Data Authentication Based on Discrete Random Convolutional Neural Network

Fei-yi XIE\textsuperscript{1}, Ai-dong XU\textsuperscript{2}, Yu-nan ZHANG\textsuperscript{2}, Yi-xin JIANG\textsuperscript{2} and Hong WEN\textsuperscript{1}.

\textsuperscript{1}University of Electronic Science and Technology of China (UESTC), Chengdu, China
\textsuperscript{2}EPRI, China Southern Power Grid Co. Ltd., Guangzhou 510080, China

*Corresponding author

Keywords: Convolutional neural network, Data authentication, Channel matrix.

Abstract. Data authentication is an efficient way to enhance the security of data transmission, especially for the resource-limited terminal devices in edge computing system. Authentication through physical layer channel characteristics is a lightweight method compared with that based on cryptography. With the popularity of deep learning, this paper proposed a discrete random convolutional neural network applicable for the channel matrix, which extracts more feature information in convolution step without increasing the required memory by discrete random movement of the convolutional kernel. It’s suitable for protecting the data transmission between micro terminals in edge computing.

Introduction

There are a massive terminals and a tremendous amount of data packages to be sent and received in an edge computing system [1]. Package forgery and vulnerability attacks through data packages are common ways of network attacks, such as attacks against the shortcomings of package filtering and application gateway technology, and Linux kernel vulnerability attacks related to data packet sockets (CVE-2017-7308) [2].

The access authentication of data package is an important method to ensure the security of data transmission in edge computing [3]. The common traditional means of defense are the composite firewall and the access authentication of sending devices, but these methods are based on cryptography, which brings heavy computing load to the low-end terminals [4]. Data authentication of physical layer channel characteristics is to judge sender’s identity information by comparing the similarity of channel information between consecutive frames, which is fast and efficient, and is very suitable for resource-limited terminals in edge computing system [6].

The traditional method of channel characteristics based data authentication uses threshold to determine the validity, of which the identification accuracy is low and unstable [6]. Machine learning and the emerging deep learning can effectively improve the identification accuracy by training a tremendous amount of samples to get classifiers [10].

With the popularity of multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM) in edge computing transmission, one-dimensional (1D) channel estimation vectors has become two-dimensional (2D) matrix samples of receiving antenna and channel estimation, i.e. channel matrix.\textsuperscript{[11]} Some techniques applicable to image processing, such as convolutional neural network (CNN), can be applied to improve the identification accuracy further [13]. However, the number of receiving antennas is far less than the number of elements in a channel estimation vector, and the receiving antennas have independence (adjacent receiving antennas do not have higher correlation), which makes CNN cannot be directly applied to these samples to get better identification results.

In this paper, a data authentication method for edge computing based on discrete random CNN is proposed. This method discretizes the convolution kernel of CNN, which can randomly select the receiving antennas and extract their features, so that the CNN can obtain enough information on the
receiving antenna dimension. By this way, the identification effect of classifier constructed by CNN in MIMO channel matrix is improved, as well as the identification accuracy.

Background

Channel Estimation

Channel Estimation is the key step in data authentication based on physical layer channel characteristics. In general, the channel can be estimated by using a preamble or pilot symbols known to both transmitter and receiver. Though greatly affected by noise, least square (LS) method is widely used in channel estimation because of its simplicity [14].

In MIMO-OFDM system, all subcarriers are orthogonal. The training symbols for $N$ subcarriers can be represented by the following diagonal matrix:

$$
Y = HX + Z
$$

(2)

where $H$ is a channel vector with the channel gain $H[k]$ and $Z$ is a noise vector.

Channel estimation method based on LS minimizes the following cost function to find channel estimates $\hat{H}$:

$$
J(\hat{H}) = \|Y - X\hat{H}\|^2 = (Y - X\hat{H})^H(Y - X\hat{H}) = Y^H Y - Y^H X\hat{H} - \hat{H}^H X^H Y + \hat{H}^H X^H X\hat{H}
$$

(3)

Set the derivative of the function with respect to $\hat{H}$ to 0:

$$
\frac{\partial J(\hat{H})}{\partial \hat{H}} = -2(Y^H Y)^T + 2(X^H X\hat{H})^T = 0 \iff X^H Y = X^H X\hat{H} \iff \hat{H}_{LS} = (X^H X)^{-1} X^H Y = X^{-1} Y
$$

(4)

Therefore, the solution to the LS channel estimation is:

$$
\hat{H}_{LS} = X^{-1} Y
$$

(5)

Convolutional Neural Network

Convolutional neural network is a kind of feedforward neural network with deep structure and convolution computation, which is one of the representative algorithms of deep learning. It has a forward propagation and a backward propagation, while the forward propagation is divided into input layer, convolutional (conv) layer, activation (mostly ReLU) layer, pooling layer and full connection (FC) layer.

The function of conv layer is to extract features from input data, including several convolutional kernels, which are generally matrices. When convolutional kernels work, they regularly sweep over input features and convolute with them, i.e. sum by matrix element multiplication and superimpose deviations with input features. The area of single convolution is called receptive field [15].
Methods

Because the number of receiving antennas is far less than the number of elements in a single channel estimation vector, the longitudinal dimension of the channel matrix is much less than the transverse dimension. Therefore, the convolutional neural network cannot extract enough feature information in this dimension. Meanwhile, the receiving antennas are on the equal status, which means the adjacent rows in the channel matrix do not have a higher correlation, the way the kernel of traditional CNN moves cannot reflect that characteristic.

A novel form of convolution kernel and moving rule are proposed, of which the kernel is discrete and the moving is random. Algorithm 1 describes the generation and movement rule of three line convolution kernel in the longitudinal direction, i.e. the receiving antenna dimension.

**Algorithm 1 The generation and movement rules of a 3 line kernel**

Generation:

\[ k_n = [I_{s_n}, I_{s_{n+1}}, I_{s_{n+2}}] ; \]

Step = \( N \) ;

lines = \( L \) ;

Initialization:

\[ n = 1; \quad s_n, s_{n+1}, s_{n+2} = 1, 2, 3; \]

\[ k_1 = [I_{1}, I_{2}, I_{3}]; \]

while \( n < N \) do

\[ k_{n+1} = [I_{s_{n+1}}, I_{s_{n+2}}, I_{s_{n+3}}] ; \quad s_{n+3} = random(1, L), s_{n+3} \neq s_n, s_{n+1}, s_{n+2}; \]

if \( k_{n+1} \neq k_n \) then

\[ n' = 1, 2, \ldots, n \]

back to while;

else

do convolution and output the result: \( R_n \); \n
\[ n = n + 1 \]

end if

end while

Convolution result: \( R = [R_1; R_2; \ldots; R_N] \)

The random discrete movement rule allows the receptive field of every convolution step to be any three lines, which greatly increases the feature information extracted in the antenna dimension. Meanwhile, the scale and number of convolutional kernels are not increased, thus the number of parameters in CNN is unchanged, which will not bring overfitting. In addition, because the random movement can output a bigger scale than the input sample, padding is not required in conv layer.

In estimated channel vector dimension, because the channel vectors have continuity, traditional translation by stride is more applicable.

Experiment and Discussion

In this experiment, an edge computing system using MIMO-OFDM with eight output antennas and eight input antennas was set indoors. The channel matrix estimated by the receiver is \( 8 \times 8 \times 128 \), of which the waveform is shown in Figure 1.
The scale of the convolutional kernel is set as $3 \times 32$, and the number of kernels are 20.

**Space Complexity**

Represents the number of parameters of the model, which is embodied in the volume of the model itself. If the kernel is square, the space complexity

$$Space \sim O\left(\sum_{l=1}^{D} K_l^2 \cdot C_{l-1} \cdot C_l\right)$$

where $K_l^2$ is the scale of the kernel, $D$ is the depth of the network and $C_l$ is the number of kernels in layer $l$. In this single layer CNN, $C_0=1$, $C_1=20$, the scale of the rectangle kernel is $3 \times 32$. So the space complexity of traditional kernel and the discrete random kernel is the same.

$$Space(trd3) = Space(dis-rand) \sim O(3 \times 32 \times 1 \times 20) = O(1920)$$

(7)

This discrete random kernel can be regarded as a $8 \times 32$ kernel with dropout, of which the space complexity is much higher.

$$Space(trd8) \sim O(8 \times 32 \times 1 \times 20) = O(5120)$$

(8)

**Time Complexity**

Time complexity is the computational effort required to implement the algorithm, which qualitatively describes the running time of the algorithm.

$$Time \sim O\left(\sum_{l=1}^{D} M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l\right)$$

where $M_l^2$ is the scale of the output feature map through convolution.

For discrete random convolution kernels, the maximum scale of the output feature map is

$$S(M_l) = A_l^3 \times 8 \times 128 = 344064$$

(10)

However, generally, it’s unnecessary to have such large output. In this experiment, set the antenna dimension to be 28.

Its time complexity is

$$Time(dis-rand) \sim O(28 \times 8 \times 128 \times 1920) \approx O\left(5.5 \times 10^7\right)$$

(11)

Correspondingly, the time complexity of traditional $3 \times 32$ (stride = 1 with same padding) and $8 \times 32$ kernel is
Summary

The space complexity of discrete random convolutional kernels is the same as that of traditional convolutional kernels of the same scale, and much lower than that of the traditional convolutional kernel with equal computational power, which can effectively reduce the requirement for running memory while avoiding over-fitting at the same time.

The time complexity of discrete random convolutional kernels is higher than that of traditional convolutional kernels of the same scale, while still much lower than that of the traditional convolutional kernel with equal computational power, which means that it can use less time to achieve the same identification result.

Conclusion

This paper proposed a discrete random convolution kernel of CNN. Its moving mode can randomly select discrete receiving antennas, so that the convolution neural network can obtain enough sample features in the dimension of receiving antenna. This method can efficiently improve the effect of feature extraction through convolution without increasing parameters, while controlling the running time in an acceptable range. It’s of positive significance to ensure data transmission between resource-limited terminals in edge computing system.

Acknowledgement

This research is supported by the National Key R&D Program of China (2018YFB0904900, 2018YFB0904905).

References

[1] R. M. Gerdes, M. Mina, S. F. Russell and T. E. Daniels, Physical-Layer Identification of Wired Ethernet Devices, IEEE Transactions on Information Forensics and Security, vol. 7, no. 4, pp. 1339-1353, Aug. 2012.

[2] N. K. Ratha, J. H. Connell and S. Pankanti, Big Data approach to biometric-based identity analytics, IBM Journal of Research and Development, vol. 59, no. 2/3, pp. 4:1-4:11, March-May 2015.

[3] R. R. Parmar, S. Roy, D. Bhattacharyya, S. K. Bandyopadhyay and T. Kim, Large-Scale Encryption in the Hadoop Environment: Challenges and Solutions, IEEE Access, vol. 5, pp. 7156-7163, 2017.

[4] H. Wen, S. Li, X. Zhu and L. Zhou, “A Framework of the PHY-layer Approach to Defense against Security Threats in Cognitive Radio Networks”, IEEE Networks, 2013, vol. 27, no.3, pp. 34-39.

[5] Y. Chen, H. Wen, H. Song, S. Chen, F. Xie, Q. Yang, L. Hu, Lightweight one-time password authentication scheme based on radio frequency fingerprinting, IET Communications, 12(12), March 2018.

[6] M Zhan, J Wu, Z Zhang, H. Wen and J Wu, Low-Complexity Error Correction for ISO/IEC/IEEE 21451-5 Sensor and Actuator Networks, IEEE Sensors Journal, 15(5), pp.2622-2630, 2015.

[7] F. Pan, Z. Pang, M. Xiao, H. Wen, R. Liao, Clone Detection Based on Physical Layer Reputation for Proximity Service, IEEE Access, vol. 7, pp. 3948-3957, Jan. 2019.
[8] L. Hu, H. Wen, B. Wu, F. Pan, R. Liao, H. Song, J. Tang, X. Wang, Cooperative Jamming for Physical Layer Security Enhancement in Internet of Things, IEEE Internet of Things Journal, Vol. 5, Issue 1, pp. 219-228, 2018.

[9] F. Pan, Z. Pang, M. Luvisotto, M. Xiao, H. Wen, Physical-Layer Security for Industrial Wireless Control Systems: Basics and Future Directions, IEEE Industrial Electronics Magazine, vol. 12, Issue 4, pp. 18-27, Dec. 2018.

[10] H. Song, H. Wen, H. Lin, R. Liao, Secure Cooperative Transmission with Imperfect Channel State Information Based on BPNN, IEEE Transactions on Vehicular Technology, 67(11), 10482-10491, Nov. 2018.

[11] H. Wen, P. H. Ho, B. Wu, Achieving Secure Communications over Wiretap Channels via Security Codes from Resilient Functions, IEEE Wireless Communication Letter, vol. 3, no. 3, pp. 273-276, 2014.

[12] H. Wen, P. H. Ho and X. Jiang, On Achieving Unconditional Secure Communications over Binary Symmetric Channels (BSC), IEEE Wireless Communication Letter, vol. 1, no. 2, pp. 49-52, April 2012.

[13] R. Liao, H. Wen, J. Wu, H. Song, F. Pan, and L. Dong, The Rayleigh Fading Channel Prediction via Deep Learning, Wireless Communications and Mobile Computing, Volume 2018, Article ID 6497340, 11 pages, https://doi.org/10.1155/2018/6497340

[14] Y. Cho, J. Kim, W. Yang and C. Kang, MIMO-OFDM wireless communications with MATLAB, IEEE PRESS, Singapore, 2010.

[15] J. Salamon and J. P. Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification," in IEEE Signal Processing Letters, vol. 24, no. 3, pp. 279-283, March 2017.