Computational Complexity Reduce for the Machine Learning Based RFF via Dimension Reduction

Wei Liu [1], Xinchen Xu[2], Haitao Wu [1], Fuquan Huang [1], Feiyi Xie[3], Honghao Liang [1], Jianfang Song[2], Hong Wen[2]

1. Shenzhen Power Supply Bureau Co, Ltd., Shenzhen 510820, China

2. School of Aeronautics and Astronautics, University of Elec. Science and Tech. of China, Chengdu 611731, China

3. National Key Lab of Comm., University of Elec. Science and Tech. of China, Chengdu 611731, China

sunlike@uestc.edu.cn

Keywords: Complexity, Dimension Reduction, Identification, Machine Learning, RFF

Abstract: In recent years, massive terminal connection and data transmission require 5G fast and high security access becoming very challenging. Fast and low cost authentication methods are neccessary. Radio frequency fingerprint (RFF) as a non-cryptographic authentication method that is based on devices' hardware information is an asymmetric authentication method under the edge computing system and almost all the computing load will be beared by the edge centre, which means that this method is low cost for the terminals and may meet the requirements of high speed, low power consumption and high security of 5G network. The existing best research on RFF authentication method is the method of RFF identification based on machine learning, which presenst high computing burden to the edge centre. In order to solve the problem of high computational complexity and time consuming, a principal component analysis (PCA) algorithm based on multi-resolution analysis and ReliefF was proposed. Experiments show that the scheme can achieve better recognition performance with lower complexity. This paper first introduces the principles of ReliefF and PCA algorithms, then introduces an improved algorithm based on multi-resolution analysis and reliefF algorithm, and proposes a principal component analysis algorithm based on the reliability of symbol features and sample selection. Finally, based on the data set analysis, the high recognition rate and low signal-to-noise ratio after data reduction are verified.

1. Introduction

The rapid development of the mobile Internet, the exponential growth of wireless data demand, the fifth generation (5G) network era has emerged as the times require. Due to its openness, terminal node mobility, and dynamically changing network topology, 5G networks are more vulnerable to eavesdropping, replay, message modification, and denial of service attacks. The low-power and large-connection scenario in 5G systems involves the connection of massive devices. How to achieve low computing complexity, fast data, and heterogeneous terminal access are its main challenges [1].

With the gradual strengthening of computer analysis and computing capabilities, which causes higher computational complexity and let it is difficult to fast access to massive terminals and data under 5G system [2]. Therefore, some systems and applications give up the protection of secure access authentication to ensure performance. Physical layer security realizes information theory security through the unpredictability and randomness of wireless channels or security identification through the essential attributes of hardware [3-5]. Radio frequency fingerprint (RFF) identification technology is a technology that recognizes wireless devices based on the unique hardware characteristics of the transmitting device to achieve security authentication [6]. These unique characteristics are mainly due to the manufacturing defects of the wireless transmitting device

  DOI: 10.25236/icicta.2019.017

hardware, the nuances of the line routing, etc. Just as each of us has a unique biometric fingerprint, it can also uniquely characterize wireless devices, so it is safe in wireless terminals. Access, Internet of Things (IoT) node security access and other fields have received more and more attention.

Radio frequency fingerprint (RFF) as a non-cryptographic authentication method that is based on devices' hardware information is an asymmetric authentication method under the edge computing system and almost all the computing load will be bared by the edge center, which means that this method is low cost for the terminals and may meet the requirements of high speed, low power consumption and high security of 5G network. The existing best research on RFF authentication method is the method of RFF identification based on machine learning, which presents high computing burden to the edge center [7-10].

In machine learning based RFF algorithms, when the number of feature dimensions is very large compared to the number of training, some algorithms are difficult to effectively train. This is called the "dimensional curse", and there are not many papers that apply dimensionality reduction ideas to radio frequency fingerprint identification. Principal Component Analysis (PCA) and Kernel Principal Component Analysis (KPCA) are widely used as classic dimension reduction algorithms in data dimension reduction. KPCA algorithm is a non-linear extension of PCA algorithm. The algorithm maps the original data to a high-dimensional space through a kernel function and then performs dimensionality reduction. In this paper PCA is employed to solve the problem of high computational complexity and time consuming, a principal component analysis (PCA) algorithm based on multi-resolution analysis and reliefF was proposed.

In the next section it is going to present a data dimensionality reduction model, the training process, and some computational experiments. Then, the paper concludes with remarks.

2. Data dimensionality reduction methods

Data dimensionality reduction has good effects in reducing data storage, speeding up machine learning, and reducing computational complexity. It can also visualize data, so it can often be seen in the fields of machine learning, data mining, and image processing. Figure. Classified by attributes, the dimensionality reduction methods can be divided into filtering, wrapping, and embedded. Filtering is the simplest and most time-efficient method. It assigns weights to each dimension feature, weights represent the importance of features, and are sorted according to weight. Wrapper is a method to search and verify the optimal subset from all subsets, but because the number of subsets in the set is exponential, this method has a particularly large amount of calculation. Aiming at this problem, a greedy search strategy of forward or backward search is generated to achieve the effect of efficiently searching for a subset of the feature space. The embedded method combines the idea of feature selection with the model training process to assign values to each feature. This section will introduce common methods for reducing the dimension of RF fingerprint

2.1 ReliefF algorithm

For each class of *K*, there are *n* vector training sample sets $\boldsymbol{T}=\{(\mathbf{x}_1,y_1),(\mathbf{x}_2,y_2),\ldots,(\mathbf{x}_{N\times K},y_{N\times K})\}$, among which, $\mathbf{x}_i\in\boldsymbol{X}\subseteq R^m$ is the *n*-dimension sample feature vector, and its corresponding label $y_i\in\{1,2,\ldots,K\}$, *K* represents the class of the sample. The features are weighted by releiff algorithm. Each time, a sample $\mathbf{x}_i$ is randomly selected from the training sample set, *K* same kind of neighbor samples are chosen to make up $\boldsymbol{H}(\mathbf{x}_i)$. At the same time, *K* different kind of neighbor samples are selected to constitute the $\boldsymbol{M}(\mathbf{x}_i)$, and then the weight value of each feature is updated according to

$$W(F_l)=W(F_l)-\frac{\sum_{j=1}^{k}diff(F_l,\mathbf{x}_i,H_j(\mathbf{x}_i))}{t*k}+\frac{\sum_{j=1}^{k}diff(F_l,\mathbf{x}_i,M_j(\mathbf{x}_i))]}{t*k}$$

(1) In formula (1), $H_j(\mathbf{x}_i)$

is the j nearest neighbor sample of the same kind of $\mathbf{x}_i$, $M_j(\mathbf{x}_i)$ is the j nearest neighbor sample of the different kind of $\mathbf{x}_i$, F_l represents the lth feature, $l = 1, 2, \ldots, m$, $diff(F_l, \mathbf{x}_i, \mathbf{x}_j)$ represents the difference between the sample vector $\mathbf{x}_i$ and the sample vector $\mathbf{x}_j$ on the feature F_l. The calculation formula is as follows:

$$diff(F_l, \mathbf{x}_i, \mathbf{x}_j) = \begin{cases} \dfrac{|\mathbf{x}_i[F_l] - \mathbf{x}_j[F_l]|}{\max(F) - \min(F)}, & if\ F\ is\ continuous \\ 0, & if\ F\ is\ discreat\ and\ \mathbf{x}_i[F_l] = \mathbf{x}_j[F_l] \\ 1, & if\ F\ is\ discreat\ and\ \mathbf{x}_i[F_l] \neq \mathbf{x}_j[F_l] \end{cases} \tag{2}$$

The specific steps of the releiff algorithm implementation are as follows:

Step 1. Give the sample set $\boldsymbol{T}$ and feature set F;

Step 2. Randomly select the samples $\mathbf{x}_i$, take the k samples of the same kind of $\mathbf{x}_i$ and record them as $H(\mathbf{x}_i)$, and record the samples of different classes as $M(\mathbf{x}_i)$;

Step 3. Update the weight of each feature according to the weight formula (1);

Step4. Repeat the process of randomly selecting samples and updating feature weights in step 2 and step 3 for many times to get the final feature weights $\boldsymbol{W}$ and rank them from high to low;

Step 5. Remove the features whose weight is less than the threshold to get a new sample set.

2.2 PCA algorithm

Principal Component Analysis (Principal Component Analysis, PCA) is a commonly used linear dimensionality reduction algorithm, which can realize the original data with fewer features after dimensionality reduction, and achieve maximum separability. The basic idea is to extract the main information of the data and abandon the redundant information to achieve the purpose of compressing the data. The form of information redundancy is that there will be a high correlation between some dimensions in the data. In order to eliminate information redundancy, the purpose of PCA is to rotate the original coordinate system to form a new coordinate space, so that the new Coordinate space can weaken the correlation or focus the main information in fewer dimensions. Assuming that the vector $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \in R^{\mathrm{m}}$ a training sample with zero mean, the goal of the PCA algorithm is to find a d-dimension ($d < m$) vector containing the largest variance in the vector data in space. The vector $\mathbf{x}_j$ is projected to the $\mathbf{v}$ axis, and the resulting projection vector is:

$$\left(\|\mathbf{x}_j\| \cos\theta\right) \frac{\mathbf{v}}{\|\mathbf{v}\|} = \left\langle \mathbf{x}_j, \mathbf{v} \right\rangle \mathbf{v} \tag{1}$$

where θ is the angle between the vector Ox_j and $\mathbf{v}$, the vector is a unit vector, and its coefficient is essentially the inner product. We have:

$$\left\langle \mathbf{x}_j, \mathbf{v} \right\rangle = \mathbf{x}_j^T \mathbf{v} = \mathbf{v}^T \mathbf{x}_j \tag{2}$$

The projection of all points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ in this normalized direction is

$$\mathbf{v}^T \mathbf{x}_1, \cdots, \mathbf{v}^T \mathbf{x}_N \tag{3}$$

The variance of the projection is:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{v}^T\mathbf{x}_i - 0\right)^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{v}^T\mathbf{x}_i\right)\left(\mathbf{v}^T\mathbf{x}_i\right)$$
$$= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{v}^T\mathbf{x}_i\right)\left(\mathbf{v}^T\mathbf{x}_i\right)^T = \frac{1}{N}\sum_{i=1}^{N}\mathbf{v}^T\mathbf{x}_i\mathbf{x}_i^T\mathbf{v} = \mathbf{v}^T\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i\mathbf{x}_i^T\right)\mathbf{v} = \mathbf{v}^T\boldsymbol{C}\mathbf{v} \tag{4}$$

where $\boldsymbol{C}$ is the covariance matrix. The objective is transformed into an optimization problem with equality constraints $\mathbf{v} = \arg\max_{\mathbf{v}\in R^d, \|\mathbf{v}\|=1} \mathbf{v}^T\boldsymbol{C}\mathbf{v}$. Using the Lagrangian multiplier method, the original optimization equation is equivalent to:

$$\begin{cases} \boldsymbol{C}\mathbf{v} = \lambda\mathbf{v} \\ \|\mathbf{v}\| = 1 \end{cases} \tag{7}$$

The optimization problem is calculated $\boldsymbol{v}$ to maximize $\boldsymbol{X}$, and can be solved by equation (7), then

$$\mathbf{v}^T\boldsymbol{C}\mathbf{v} = \mathbf{v}^T\lambda\mathbf{v} = \lambda\mathbf{v}^T\mathbf{v} = \lambda \tag{8}$$

Solve $\boldsymbol{C}\mathbf{v} = \lambda\mathbf{v}$ to get the largest eigenvalue λ. The eigenvector $\boldsymbol{v}$ corresponding to this value. The eigenvector is the largest principal component direction. If you need to add one more direction, you need to continue to find a second largest λ, the direction pointed by the corresponding feature vector $\boldsymbol{v}$ is the direction that makes the variance the second largest, and so on to obtain a multi-dimensional vector space. The algorithm steps of principal component analysis are as follows:

Step 1. The data vector of K class of each class is composed into N data set to perform column zero mean processing to obtain a matrix $\boldsymbol{X} \in R^{p\times m}$, where $p = K \times N$ is the number of samples requiring dimensionality reduction and $\boldsymbol{n}$ is the feature dimension of the original sample;

Step 2. Calculate the covariance matrix of $\boldsymbol{X}$:

$$Cov(\boldsymbol{X}) = \frac{\boldsymbol{X}\boldsymbol{X}^T}{p} \tag{9}$$

Step 3. Calculate the eigenvectors $\mathbf{u}_i$ and eigen values λ_i of the covariance matrix, and sort the eigen values in ascending order:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \tag{10}$$

Step 4. The contribution rate of the d previous principal components can be represented by the ratio of the sum of the previous eigenvalues to the sum of the total eigen values $\boldsymbol{X}$. Set the threshold Q (usually **Q** > 85), the dimension value d is determined after the dimension reduction, and the transformation matrix $\boldsymbol{R}$ is:

$$\boldsymbol{R} = (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_d) \tag{11}$$

Step5. Then the original matrix can be reconstructed as:

$$\boldsymbol{X}' = \boldsymbol{R}\boldsymbol{X} \tag{12}$$

Most of the feature information of $\boldsymbol{X}'$ extracted from the reconstructed matrix $\boldsymbol{X}$ and achieves the purpose of reducing the data dimension.

2.3 Improved algorithm based on multi-resolution analysis and reliefF algorithm

Aiming at the problem of high dimensionality, a feature selection method (PCA-RF) combining PCA and reliefF algorithm was proposed in [18]. However, the PCA algorithm was first used to re-

duce the dimensionality and to filter out the features that are favorable for classification risk, which reduces the classification recognition rate. Therefore, based on RF fingerprint recognition, a principal component analysis algorithm (MulRF-PCA) based on multi-resolution analysis and reliefF is proposed in this section. The multi-resolution analysis is used to extract the RF fingerprint characteristics of the signal. The features are weighted to remove features that are not conducive to classification. Finally, the PCA algorithm is used to reduce the dimensions and remove the correlation. The process of the dimensionality reduction and identification algorithm for the RF fingerprint features combined with multi-resolution analysis, reliefF, and PCA algorithms is shown in Figure 1. The multi-resolution analysis is performed on the collected RF fingerprint signals to obtain the initial features. ReliefF feature selection is followed by PCA dimensionality reduction, and finally SVM is used for classification and recognition. Using multi-resolution analysis can effectively extract RF fingerprints from the signal. However, the extracted feature dimensions are high and the time overhead of machine learning is large. Therefore, the reliefF algorithm is used to eliminate features that are not conducive to classification. However, the existence of feature redundancy will still increase the computational overhead, and then the PCA algorithm is used to remove the correlation between the features. The SVM calculation is reduced while ensuring the recognition rate the complexity.

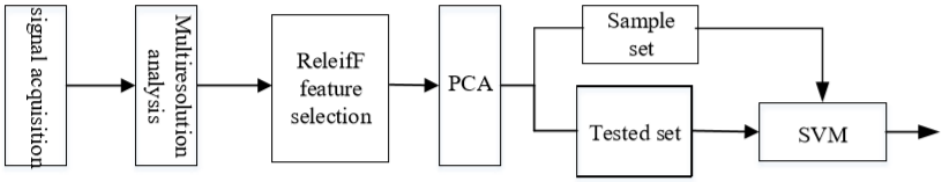

Figure 1. The results obtained by four basic predictors

The symbol characteristics based on the signal amplitude information can be expressed as:

$$\mu = \begin{cases} 0, a \geq 0 \\ 1, a < 0 \end{cases} \tag{13}$$

When the amplitude is non-negative, the symbol eigenvalue is 0, and when the amplitude is negative, the symbol eigenvalue takes 1. The signal-to-noise ratio is the ratio of the useful signal power (Power of Signal) to the noise signal power (Power of Noise), and the signal power is proportional to the square of the amplitude.

$$SNR = \frac{P_{signal}}{P_{noise}} = \left(\frac{A_{signal}}{A_{noise}} \right)^2 \tag{14}$$

Let $SNR = d$, then $\left|A_{noise}\right| = \sqrt{d}\left|A_{signal}\right|$, the amplitude of the noisy signal is expressed as:

$$A = A_{signal} \pm \left|A_{noise}\right| = \left(1 \pm \sqrt{d}\right)\left|A_{signal}\right| \tag{15}$$

With the increase of the signal-to-noise ratio, the value of $\frac{1}{\sqrt{d}}$ is reduced accordingly, so the effect of signal noise on the actual signal amplitude becomes smaller and smaller, explaining that as the SNR increases, the recognition rate of signal amplitude characteristics will gradually increase ; when $d > 1$, the value of $\frac{1}{\sqrt{d}}$ is always less than 1, so the noise will not have any effect on the symbol characteristics of the noisy signal, which explains that the symbol features have a good recognition effect at a low signal-to-noise ratio.

2.4 Experimental analysis

The experimental parameters of the relief algorithm are configured as follows. The number of

selected samples is set to 30, the number of iterations is set to 50, and the number of nearest neighbors *k* is set to 10. As shown in Figure 2, among the 14-dimensional features whose weight values are greater than 0, there are 3, 4, 8, 10, 11, 12, and 14 dimensional features. The remaining features have weight values less than 0, and the threshold is set to 0.Therefore,the algorithm will remove the 5-dimensionalfeaturesandretain the remaining 7-dimensional features, so as to describe the original 243-dimensional data, as shown in Figure 3.

The following three dimensionality reduction algorithms use SVM for training and testing, and the performance comparison between algorithms is based on runtime, test correctness (recognition) rate, and so on. Compared with the MulRF-KPCA algorithm, the running time of the MulRF-PCA algorithm is much shorter. Compared with the MulPCA-RF algorithm, the running time is not significantly improved. The correct recognition rate of the MulRF-PCA algorithm is 99.6% in the absence of noise, which is much higher than the recognition rate of the MulPCA-RF algorithm, which is 2.4% higher than the recognitionrate of the MulRF-KPC algorithm. MulRF -PCA algorithm outperforms the other two algorithms in various signal-to-noise ratios. With a recognition rate of 90% as the standard, the MulRFPCA algorithm improves 4 signal-to-noise ratios compared to the MulRF-KPCA algorithm. It can be shown that the PCA algorithm is better than KPCA. The algorithm is more suitable for application in the dimensionality reduction algorithm of RF fingerprint features. Therefore, based on the above analysis, the experiment proves that the proposed MulRF-PCA algorithm can describe the original data with less dimensional features to a great extent, and is effective in both classification performance and time efficiency.

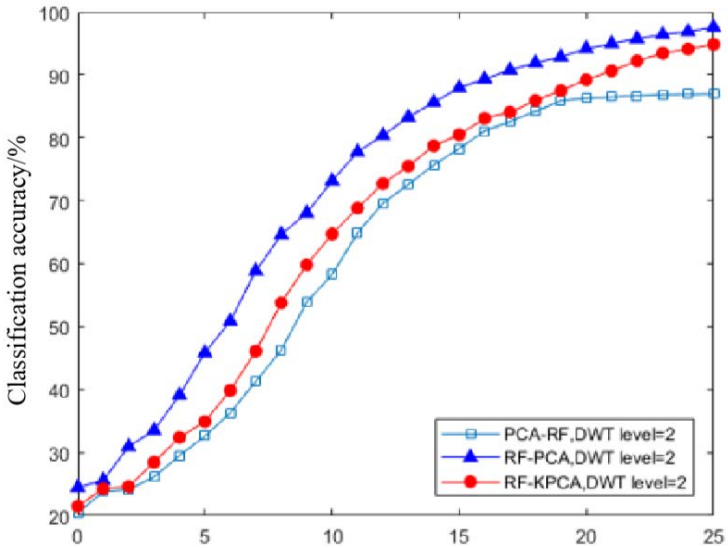

Figure 2. Test recognition rate of three algorithms

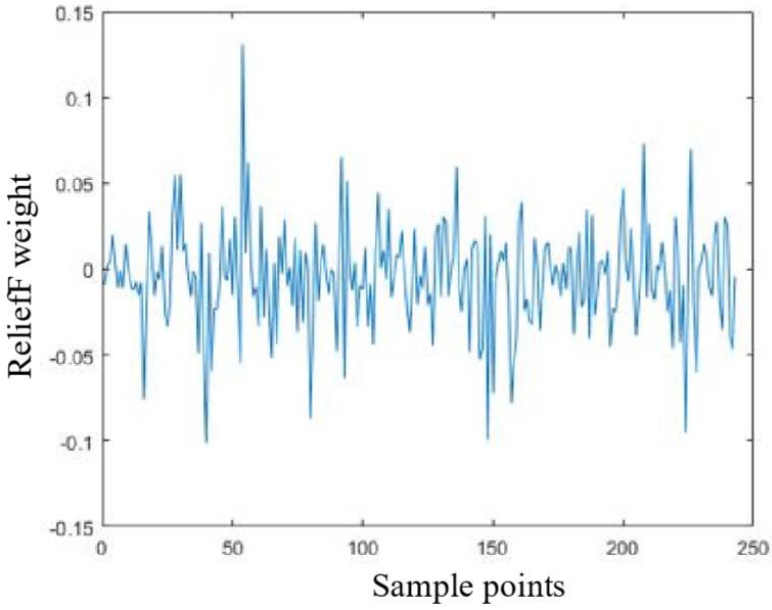

Figure 3. Weight distribution of 243 features

3. Conclusion

This paper first gives an overview of the dimensionality reduction methods, classifies them by attributes and mapping methods, and then introduces the dimensionality reduction methods commonly used in radio frequency fingerprint identification and authentication, mainly including the reliefF algorithm. The theoretical basis method is provided for the improved dimensionality reduction algorithm described later. Aiming at the problem of high feature dimension of radio frequency fingerprint recognition, this paper has presented a principal component analysis algorithm combining multi-resolution analysis and reliefF based on the feature extraction algorithm, which has achieved a higher recognition rate and lower computational overhead. By combining with an improved feature transformation method-symbol feature extraction, a reliefF based on symbol features and a principal component analysis algorithm for sample selection were proposed to ensure High recognition rate with low signal-to-noise ratio after dimensionality reduction.

Acknowledgments

This work is supported by National major R&D program (2018YFB0904900, 2018YFB0904905).

References

[1]Fei Pan, Hong Wen, Huanhuan Song, Tang Jie, Longye Wang, "5G Security Architecture and Light Weight Security Authentication," in Proc. ICCC/GSCT, Nov. 2015.

[2]N. Yang, L. Wang, G. Geraci, et al. Safeguarding 5G wireless communication networks using physical layer security. IEEE Communications Magazine, 2015, 53(4): 20-27

[3]H. Wen, J. Tang, J. Song Wu, Huanhuan Song, T. Wu, B. Wu, P-H Ho, S. Lv, L. Sun, A Cross-Layer Secure Communication Model Based on Discrete Fractional Fourier Fransform (DFRFT), IEEE Transactions on Emerging Topics in Computing, vol.3, issue 1, March 2015, pp.119-126

[4]Jie Tang, Hong Wen, Lin Hu, Huan Huan Song, Fei Pan, Hongbin Liang, Associating MIMO beamforming with security codes to achieve unconditional communication security, IET Communications 10(12): 1522-1531

[5]Lin Hu, Hong Wen, Bin Wu, Fei Pan, Run-Fa Liao, Huanhuan Song, Jie Tang, Xiumin 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.

[6]J. Hall, M. Barbeau, E. Kranakis. Enhancing intrusion detection in wireless networks using radio frequency fingerprinting. Communications, Internet and Information Technology, Thomas, US Virgin Islands, 2004, 201-206.

[7]Feiyi Xie, Hong Wen, Yushan Li, Songlin Chen, Lin Hu, Yi Chen, and Huanhuan Song，Optimized Coherent Integration-Based Radio Frequency Fingerprinting in Internet of Things, IEEE Internet of Things Journal, Vol. 5, Issue 5, pp. 3967-3977, Oct. 2018.

[8]Songlin Chen, Feiyi Xie, Yi Chen, Huanhuan Song and Hong Wen, “Identification of Wireless Transceiver Devices Using Radio Frequency (RF) Fingerprinting Based on STFT Analysis to Enhance Authentication Security,”, IEEE EMC 2017.

[9]R. Zhang, L. Zhu, C. Xu, Y. Yi. An efficient and secure RFID batch authentication protocol with group tags ownership Transfer. IEEE Conference on Collaboration and Internet Computing (CIC), Hangzhou, 2015, 168-175.

[10] Y. F. Yin, M. M. Xu, Q. Zhang, et al. Cryptanalysis of a new lightweight RFID mutual authentication protocol with cache in reader for IoT. IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference, Chengdu, China, 2017, 909-913.

[11] Lin Hu, Hong Wen, Bin Wu, Jie Tang, and Fei Pan, “Adaptive Secure Transmission for Physical Layer Security in Cooperative Wireless Networks,” IEEE Communications Letters, vol. 21, no. 3, pp. 524-527, Mar. 2017.

[12] Luping Zhang, Hong Wen, Huanhuan Song, Jie Tang, Yixin Jiang, Weihua Dong, Xiping Zhu, The Performance of the MIMO Physical Layer Security System with Imperfect CSI, CNS 2016, 17-19 October 2016 // Philadelphia, PA USA

[13] Jinling Zhang, Hong Wen, Huanhuan Song, Yixin Jiang, Zhengguang Zhang, Luping Zhang, Xiping Zhu, Using Basis Expansion Model for Physical Layer Authentication in Time-Variant System, CNS 2016, 17-19 October 2016 // Philadelphia, PA USA

[14] Fei Pan, Zhibo Pang, Michele Luvisotto, Ming Xiao, Hong 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.

[15] Yi Chen, Hong Wen, Huanhuan Song, Songlin Chen, Feiyi Xie, Qing Yang, Lin Hu, Lightweight one-time password authentication scheme based on radio frequency fingerprinting, IET Communications, 12(12): 14477-1484, March 2018.

[16] R. Liao, H. Wen, J. Wu, F. Pan, A. Xu, Y. Jiang, F. Xie, and M. Cao, “Deep-learning-based physical layer authentication for industrial wireless sensor networks,” sensors, vol. 19, no. 11, pp. 1–17, 2019.

[17] Fei Pan, Zhibo Pang, Michele Luvisotto, Xiaolin Jiang, Roger N. Janssony, Ming Xiao, and Hong Wen, "Authentication Based on Channel State Information for Industrial Wireless Communications", The 44th Annual Conference of the IEEE Industrial Electronics Society The IECON'18, 2018.10.21, Washington D.C., USA.

[18] R. W. Klein, M. A. Temple, and M. J. Mendenhall, “Application of wavelet-based RF fingerprinting to enhance wireless network security,” IEEE/KICS J. Commun. Netw., vol. 11, no. 6, pp. 544–555, Dec. 2009.