaluating the security performance of the physical layer is to estimate the channel. Using channel estimation to evaluate physical layer security in Ref[8-9]. However, due to channel noise, channel time-varying and other reasons, the channel estimation results may have errors.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

Whether the CSI of the channel can be obtained, whether there is error in the CSI, and the size of the error have an important impact on the design of the wireless communication system transmission scheme and the system performance. Estimating channel under the condition of unknown eavesdropping channel state in Ref[10-11]. This is because the ideal CSI can not only provide a strong transmission power focusing capability in the downlink of the base station, but also provide an ideal uplink power receiving scheme for it. At the same time, it can also more effectively prevent eavesdropping channels from obtaining information and improve the confidentiality of the system. In the research of physical layer security, it is generally assumed that the CSI of legitimate channels is accurately known, while the CSI of eavesdropping channels can be divided into known, partially known or unknown cases. Different security schemes are required for different situations. On this premise, this paper constructs a model related to the main stealing channel. After the estimation of the legitimate channel, the relationship between the main stealing channels is used to estimate the channel characteristics of the eavesdropping channel, and then the channel capacity of the main stealing channel is calculated separately to obtain the security channel capacity, so as to evaluate the security performance of the physical layer.

OFDM system is widely used in time-varying channel estimation due to its strong bandwidth expansion, high spectrum utilization, strong anti-interference ability, and good adaptability to frequency selective fading. In recent years, deep learning has made great progress in channel estimation of wireless communication. Deep neural network is widely used in time-varying channel estimation because of its strong nonlinear mapping ability and data driven characteristics. Researching how to introduce deep learning network into OFDM system in Ref[12-13]. Instead of the traditional channel estimation method that assumes that the channel is a channel estimation method that meets the fixed change rule, the deep learning method can use multiple stack layers to learn the hidden nonlinear rule, so it has achieved good estimation performance. The OFDM time-varying channel estimation based on deep learning adopted in this paper increases the pre-training process and further avoids the error caused by random initialization.

2. Channel Model

2.1 OFDM Time-varying Channel Model

The downlink multi-user system model studied in this paper is shown in Figure 1. The system consists of a base station (Alice) and K users. Alice sends information to K users with a time-division manner in turn, and the information which is sent to each user needs to be kept confidential. Without losing generality, the target user of the current time information is called Bob, and other users are regarded as eavesdroppers (Eves).

Suppose that the source node (Alice), target node (Bob) and eavesdropping node (Eve) are all equipped with a single antenna. the main channel from Alice to Bob and the eavesdropping channel from Alice to Eve are related. h_d and h_e represent the channels from Alice to Bob and Alice to Eve respectively, then the correlation between the main channel and stealing channel can be described by their joint probability density function, which is represent.

$$f_{h_e^2,h_d^2}(x,y)=\frac{I_0\left(\frac{2}{1-\rho}\sqrt{\frac{\rho xy}{\alpha_e^2\alpha_d^2}}\right)}{(1-\rho)\alpha_e^2\alpha_d^2}e^{-\frac{\frac{x}{\alpha_e^2}+\frac{y}{\alpha_d^2}}{1-\rho}} \tag{1}$$

where, α_d^2 and α_e^2 gain variance of h_d and h_e Channel. $I_0(x)$ Is the first kind of zero order modified Bessel function, namely.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

$$I_0(x)=\sum_{k=0}^{\infty}\left(\frac{x^k}{2^k k!}\right)^2 \tag{2}$$

The power correlation coefficient between the main stealing channels is $\rho=\mathrm{cov}(h_d^2,h_e^2)/\sqrt{\mathrm{var}(h_d^2)\,\mathrm{var}(h_e^2)}$. In the wireless environment, the power correlation coefficient between the main stealing channels, as a statistical variable that remains stable for a long period of time, can be obtained by measuring the environment in advance. without doubt, the correlation between the main stealing channels is $\rho\in[0,1)$. The correlation between the main stealing channels depends not only on the distance between Bob and Eve, but also on the abundance of surrounding scatters. Specifically, $\rho=0$ corresponding the independence of h_d and h_e, but under extreme conditions $\rho=1$ express the full relevant of h_d and h_e.

From many relevant time intervals, h_d^2 is a two degree of freedom Random variable called χ^2, whose probability density can be expressed as:

$$f_{h_d^2}(y)=\begin{cases}\dfrac{1}{\alpha_d^2}\mathrm{e}^{-\frac{y}{\alpha_d^2}}, & y>0\\ 0, & y<0\end{cases} \tag{3}$$

In any coherent time interval, h_d remains unchanged and can be obtained by Alice through pilot training estimation at Bob and CSI feedback. For h_e, since Eve generally does not feedback its CSI, Alice cannot obtain its estimated value. Therefore h_e is an unknown variable for Alice. Since the main channel is determined and there is a certain correlation between the eavesdropping channel and the main channel, it is not accurate to use Rayleigh distribution to describe the eavesdropping channel. Instead, the conditional probability density should be used to describe the probability distribution of eavesdropping channels.

$$f_{h_e^2|h_d^2}(x\,|\,y)=\frac{f_{h_e^2|h_d^2}(x,y)}{f_{h_d^2}(y)} \tag{4}$$

So we can get:

$$f_{h_e^2|h_d^2}(x\,|\,y)=\begin{cases}\dfrac{\mathrm{e}^{-\frac{b^2+x}{2a^2}}I_0\left(\frac{b}{a^2}\sqrt{x}\right)}{2a^2}\\ 0, x\le 0\end{cases} \tag{5}$$

where, $a^2=\dfrac{(1-\rho)\alpha_e^2}{2}$ $b=\sqrt{\dfrac{y\rho\alpha_e^2}{\alpha_d^2}}$.

From the probability distribution of eavesdropping channels, h_e^2 is a noncenter with 2 degrees of freedom Random variable called χ^2, so eavesdropping channel can be modeled as:

$$h_e=\bar{h}_e+h_e \tag{6}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

where, $\bar{h}_e = \frac{\alpha_e \sqrt{\rho} h_d e^{j\theta_e}}{\alpha_d}$ $h_e = \alpha_e \sqrt{1-\rho} z_e$.

θ_e and $z_e \sim \mathbb{CN}(0,1)$ are phase variable and zero mean unit variance complex Gaussian variable respectively. Obviously, the above formula represents the correlation between the main channel and stealing channels. Therefore, it can be found that h_e is a complex Gaussian variable, whose mean and variance are $\frac{\alpha_e \sqrt{\rho} h_d e^{j\theta_e}}{\alpha_d}$ and $\alpha_e \sqrt{1-\rho}$.

Alice transmits information to each user in turn. Before the information transmission starts, Alice sends pilot signals to the target users through the OFDM system transmitter(X_m^p) to perform channel estimation, and transmit the signals received by all sub channels(Y_m) to send back to Alice through the receiver.

OFDM system is divided into transmitter and receiver. At the transmitting end, the data is subject to cyclic redundancy check (CRC), channel coding, punching, interleaving, constellation mapping, pilot insertion and subcarrier scrambling, then reverse Fourier transform (IFFT) is performed, and cyclic prefix (CP) is added for OFDM symbol transmission. The receiving end is the inverse process of the transmitting end. After removing the CP, fast Fourier transform (FFT) is performed, and the transmitted data can be recovered through equalization, demodulation and other operations in the channel estimation link.

After the OFDM system passes through the time-varying channel and noise interference, the frequency form of the receiver signal can be expressed as:

$$\begin{aligned} &Y(\mathrm{m}) = H(m)X(m) + N_w(m) \\ &\mathrm{m} = 0,1,\cdots,\mathrm{N}-1 \end{aligned} \tag{7}$$

where, N is the number of an OFDM symbol subcarriers, and m is the serial number of the subcarrier. H_m is the channel response at the m-th frequency point, $X(m)$ is the symbol sent, $N_w(m)$ is additive white Gaussian noise (AWGN).

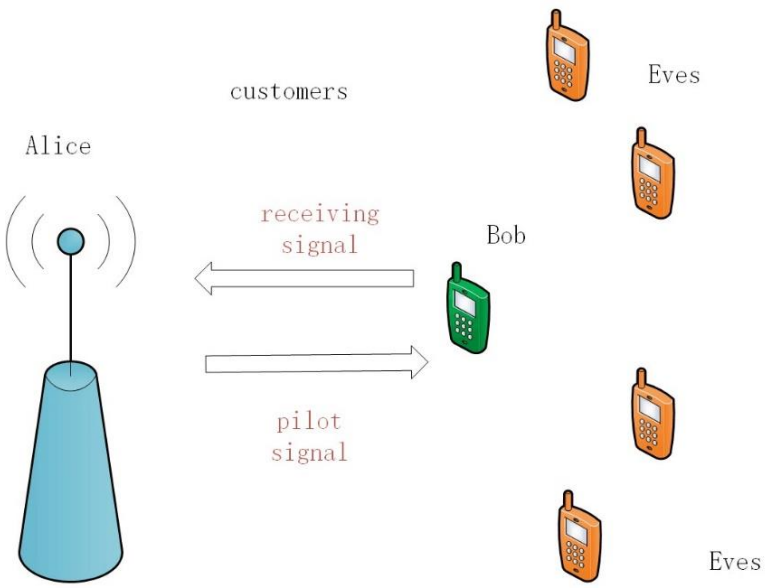

Figure 1. Time varying channel model

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

3. Algorithm Design

3.1 Basic Structure of DNN Model

The neural network model used in this paper consists of an input layer, a 3-layer hidden layer and an output layer. The number of neurons in each layer was 32,30,60,120,224. The number of OFDM subcarriers used in this paper is 128, of which the number of pilot subcarriers is 16, and the number of data subcarriers is 112. Because the real and imaginary parts of the complex number need to be separated, the number of neurons in the input layer of the neural network is set to 32, and the number of neurons in the output layer is set to 224. The number of neural elements in the hidden layer is enhanced layer by layer to increase the DNN nonlinear fitting ability. The basic learning process of DNN channel estimation algorithm consists of two processes: forward propagation of signal and back propagation of error.

During forward propagation, the input data received by the input layer is the frequency response at the pilot position estimated by LS H_p. Since the pilot frequency response is complex, the real part needs to be extracted H_{p_re} and imaginary part H_{p_im}, combined to get the input vector x:

$$x = [H_{p\_re}^T, H_{p\_im}^T]^T \tag{8}$$

where, $H_{p\_re}^T$ and $H_{p\_im}^T$ are N_p Dimension vector.

Then the input vector of the first hidden layer is obtained by the following formula $u^{(1)}$:

$$u^{(1)} = W^{(1)}x + b^{(1)} \tag{9}$$

where, $u^{(1)}$ by n_1 Dimension line vector, n_1 is the number of neurons in the first hidden layer; $n_1 \times 2N_p$ matrix $W^{(1)}$ is the weight matrix between the input layer and the first hidden layer, $b^{(1)}$ by n_1 Dimension offset vector.

Then, under the action of nonlinear activation function, the output vector of the first hidden layer is obtained $x^{(1)} = f(\mathrm{u}^{(1)})$, activation function $f(\cdot)$ is a sigmoid function, which can be expressed as $f(x) = \frac{1}{1+e^{-x}}$ Similarly, the output vector of the whole neural network model is:

$$y = f(W^{(l-1)}(\cdots f(W^{(1)}x + b)\cdots) + b^{(l-1)}) \tag{10}$$

where, l is the number of layers of the neural network; y is $2N_d$ Dimension line vector, N_d is the number of data subcarriers in an OFDM symbol.

The frequency response at the data symbol predicted by the neural network can be obtained by recombining the output results of the output layer into the form of complex numbers.

Since the channel frequency response has positive and negative values, and the output range of the Sigmoid function is [0,1], the activation function of the output layer selects the Tanh function, whose output range is [- 1,1], which can be expressed as:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{11}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

The loss function can be expressed as:

$$L = \frac{1}{2m}\sum_{j=1}^{m} \| y_j - z_j \|^2 \tag{12}$$

where, m is the number of neurons in the output layer,$y_j$is the value of the jth neuron in the output layer, z_j is the corresponding output layer label value.The weight value to be adjusted each time can be expressed as:

$$\triangle w_{ij} = -\eta \frac{\partial L}{\partial w_{ij}} \tag{13}$$

where, w_{ij} is the weight value between the i-th hidden layer neuron and the jth output layer neuron, η is the learning rate of neural network.According to the chain rule, the above formula can be rewritten as:

$$\triangle w_{ij} = -\eta \frac{\partial L}{\partial y_j} \cdot \frac{\partial y_j}{\partial u_j^{(2)}} \cdot \frac{\partial u_j^{(2)}}{\partial w_{ij}} = -\eta (y_j - z_j) f^{'}(u_j^{(2)} + b_j^{(2)}) x_i^{(1)} \tag{14}$$

where, $u_j^{(2)}$ is the offset term of the jth neuron in the output layer; $x_i^{(1)}$ is the output value of the i-th neuron of the hidden layer.Similarly, we can get the expression of offset term update as:

$$\triangle b_j^{(2)} = -\eta (y_j - z_j) f^{'}(u_j^{(2)} + b_j^{(2)}) x_i^{(1)} \tag{15}$$

The update method of weight and offset items between hidden layer and output layer is the same as above. Each iteration updates the value of the weight and offset term according to the gradient descent algorithm until the set number of iterations is reached.

3.2 Time Varying Channel Estimation Algorithm based on Deep Learning

In the process of channel estimation, a depth learning assisted time-varying channel estimation algorithm can be used, which includes two stages: offline training and online estimation. When constructing samples, collect the received signals of all subchannels Y_m, transmit pilot X_m^p and channel estimation at historical time $H_{m-1}$to construct the input samples of the DNN network and extract the channel characteristics, which is:

$$u_{tr} = \{(x_{tr}^{(1)}, g_{tr}^{(1)}), (x_{tr}^{(2)}, g_{tr}^{(2)}), \cdots (x_{tr}^{(v)}, g_{tr}^{(v)}), \cdots, (x_{tr}^{(V)}, g_{tr}^{(V)})\} \tag{16}$$

where, V represents the number of samples, $x_{tr}^{(1)}$, $g_{tr}^{(1)}$ respectively represent the input and output samples of the network, of which the v input sample:

$$x_{tr}^{(v)} = [Y_m, X_m^p, H_{m-1}]^T \tag{17}$$

The output sample is:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

$$g_{tr}^{(v)} = H_m^T \tag{18}$$

where, H_m indicates the current real channel status.

Then the input and output samples are fed into the DNN network for offline learning, where the input vector is x_m:

$$x_m = [H_{x\_re}^T, H_{x\_im}^T]^T \tag{19}$$

Input vector from the first hidden layer $u_m^{(1)}$:

$$u_m^{(1)} = W_m^{(1)} x_m + b_m^{(1)} \tag{20}$$

Then, under the action of nonlinear activation function, the output vector of the first hidden layer is obtained $x_m^{(1)} = f(u_m^{(1)})$ by analogy, the output vector of the entire neural network, that is, the initialization parameters θ.

In the online test phase, based on the trained DNN network, the channel estimation matrix at the current time can be obtained:

$$H_m = \Phi(X_m) \tag{21}$$

where, $\Phi(\cdot)$ represents the nonlinear mapping of the DNN network.

Because this method is based on the trained neural network, the expected channel estimation matrix can be obtained by inputting the target samples, making full use of the data-driven characteristics of the neural network, which has high practicability in practical applications, and its performance has considerable advantages over the traditional LS estimation.

3.3 Algorithm Improvement

In order to avoid the error caused by random initialization, this paper proposes a time-varying channel estimation algorithm based on deep neural network DNN. Compared with the above method, this method adds a training process, before training, the network has learned the channel characteristics of time-varying channels, and then trains the network again, greatly accelerating the convergence speed of the network.

This method is divided into three steps: pre-training, training and testing. The DNN network with three hidden layers adopts the same structure in different stages.

In the pre training phase, the network inputs used include sending signals X_m (assume known in the pre training phase), receive the signal Y_m, channel estimation at the previous time H_{m-1}, and channel estimation at the current time obtained by LS algorithm $H_{m,L,S}$:

$$u_{pre} = \{(x_{pre}^{(1)}, g_{pre}^{(1)}), (x_{pre}^{(2)}, g_{pre}^{(2)}), \cdots (x_{pre}^{(v)}, g_{pre}^{(v)}), \cdots, (x_{pre}^{(V)}, g_{pre}^{(V)})\} \tag{22}$$

where, $x_{pre}^{(v)}$ represents the v th input sample, which can be expressed as:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

$$x_{pre}^{(v)} = [Y_m, X_m^p, H_{m-1}, H_{m,LS}^p]^T \tag{23}$$

In the training phase, the network inputs used include sending signals X_m^p (only the pilot is known), receive the signal Y_m, channel estimation at the previous time H_{m-1}, and channel estimation at the current time obtained by LS algorithm $H_{m,L,S}$ (only the information at the pilot position is known, and the data subcarrier is disposed of as 0), which is:

$$u_{tr} = \{(x_{tr}^{(1)}, g_{tr}^{(1)}), (x_{tr}^{(2)}, g_{tr}^{(2)}), \cdots (x_{tr}^{(v)}, g_{tr}^{(v)}), \cdots, (x_{tr}^{(V)}, g_{tr}^{(V)})\} \tag{24}$$

The v th input sample can be expressed as:

$$x_{tr}^{(v)} = [Y_m, X_m^p, H_{m-1}, H_{m,LS}^p]^T \tag{25}$$

After network training, initialization parameters more suitable for current channel characteristics can be obtained θ_1.

After getting the parameters θ_1, the pilot can be transmitted with the same type of input data as in the training phase X_m^p, receive signal Y_m, channel estimation at the previous time H_{m-1}, and channel estimation at the current time obtained by LS algorithm $H_{m,LS}^p$, together as the input of the DNN network, the desired channel value can be obtained:

$$H_m = \Phi(x_m) \tag{26}$$

This method uses the pre-training method to effectively avoid the impact of random initialization and improve the accuracy of channel estimation. After the channel estimation matrix is obtained through channel estimation. the channel capacity is:

$$C_d(H) = \frac{1}{2}\log_2\left\{\det[I_{n_R} + \frac{HR_{xx}H^H}{\sigma^2}]\right\} \tag{27}$$

Because there is correlation between the legitimate channel and the eavesdropping channel, the estimation matrix of the eavesdropping channel can also be estimated according to the estimation matrix of the legitimate channel, thus obtaining the channel capacity of the eavesdropping channel $C_e(H)$, the channel security capacity $C(H)$ can be defined as:

$$C(H) = C_d(H) - C_e(H) \tag{28}$$

4. Simulation Test

4.1 System Simulation

During the simulation test, assume that the total number of samples (V) is 100, OFDM time-varying channel is used, and the noise type is AWGN. Among them, the number of points in each layer of the neural network is set to 32,30,60,120,224, the Sigmoid function is selected as the activation function of the hidden layer, the Tanh function is selected as the activation function of the output layer, the L2

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

function is selected as the loss function, the learning rate is set to 0.5, the number of iterations is set to 200, and the number of Monte Carlo simulations is set to 1000.

In the simulation process, SNRs of 10DB and 20DB are selected, and the channel security capacity changes with the increase of transmission power when there is no deep learning algorithm and there is deep learning algorithm are plotted respectively for comparison.

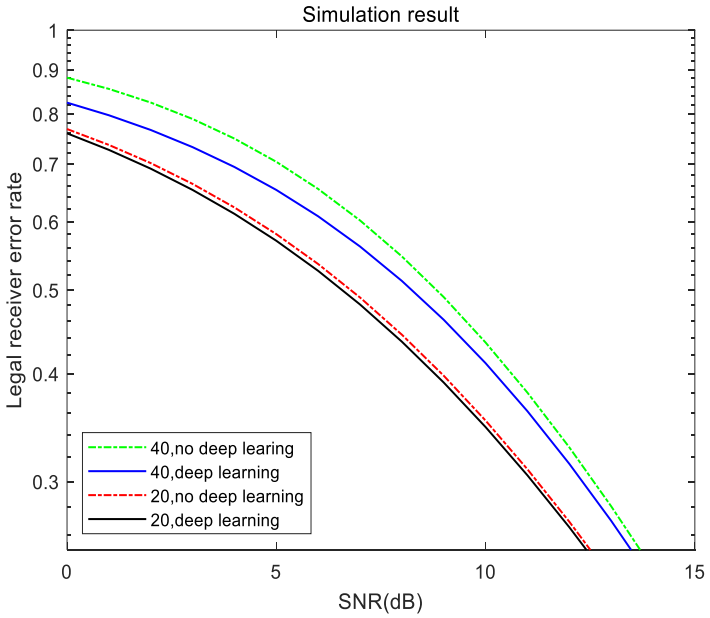

Figure 2. Simulation result of Legal receiver error rate

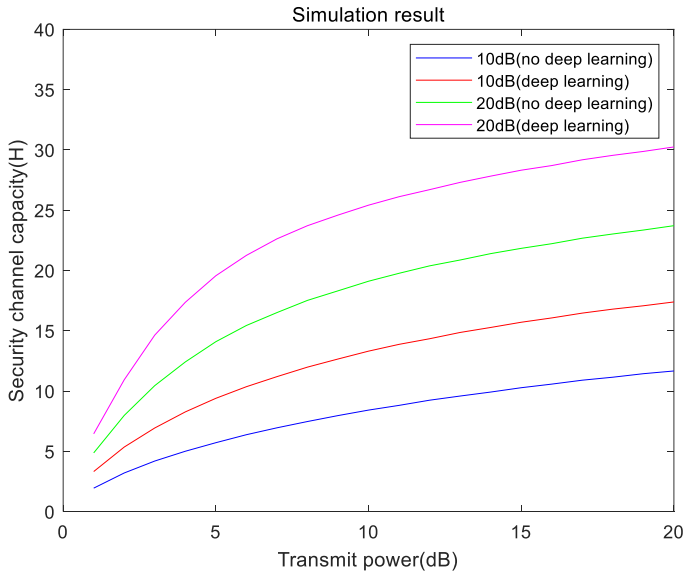

Figure 3. Simulation result of Security channel capacity

The simulation results show that the capacity of the secure channel changes with the transmission power under different SNR conditions and whether the deep learning algorithm is used. Longitudinal comparison shows that the secure channel capacity from the best to the worst is 20dB (machine learning), 20dB (no machine learning), 10dB (machine learning), 10dB (no machine learning). With the continuous increase of transmission power, the secure channel capacity will also increase. At the same time, the introduction of deep learning algorithm to estimate the channels of legitimate users

ISSN: 2414-1895 DOI: 10.6919/ICJE.202307_9(7).0017

can effectively improve the accuracy of channel estimation, The feasibility of the algorithm is proved by increasing the capacity of the secure channel and effectively improving the security performance of the physical layer.

5. Conclusion

This paper considers the correlation between the main stealing channels, introduces a deep learning algorithm, estimates the characteristics of the eavesdropping channel by estimating the legitimate channel, calculates the channel security capacity, and evaluates the security performance of the physical layer. At the same time, the deep learning algorithm is improved by adding the pre training process to estimate the characteristics of the time-varying channel, avoid the error caused by random initialization, and further improve the accuracy of channel estimation. From the simulation results, it can be seen that the accuracy of channel estimation using the deep learning algorithm is significantly greater than that without the deep learning algorithm. At the same time, the use of the deep learning algorithm can significantly increase the channel security capacity and effectively improve the security performance of the physical layer. It proves the feasibility of this algorithm in channel estimation.

References

[1] Xu Chaofan Research on physical layer security authentication based on deep learning [D]. Nanjing University of Posts and Telecommunications, 2022.

[2] Jia Fan Research on security technology of NOMA physical layer for internal eavesdropping users [D]. Xi'an University of Electronic Science and Technology, 2021.

[3] Liu Wenfei Research on physical layer security method based on OFDM system [D]. Dalian University of Technology, 2018.

[4] Wang Jialing Research on Physical Layer Security Implementation Technology Based on Channel Difference [D]. University of Electronic Science and Technology of China, 2022.

[5] Zhu Yuxuan Research on Wireless Communication Channel Estimation Algorithm Based on Machine Learning [D]. University of Electronic Science and Technology of China, 2022.

[6] Li Hong Research on optimization of channel estimation algorithm based on machine learning [D]. University of Electronic Science and Technology of China, 2020.

[7] Li Hang Research on Physical Layer Security Authentication Technology Based on Machine Learning [D]. University of Electronic Science and Technology of China, 2019.

[8] Li Huan Research on Physical Layer Security Technology Based on Neural Network [D]. Chongqing University of Posts and Telecommunications, 2019.

[9] Xi Chenjing, Gao Yuanyuan, Sha Nan. Impact of channel estimation error on physical layer security encryption scheme [J]. Computer Engineering, 2020,46 (06): 122-129.

[10] Zhuo Guofeng, Du Shengdong, Lin Shengbin. Physical layer secure transmission scheme based on segmented pilot transmission and artificial noise [J]. Electronic Technology Application, 2017,43 (01): 129-132.

[11] Liu Jue, Cheng Kaixin, Yang Weiwei. Research on physical layer security technology under intelligent eavesdropping attack [J]. Information Network Security, 2023,23 (02): 45-53.

[12] Zeng Lingqing, Cai Xiaoxia, Chen Hong, Zhu Wenli. Power allocation scheme for physical layer security multiple antenna selection based on channel estimation [J]. Computer Application Research, 2017,34 (03): 879-882.

[13] Xian Xiaoxiao, Chen Di, Gao Hui, Cao Ruohan, Bie Zhisong. Machine learning assisted efficient beam training and channel estimation method for RIS mmWave system [J]. Signal Processing, 2022,38 (08): 1610-1619.