A Lightweight Intelligent Authentication Approach for Intrusion Detection

Xiaoying Qiu*, Zhidu Li†, Xuan Sun*, Tongyang Xu‡
* School of Information Management, Beijing Information Science and Technology University
Beijing 100192, China, Email: {xiaoyingqiustu, xuansun}@163.com
† School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications
Chongqing, 400065, China, Email: lizd@cqupt.edu.cn
‡ Department of Electronic and Electrical Engineering, University College London
London, UK, Email: tongyang.xu.11@ucl.ac.uk

***Abstract*—Internet of things (IoT) offers advanced and intelligent services for our life. However, smart IoT devices also bring various security vulnerabilities. Traditionally, attacks are solved by conventional authentication and authorization schemes, requiring extensive time and computational resources. In addition, it is possible to exploit artificial intelligence (AI) to provide countermeasures while enabling lightweight authentication. In this paper, we explore a solution on modelling a spoofing detection system based on machine learning and we propose a deep learning method using Auto-Extractor/Classifier Neural Network. Our scheme operates on the physical layer without causing computational overhead. Therefore, the lightweight authentication can be achieved and spoofing attacks are well-controlled in IoT scenarios.**

***Index Terms*—Physical layer authentication, artificial intelligence, deep learning, CNN, classification, prototyping, software defined radio.**

I. Introduction

With the ever growing interest in the large-scale Internet of Things (IoT) paradigm, billions or trillions of small objects are being connected. Low-cost IoT devices constitute a new communication scenario that could benefit from artificial intelligence (AI) [1]–[3]. However, due to the limited signal processing power in each resource-constrained IoT device, unforeseen attacks and invasion could cause devastating impacts on IoT communication security. Existing physical layer security (PLS) solutions [4] such as beamforming, artificial noise and directional modulation are limited to pre-known channel state information (CSI). Recently, a waveform-defined security (WDS) framework [5] is proposed to fundamentally enhance security without the CSI and additional hardware requirements. However, there are still security challenges to IoT networks. Firstly, existing security standards and protocols may not be sufficient to completely protect wireless devices. Secondly, the overhead and complexity of available PLS algorithms consume limited resources in IoT networks.

Physical layer authentication (PLA), which safeguards communications by using the intrinsic characteristics of wireless channels, is a promising lightweight security method [6]. In previous studies, a number of approaches based on artificial intelligence algorithms, including Support Vector Machine (SVM) [7], Gaussian Mixture Model (GMM) [8], [9], Genetic Algorithms [10], Random Forest [11] and others [12]–[14], have been widely used in PLA technologies. For example, an SVM-based learning method was designed to improve the detection accuracy of the authentication scheme [15]. In [16], an adaptive PLA scheme using a kernel machine was studied to deal with N-dimensional channel characteristics. Furthermore, the reinforcement learning-assisted security architecture in [17] enables an optimal test threshold. In [18], a novel deep learning authentication solution was proposed to address current security challenges. Moreover, in [19], an intrusion detection method based on k-nearest neighbors was proposed for industrial wireless networks.

Nevertheless, machine learning-assisted PLA solutions are facing both new challenges and opportunities when we consider the time-varying nature of wireless networks. For example, the channel-based PLA scheme, which applies two-dimensional static features, has severely degraded performance in a time-varying communication environment [9]. In addition, the independent design of static features extraction and clustering will increase the complexity of security authentication. In this paper, we propose a lightweight authentication approach based on auto-extractor/classifier neural networks. Specifically, low-complexity physical layer attributes are used for security authentication since the physical layer characteristics will not cause extra computational overhead to power/memory-limited IoT devices. Compared with the previous work that requires additional manual feature extractions [8], [9], [20], we use neural networks to learn the deep features of each legitimate device automatically and authenticate the attacker simultaneously.

The organization of this paper is as follows: Section II presents the system model. The learning-based PLA scheme is proposed in Section III. Following up from that, the simulation results of the proposed authentication solution are discussed in Section IV. Finally, Section V concludes the paper.

II. System model and problem statement

A. System Model

In our system, we consider three types of devices. The first type is smart device, which is capable of performing advanced tasks, such as data storage, transmission, and communication; in our security model, this type of device refers to smart device "Alice" for collecting data from other sensors. The second type is called "Bob", and its resources are limited. For the persistence of life, Bob spends most of his time on sensing and data recording. Once activated, Bob will run in data transfer mode, transfer its stored data to Alice, and clear memory for upcoming data. The remaining one, called "Cathy", is an "unknow terminal" trying to send misleading message to the smart device Alice. The various mutual authentication processes are shown in Fig. 1.

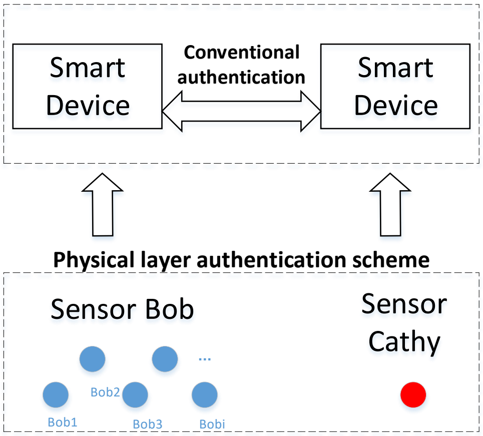


Fig. 1: The system model diagram describes the various types of authentication schemes to consider.

B. Problem Statement

Smart Alice must verify the received message at time $t+1$ and authenticate whether it is coming from Bob_i.

$$\mathrm{H_0 : |F(H_{Bob_i}(t) - H(t+1))| \leq \gamma} \quad (1)$$

$$\mathrm{H_1 : |F(H_{Bob_i}(t) - H(t+1))| > \gamma} \quad (2)$$

where $\mathrm{H_0}$ represents that the received message come from the sender Bob_i, $\mathrm{H_1}$ indicates the hypothesis that the sensor is Cathy, F is the proposed learning authentication function, $\boldsymbol{H}$ denotes the estimated physical properties of the channel, and γ is the threshold.

III. Intelligent Authentication Process

The problem with conventional PLA methods is that physical layer attributes are likely to be time-varying, but the estimated channel characteristics are static, which greatly reduces the accuracy of the authentication scheme. It is for this reason that we use neural networks to learn deep channel characteristics and perform spoofer detection concurrently. The step involved in PLA process based on deep learning is shown in Fig. 2.

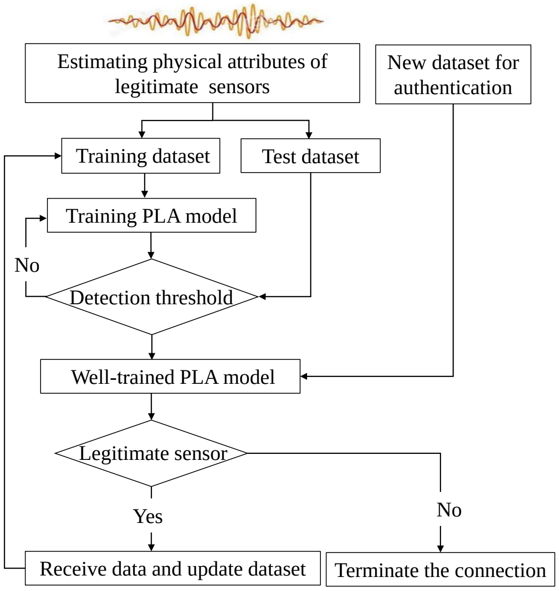


Fig. 2: Block diagram of proposed PLA.

A. PCA Reconstruction Feature for Extraction

We use principal component analysis (PCA) reconstruction as the input vector of the extractor in the proposed Auto-extractor/classifier neural network. Explicitly, the estimate channel vector is denoted as $H(t) = (h_1, h_2, \ldots, h_{256})^T$, where h_i denotes the estimated channel value. In the proposed PLA scheme, we recombine the channel vectors as shown in Fig. 3 to depress the influence of noise and make it particularly suitable for authentication classification.

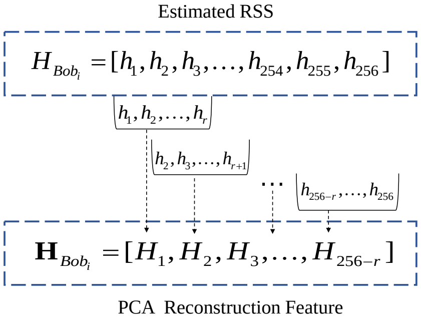

Estimated RSS
$H_{Bob_i} = [h_1, h_2, h_3, \ldots, h_{254}, h_{255}, h_{256}]$
$h_1, h_2, \ldots, h_r$
$h_2, h_3, \ldots, h_{r+1}$
$\cdots$
$h_{256-r}, \ldots, h_{256}$
$\mathbf{H}_{Bob_i} = [H_1, H_2, H_3, \ldots, H_{256-r}]$
PCA Reconstruction Feature

Fig. 3: Reconstruction Process.

The reconstruction process transforms the estimated channel vector H_{Bob_i} into a finite-dimensional matrix $\mathbf{H}_{Bob_i}$. More details of the PCA features we used can be found in [8], [21]. In this paper, we reconstruct the channel vectors as the input of the Auto-extractor/classifier neural network.

B. Auto-Extractor/Classifier Network

As discussed in the overview, it is difficult to manually capture channel characteristics with high robustness, so the back-end utilizes neural networks to learn deep features. Inspired by image classification, the convolutional neural network is used to automatically extract channel characteristics and detect the attacker. The structure of an Auto-extractor/classifier network consists mainly of two parts, including an extractor and a classifier. In order to transform physical layer properties from simple fixed features to highly robust features, here we input the reconstruction $\mathbf{H}$ into the proposed convolutional layer. The extractor captures deeper salient information and enlarges the size of channel features. We characterize the inconsistency between the predicted value of the model and the true label as follows:

$$L = -\sum_{k=1}^{K} y_k \, \mathrm{log} s_k \tag{3}$$

where K denotes the number of transmitter classes, y_k indicates that the corresponding class is set to 1 for the label k, and s_k is the kth value of the output vector s of softmax. More specifically,

$$s_k = \frac{e^{a_k}}{\sum_j^K e^{a_j}} \tag{4}$$

where a_k represents the predicted value belong to the kth class, which is the output of the final fully connected layer.

The purpose of the extractor is to learn the deep features of the input, which is beneficial to the target classification. The first layer may only learn some low-level features. The extractor maps the reconstruction matrix to a deeper representation. The mapping operation of our Auto-extractor/classifier network scheme is given by

- input map $H^{conv} \in R^{V \times W \times D}$.
- the Conv filter $F^{conv} \in R^{V' \times W' \times D \times D''}$
- the output map $y^{conv} \in R^{V'' \times W'' \times D''}$

where $V(V'$ or $V'')$ is the height, $W(W'$ or $W'')$ represents the width, and $D(D'')$ is denoted as the depth. Then the output y^{conv} can be formulated as

$$y^{conv}_{i'',j'',d''} = b_{d''} + \sum_{i'=1}^{V'} \sum_{j'=1}^{W'} \sum_{d=1}^{D} F^{conv}_{i',j',d,d''} \times H_{S_v(i''-1)+i'-P_v^-, S_w(j''-1)+j'-P_w^-, d} \tag{5}$$

where $b_{d''}$ is the bias $(d'' \in [1, D''])$, (S_v, S_w) denote the vertical (v) and horizontal (w) input samples, respectively, $(P_v^-, P_v^+, P_w^-, P_w^+)$ represent paddings.

At the training stage, each reconstruction matrix is fed into our proposed learning scheme to extract deep features. The fully-connected layer is designed to make decisions about physical layer authentication. It is observed from (3) that the loss function is used to optimize the Auto-extractor/classifier network during training phase. The proposed PLA scheme will perform a forward propagation and backward propagation iteratively.

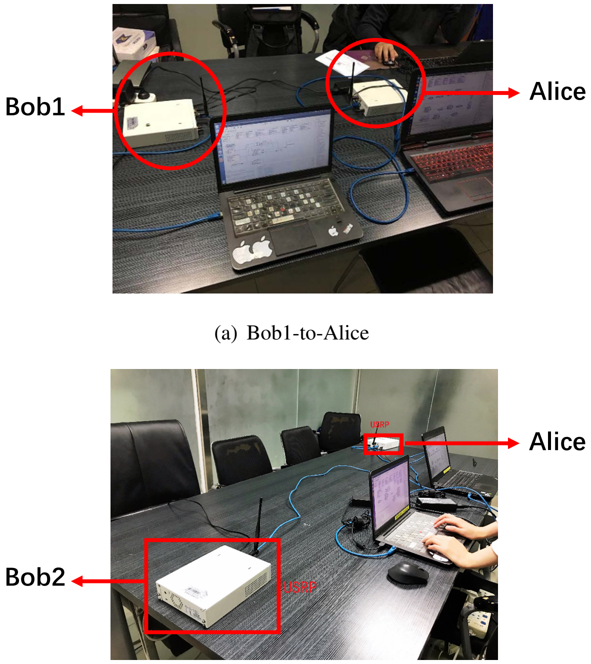
For the testing purpose, we input the test data into the proposed Auto-extractor/classifier scheme and calculate the probability that the estimated channel attributes belong to different senders. More specifically, the prediction function of the classifier is defined as

$$P = \frac{e^{a_k}}{\sum_j^K e^{a_j}} \tag{6}$$

IV. Prototype and Performance Evaluation

A. Prototype System Setup

To emulate the learning-based PLA method, we set the Universal Software Radio Peripheral (USRP) transceiver to operate in IEEE 802.11a/g mode, working at 2.4 GHz and having a bandwidth of 20 MHz. Fig. 4(a) and Fig. 4(b) show the experiment setup in an indoor conference room. We investigate the performance of the proposed PLA model in the binary classification (Bob1 and Bob2). All estimates of received signal strength constitute a park with two classification targets. The sampled data set used for authentication provides a more realistic basis for theoretical verification. To automate the training model generation and detection authentication process, we created the script using the Python language.

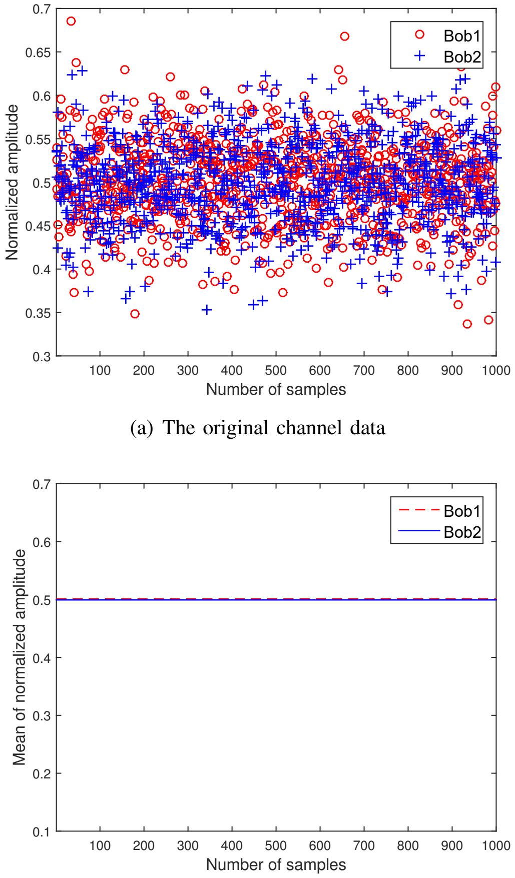


(a) Bob1-to-Alice

(b) Bob2-to-Alice

Fig. 4: Experimental areas.

We collect data for each transmitter-receiver combination. As shown in Fig. 5, two types of samples are involved in the proposed USRP data set, each of which contains 2000 × 256 channel feature sampling points, a total of 4000 × 256 samples. During the training process, 1000 × 256 sampling points are randomly selected for each type of channel feature data. Since the memory is limited, it is not possible to load all the data at once, so set each feature matrix size to 8×256. The steps involved in feature extraction are shown in Fig. 2 and



(a) The original channel data

(b) The average amplitude of channel data

Fig. 5: The distribution of physical layer features.

Fig. 3. According to the feedback network, the structure of the proposed extractor/classifier network is shown in Table I.

TABLE I: Parameters setting

Description	Input Size	Output Size
Convolution 1	$16 \times 16 \times 8$	$8 \times 8 \times 32$
Pooling 1	$8 \times 8 \times 32$	$4 \times 4 \times 32$
Convolution 2	$4 \times 4 \times 32$	$2 \times 2 \times 128$
Pooling 1	$2 \times 2 \times 128$	$1 \times 1 \times 128$
Fully Connected	$1 \times 1 \times 128$	$1 \times 1 \times 1$

B. Security Analysis

1) The Effect of SNRs on Convergence: Fig. 6 characterizes the training performance of the proposed PLA approach. We can observe from Fig. 6 that the Softmax loss value of the proposed authentication scheme reaches its steady-state value after 150 iterations when SNR $=$ 8 dB, while that relying on the higher SNR has the lowest loss value. In other words, increasing the SNR of the sample in a conference room accelerates the convergence. This is because the training performance of the proposed approach depends on the wireless communication environment. Additionally, the loss curves of the proposed algorithm show a decreasing trend during the initial training process. However, after the iteration of 30, its training performance deteriorates. The reason for this trend is that the signal collected in the conference room is an estimated degradation value, and its signal strength may be interfered by the surrounding wireless signals, resulting in incomplete channel estimation values. In the actual communication scenario, the power of the transmitter will be much higher than that of this USRP transmitter. Therefore, it is expected that the performance of our scheme will be better.

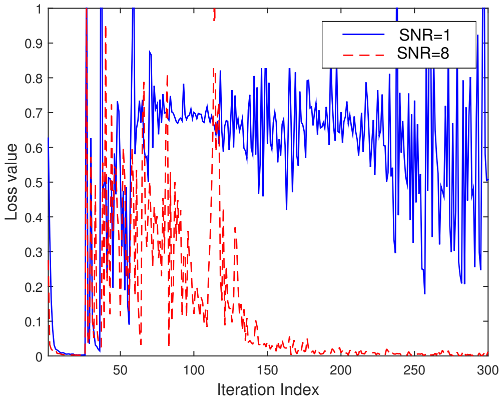


Fig. 6: Training performance curve of the proposed authentication scheme.

2) Performance Comparison: In order to verify the advantages of the Auto-extractor/classifier scheme, we also need to compare and simulate the existing detection approaches, namely SVM and GMM. Table II lists the authentication results. Compared with the SVM method, the proposed method is significantly better than the SVM method for all SNRs between 2 dB and 10 dB, and has greater reliability (100% vs. 98.88% when SNR is 10 dB). It is also observed from Table II that the Auto-extractor/classifier method shows its advantages in tolerating low SNR. More importantly, compared with conventional methods, the Auto-extractor/classifier approach evaluates the trustworthiness between the sending and receiving devices. The results show that the accuracy of the learning-based authentication method is related to the dimensions of the channel characteristics used.

Since communication conditions involve interference that may reduce the received signal strength, the authentication task becomes uncertain in non-ideal situations. In this experiment, we try to verify the convergence of the proposed method under different communication channel conditions. Fig. 7 shows the effectiveness of the Auto-extractor/classifier approach in two different situations, namely line-of-sight (LOS) and non-line-of-sight (NLOS). It can be clearly observed that

TABLE II: The performance comparison results.

SNR	Detection Rate (%)		
	SVM	GMM	Proposed PLA
4dB	91.95	95.01	**95.89**
6dB	97.50	95.99	**100.00**
8dB	98.01	96.56	**100.00**
10dB	98.88	98.01	**100.00**

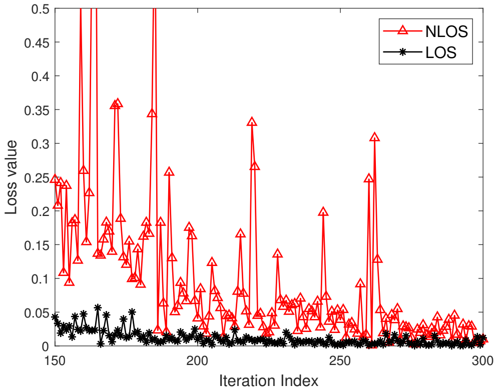


Fig. 7: Training performance of intelligent authenticator in LOS and NLOS scenarios.

the Auto-extractor/classifier scheme has better convergence performance in the case of LOS. This is because the inherent characteristics of the estimated signal are weakened in the NLOS scenario. This indicates that the proposed approach is less robust in NLOS conditions. Actually, existing PLA schemes are difficult to extract robust features from time-varying attributes, which results in poor detection results. In contrast, the Auto-extractor/classifier approach can effectively learn deep features, which is a better authentication scheme.

V. Conclusion

In this paper, we proposed a lightweight intelligent authentication scheme based on Auto-Extractor/Classifier Neural network. Since the independent design of feature extraction and clustering will increase the complexity of authentication, therefore we use neural networks to learn the channel characteristics and conduct the authentication task concurrently. To validate the effectiveness of utilizing the intelligent scheme, USRP prototype systems are set up in an indoor conference room. Furthermore, the rigorous security analysis and convergence of the proposed approach under different SNRs are comprehensively evaluated.

References

[1] B. Chatterjee, D. Das, S. Maity and S. Sen, “RF-PUF: Enhancing IoT Security Through Authentication of Wireless Nodes Using In-Situ Machine Learning,” in IEEE Internet of Things Journal, vol. 6, no. 1, pp. 388-398, Feb. 2019.

[2] G. Gui, F. Liu, J. Sun, J. Yang, Z. Zhou, and D. Zhao, “Flight delay prediction based on aviation big data and machine learning,” IEEE Transactions on Vehicular Technology, vol. 69, no. 1, pp. 1065-1069, 2020.

[3] T. Xu and I. Darwazeh, “Design and Prototyping of Neural Network Compression for Non-Orthogonal IoT Signals,” 2019 IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, 2019, pp. 1-6.

[4] Y. Zou, J. Zhu, X. Wang and L. Hanzo, “A Survey on Wireless Security: Technical Challenges, Recent Advances, and Future Trends,” in Proceedings of the IEEE, vol. 104, no. 9, pp. 1727-1765, Sept. 2016.

[5] T. Xu, “Waveform-defined security: a framework for secure communications,” in IEEE/IET 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP2020), Porto, Portugal, Jul. 2020.

[6] N. Wang, P. Wang, A. Alipour-Fanid, L. Jiao and K. Zeng, “Physical-Layer Security of 5G Wireless Networks for IoT: Challenges and Opportunities,” in IEEE Internet of Things Journal, vol. 6, no. 5, pp. 8169-8181, Oct. 2019.

[7] T. M. Hoang, T. Q. Duong and S. Lambotharan, “Secure Wireless Communication Using Support Vector Machines”, Proc. IEEE Conference on Communications and Network Security (CNS), Washington DC, DC, USA, 2019, pp. 1-5.

[8] X. Qiu et al., “Wireless User Authentication Based on KLT and Gaussian Mixture Model,” 2019 IEEE Wireless Communications and Networking Conference (WCNC), Marrakesh, Morocco, 2019, pp. 1-5.

[9] A. Weinand, M. Karrenbauer, J. Lianghai, H. D. Schotten, “Physical Layer Authentication for Mission Critical Machine Type Communication Using Gaussian Mixture Model based Clustering”, Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring), pp. 1-5, Jun. 2017.

[10] A. L. Buczak and E. Guven, “A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection,” in IEEE Communications Surveys and Tutorials, vol. 18, no. 2, pp. 1153-1176, Secondquarter 2016.

[11] N. Farnaaz and M. A. Jabbar, “Random Forest Modeling for Network Intrusion Detection System,” Procedia Comput. Sci., vol. 89, pp. 213-217, 2016.

[12] N. Van Huynh, D. N. Nguyen, D. T. Hoang and E. Dutkiewicz, “Jam Me If You Can: Defeating Jammer With Deep Dueling Neural Network Architecture and Ambient Backscattering Augmented Communications,” in IEEE Journal on Selected Areas in Communications, vol. 37, no. 11, pp. 2603-2620, Nov. 2019.

[13] S. Choudhary, N. Kesswani, “A Survey: Intrusion Detection Techniques for Internet of Things,” in International Journal of Information Security and Privacy, pp. 86-105, 2019.

[14] J. M. Hamamreh, H. M. Furqan and H. Arslan, “Classifications and Applications of Physical Layer Security Techniques for Confidentiality: A Comprehensive Survey,” in IEEE Communications Surveys and Tutorials, vol. 21, no. 2, pp. 1773-1828, Secondquarter 2019.

[15] C. Pei, N. Zhang, X. S. Shen and J. W. Mark, “Channel-based physical layer authentication,” 2014 IEEE Global Communications Conference, Austin, TX, 2014, pp. 4114-4119.

[16] H. Fang, X. Wang and L. Hanzo, “Learning-Aided Physical Layer Authentication as an Intelligent Process,” in IEEE Transactions on Communications, vol. 67, no. 3, pp. 2260-2273, March 2019.

[17] L. Xiao, T. Chen, G. Han, W. Zhuang and L. Sun, “Game Theoretic Study on Channel-Based Authentication in MIMO Systems,” in IEEE Transactions on Vehicular Technology, vol. 66, no. 8, pp. 7474-7484, Aug. 2017.

[18] H. Yu, Z. Tan, Z. Ma, R. Martin and J. Guo, “Spoofing Detection in Automatic Speaker Verification Systems Using DNN Classifiers and Dynamic Acoustic Features,” in IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 10, pp. 4633-4644, Oct. 2018.

[19] S. Henningsen, S. Dietzel, and B. Scheuermann, “Misbehavior Detection in Industrial Wireless Networks: Challenges and Directions,” in Mobile Netw Appl, 23: 1330, 2018.

[20] A. Mahmood, W. Aman, M. O. Iqbal, M. M. U. Rahman, and Q. H. Abbasi, “Channel impulse response-based distributed physical layer authentication”, Proc. IEEE 85th Vehicular Technology Conference (VTC Spring), pp. 1-5, Jun. 2017.

[21] X. Qiu, J. Dai and M. Hayes, “A Learning Approach for Physical Layer Authentication Using Adaptive Neural Network,” in IEEE Access, vol. 8, pp. 26139-26149, 2020.