Automated Labeling and Learning for Physical Layer Authentication Against Clone Node and Sybil Attacks in Industrial Wireless Edge Networks

Songlin Chen, *Student Member, IEEE*, Zhibo Pang, *Senior Member, IEEE*,
Hong Wen, *Senior Member, IEEE*, Kan Yu, *Member, IEEE*,
Tengyue Zhang, *Student Member, IEEE*, and Yueming Lu

***Abstract*—In this article, a scheme to detect both clone and Sybil attacks by using channel-based machine learning is proposed. To identify malicious attacks, channel responses between sensor peers have been explored as a form of fingerprints with spatial and temporal uniqueness. Moreover, the machine-learning-based method is applied to provide a more accurate authentication rate. Specifically, by combining with edge devices, we apply a threshold detection method based on channel differences to provide offline training sample sets with labels for the machine learning algorithm, which avoids manually generating labels. Therefore, our proposed scheme is lightweight for resource constrained industrial wireless devices, since only an online-decision making is required. Extensive simulations and experiments were conducted in real industrial environments. Both results show that the authentication accuracy rate of our strategy with an appropriate threshold can achieve 84% without manual labeling.**

***Index Terms*—Cyber physical security, physical layer authentication, supervised machine learning.**

Manuscript received October 31, 2019; revised December 4, 2019 and December 20, 2019; accepted December 28, 2019. Date of publication January 15, 2020; date of current version November 20, 2020. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61572114, in part by the National Major R & D Program (2018YFB0904900 and 2018YFB0904905), in part by Sichuan Science and Technology Service Development Project (18KJFWSF0368), and in part by Sichuan Science and Technology Basic Research Condition Platform Project (2018TJPT0041). Paper no. TII-19-4819. *(Corresponding author: Hong Wen.)*

S. Chen is with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China (e-mail: songlinch0061@163.com).

Z. Pang is with the Department of Automotive Solutions, ABB Corporate Research, Västerås 72226, Sweden (e-mail: pang.zhibo@se.abb.com).

H. Wen and T. Zhang are with the Department of Aeronautics and Astronautics, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China (e-mail: sunlike@uestc.edu.com; uestczty@163.com).

K. Yu is with the La Trobe University, Bundoora VIC 3086, Australia (e-mail: kan.yu@hotmail.com).

Y. Lu is with the Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: ymlu@bupt.edu.cn).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2020.2963962

I. Introduction

INDUSTRIAL wireless networks offer unprecedented opportunities for Industry 4.0. However, industrial wireless communications are vulnerable to probing free attacks, which are not possible in wired communication systems [1]. Security is of vital importance, especially in industrial control systems. If malicious information or commands are sent to control devices by attackers, unexpected results may occur, such as an interruption in the industrial process, leading to economic losses or even safety accidents. For example, an attacker can launch a clone node attack in an unsupervised industrial wireless network. The attacker first hijacks legitimate node and extracts its ID, key, and other confidential information. Subsequently, a large number of cloned nodes are deployed by the attacker in the industrial wireless network, which seriously threaten the industrial network security in a way of amending routing information, collecting sensitive information, or interrupting key distributions. In addition, a Sybil attack is another type of security threats to industrial wireless networks, where a malicious node claims multiple fake identities. Since the industrial control center may not be able to distinguish those multiple fake identities, the malicious node will be able to attack industrial networks.

Owing to the resource constraints of wireless sensor network devices, a lightweight and effective detection method is needed in industrial wireless networks. Especially for clone attacks, as attackers are able to obtain all sensitive information from a compromised legitimate node, it is difficult to detect and deter these attacks by relying on cryptographic algorithms. In [2]–[8], physical layer authentication based on generalized channel responses with spatial variability was put forward, which is based on generalized channel response with spatial variability and considers the correlation of time, frequency, and spatial domains. On account of this, we propose an attack detection method based on channel differences. To the best of our knowledge, the existing methods are not able to detect both clone and Sybil attacks at once. In addition, in order to further improve the detection rate, a spoofing attack detection strategy combined with machine learning algorithm is proposed in [9]. However, the offline training model of the machine learning needs strong computing power. Moreover, another premise of this work is that the machine learning algorithm needs to be optimized by providing samples with labels in the offline training phase. Previous studies about machine learning have focused on artificial injection of label samples, and it is difficult to obtain attack samples in advance [10].

To further improve the network efficiency, edge computing or fog computing is introduced into industrial wireless networks to meet the needs of low latency of industrial applications such as real-time monitoring and control for critical industrial devices [11]–[13]. Both edge and fog computing acquire more devices to process data from nearby devices on the basis of traditional centralized cloud platform data processing and analysis methods, thus it can relieve the traffic burden and enable ultralow latency response. Regarding industrial edge networks, a lot of studies focus on optimizing resource allocation by utilizing edge computing networks and security protection [14]–[17]. By this means, the edge computing can provide the proximal support platform that makes use of machine-learning-based method to improve the channel state information (CSI)-based attack detection rate. Therefore, in this article, we proposed an automated labeling and learning method for physical layer authentication to detect clone and Sybil attacks in edge computing industrial wireless networks. First, we provide a method of utilizing channel differences and threshold detection, as well as the offline training sample sets with labels, for the machine learning classification algorithm. Second, a support vector machines (SVMs) algorithm is utilized to realize online decision of two types of attack detection at once, since the SVM algorithm has better performance under small size offline training sample sets compared with other machine learning algorithms. Our main contributions are summarized as follows.

1) We propose a physical layer authentication strategy based on channel differences to detect clone attacks and Sybil attacks simultaneously in industrial wireless environments.
2) Compared with the existing machine-learning-based physical layer authentication schemes, the labels of offline training sample sets can be generated automatically without manual operation by our proposed strategy.
3) The simulations of both clone and Sybil attacks detection are conducted by using open datasets from the National Institute of Standards and Technology. Field validations are carried out in real industrial environments to verify the feasibility of our proposed method [18].

The rest of this article is organized as follows. Section II illustrates related work. Attack modeling and analysis are introduced in Section III. In Section IV, the proposed scheme is described in detail, followed by simulation and experimental verification for our proposed strategy in Sections V and VI. Finally, Section VII conclude this article.

II. Related Work

A high number of research efforts have been put on the detection of clone and Sybil attacks by utilizing physical layer authentication. CSI-based Sybil attacks detection in wireless networks was proposed in [19]. This approach is an enhanced physical layer authentication scheme to detect Sybil attacks by exploiting the spatial variability of radio channels. Physical layer authentication is based on judging the channel information difference of nodes, which is caused by their different spatial locations. In [2], the feature-based physical layer authentication protocol is highly effective in detecting Sybil attacks for mission-critical machine-type communication applications. The existing physical layer authentication methods are mainly studied from three perspectives. First, appropriate channel estimation algorithms are selected to improve the accuracy of CSI. Due to channel differences of communications peers need to be compared, it is essential to obtain accurate CSI through channel estimation. Second, a better channel difference is constructed between different nodes, and which can be constructed as the test statistic in the binary hypothesis testing. A variety of channel difference test statistics has been proposed by other scientists. For instance, the channel-based likelihood ratio test and sequential probability ratio test are exploited for authentication in smart systems in [20]. Third, a suitable threshold also affects the effect of channel-based physical layer authentication. User authentication in the practical leveraging CSI-based threshold method is studied in [5] for spoofing detection. However, it is prone to multipath effect and Doppler frequency shift. At present, in order to improve the accuracy of detection rate, some researchers devoted to study machine learning algorithms based on CSI. In [9] and [10], various supervised learning algorithms are applied to detect spoofing attacks. Q-learning and game theory are utilized to adjust and determine the threshold in channel authentication. All of these approaches need offline training sample sets with labels. Those are provided to the machine learning algorithm for continuous optimization of the learning model.

To the best of our knowledge, none of the previous studies is able to detect clone and Sybil attacks at the same time. Moreover, the existing CSI-based attack detection methods combined with machine learning algorithms need offline training samples, requiring labels with attack identities or legitimate identities. However, the existing methods need to label samples manually. They fail to satisfy the requirement of offline sample training for detection model without knowing the label attributes of attack nodes beforehand [10]. To overcome this drawback, we propose an automated labeling and learning strategy for physical layer authentication. This strategy utilizes the channel difference threshold method to generate labels of learning samples into the offline training sample sets when the machine learning algorithm is trained offline. In addition, our method can also detect clone and Sybil attacks at the same time in wireless industrial networks.

III. Attack Modeling and Analysis

In this section, we first introduce industrial wireless edge networks (IWENs), and then, give an overview of clone and Sybil attacks in IWENs.

A. Industrial Wireless Edge Networks (IWENs)

An IWEN is regarded as a bridge between industrial wireless devices and the remote cloud. It can provide real-time edge intelligent services to meet the critical needs of industry digitization in terms of agile connectivity, real-time services,

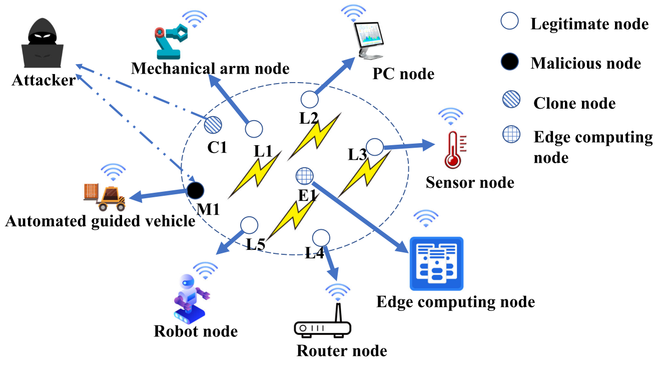


Fig. 1. Clone attack in IWEN: The attacker replicates the clone node C1 based on M1 in large quantities and deploys C1 in different positions in the industrial control system based on edge computing.

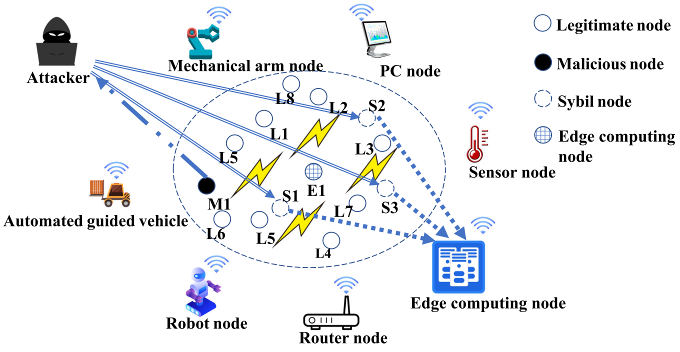


Fig. 2. Sybil attack in IWEN: The attacker deploys multiple Sybil nodes, including S1, S2, and S3, in the industrial control system, but their true position is M1.

data optimization, application intelligence, security and privacy protection [16]. In order to improve the work efficiency of assembly lines in automotive factories, a local wireless autonomous network with an industrial edge computing node is established. The network consists of local industrial wireless sensors for industrial environment detection and wireless control nodes (e.g., mechanical arm, unmanned logistics vehicles, robots, etc.). An industrial edge computing node can work in real time and efficiently with other local terminal devices. Moreover, this node ensures a smooth and efficient collaboration among devices by dispatching the operation status of all kinds of terminals reasonably, thus realizing a real time and efficient operation in industrial environments [15]. In addition, the industrial edge computing node is able to work with a remote cloud computing. The remote cloud computing is responsible for the analysis of nonfactual and long-period data, while the industrial edge computing node focuses on real-time, short-period data analysis. An industrial edge computing network can support real-time responses and high mobility in an industrial operating environment. Clone and Sybil attacks may occur in the IWEN. Fortunately, the industrial edge computing is more conducive to providing sufficient computing resources, and utilizing machine learning algorithms based on channel identification attack detection in a proximal location.

B. Overview of Clone Attacks and Sybil Attacks

We consider that an industrial edge computing node controls multiple wireless device nodes as local real-time decision center in an edge-computing-based industrial control environment. There are various wireless nodes in the wireless network, such as manipulators, wireless personal computers, temperature sensors, automatic steering vehicles, and wireless control nodes, shown in Fig. 1.

The industrial edge computing node communicates with multiple nodes to achieve the coordination and integration of all kinds of wireless sensors. It needs to ensure the network security of all access wireless nodes and provide resources for various devices to ensure stable operation.

1) Clone Attack Scenario: The scenario of clone attacks in industrial control systems based on edge computing is shown in Fig. 1. When an attacker hijacks the legitimate node M1 (automated guided vehicle), he is able to access all sensitive information related to this node (e.g. key, control information, data information, ID information, etc.). Therefore, the attacker is able to replicate a clone node C1 based on the extracted information from M1 and deploys C1 in a different location in this IWEN. Then, the clone node C1 is able to act as the legitimate node M1 and participate in data interaction with the industrial edge computing node E1. Meanwhile, the other nodes L1,..., L5 are legitimate nodes.

A clone attack will bring a series of serious consequences. For example, when a great number of clone nodes exist in the IWEN, they are able to launch a denial of service attack by occupying the communication channel at all times and sending access requests to the industrial edge computing node continuously to prevent the industrial edge computing node from receiving any information from legitimate nodes. Moreover, by broadcasting malicious information in the network, attackers can inject malicious data or change transmission information of legitimate nodes, leading to a collapse or breakdown of the industrial edge computing node.

2) Sybil Attack Scenario: As the local real-time processing center, the industrial edge computing node needs to interact with other nodes shown in Fig. 2. Sybil attacks are malicious nodes pretending to be other nodes or claiming faked IDs in the industrial wireless network. As shown in Fig. 2, an attacker deploys multiple Sybil nodes, including S1, S2, and S3, in the industrial control system, but their true location is M1. Meanwhile, L1,..., L8 are legitimate nodes. The edge control node mistakenly believes that S1, S2, and S3 are located in different locations in the network topology, but in fact, they are in the same geographical location. For example, if nodes S1 and S2 are generated from the node M1 and they are not in the positions being claimed in Fig. 2, the nodes S1, S2, and M1 are physically in the same position.

Sybil attackers fake multiple identities to communicate with the industrial edge computing node, affecting the communication network of other legitimate nodes, thus intercepting and tampering with information, and decreasing network availability and undermining data integrity. In particular, Sybil nodes may

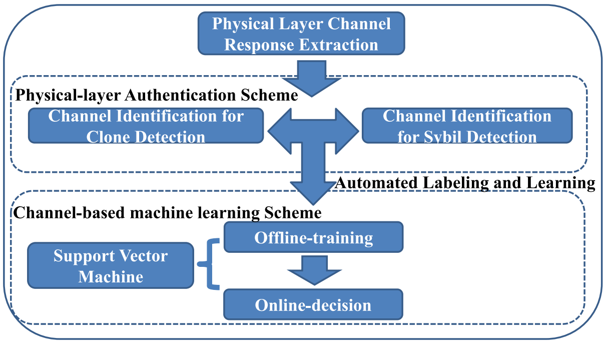

Physical Layer Channel Response Extraction
Physical-layer Authentication Scheme
Channel Identification for Clone Detection
Channel Identification for Sybil Detection
Automated Labeling and Learning
Channel-based machine learning Scheme
Support Vector Machine
Offline-training
Online-decision

Fig. 3. Overall framework of automated labeling and learning for physical layer authentication.

continuously send a huge amount of messages to the industrial edge computing node by imitating legitimate devices or with fake IDs to request network and storage resources. Consequently, due to the network resource exhaustion, the IWEN may collapse, leading to more serious economic losses or safety accidents.

IV. Automated Labeling and Learning for Physical Layer Authentication

We propose an automated labeling and learning scheme for physical layer authentication to detect both clone and Sybil attacks, consisting of the physical layer channel response extraction, physical layer authentication scheme, and channel-based machine learning scheme.

A. Overall Framework

Within the framework of Fig. 3, the physical layer channel response extraction is the key step. In [6], several approaches of channel estimation are introduced. Also, we can apply least square (LS) and normalized minimum square error (NMSE) to extract channel information in our work. Then, a physical layer authentication strategy is taken to, respectively, detect clone attacks and Sybil attacks through channel differences, generating offline training sample sets. Finally, we develop two attack detection models including offline training and online decision making by the SVM machine learning method based on channel. Offline training first utilizes training sample sets automatically generated by the physical layer authentication strategy to train the offline model of the machine learning algorithm. In the end, a real-time online decision will be carried out. In this article, all these processes will be investigated, as discussed in detail in the remainder of this section.

B. Physical Layer Channel Response Extraction

The extraction of CSI is undertaken by legitimate receivers. The signal model of the legitimate receiver can be given by

$$r(t) = hx(t) + n(t) \tag{1}$$

where t, h, and x, respectively, refer to the time slot, which means the time interval between every data frames, a time-domain channel matrix that is the matrix value of channel coefficients, a pilot signal known by transmitters and receivers and used to estimate channel information, and $n(t)$ is additive white Gaussian noise with variance σ^2. The channel frequency response generated by the receiver through the channel estimation is as follows:

$$\hat{H}_k = RX^{-1} = H_k + NX^{-1} \tag{2}$$

where H_k is the channel frequency response, R, N, and X are obtained by discrete Fourier transform in time domain r, n, and x. LS and NMSE methods can be used to estimate the channel response $\hat{H}_k$ from pilots [6]. According to [6], $\hat{H}_k$ can be expressed by the following:

$$\hat{H}_k(n) = [\hat{H}_k(n,1), \hat{H}_k(n,2), \ldots, \hat{H}_k(n,M)]^T \tag{3}$$

where k is the frame index, n is the symbol index, and M is the dimension of channel information.

C. Physical Layer Authentication Scheme

Different node locations indicate different channel states. Therefore, the CSI characteristics of a legitimate node are supposed to be different from those of cloned ones. Similarly, for the Sybil attack detection, since a malicious node has multiple identities but only one physical device, when this malicious node launches Sybil attacks, the CSI from this node remains the same with different identities. Thereby, it is possible to determine whether a Sybil attack occurs. The physical layer authentication strategy, at the beginning of detection, requires information exchange between nodes to realize attack recognition and channel estimation. For example, the industrial edge computing node periodically broadcasts request messages B to other nodes with timestamp T_w. All nodes, after receiving the message B, send response signals to the industrial edge computing node at the time interval $t + \tau$, where τ represents the response delay. Meanwhile, τ must be less than channel coherence time γ, [6]. The response signal R is given by

$$R|\{\text{Pilot}, \text{Identity}\}. \tag{4}$$

After receiving the response signals, the industrial edge computing node initially distinguishes the attack type and recognizes whether there is an identity conflict by comparing with saved legitimate identity information. If there exist conflicts of identity declaration, it is possible that node with multiple IDs initiates a Sybil attack. If the identity remains the same, it is possible that node with the same ID initiates a clone attack. Therefore, the authentication process is expressed as

$$f(\text{ID}) = \begin{cases} \text{ID}_i = \text{ID}_j, & \text{Legitimate or clone node} \\ \text{ID}_i \neq \text{ID}_j, & \text{Legitimate or Sybil node} \end{cases}. \tag{5}$$

1) Channel Identification for Clone Detection: The channel identification algorithms for clone attacks are launched after the communication and interaction between the industrial edge computing node and terminal nodes to be tested. The industrial edge computing node initially interacts with terminal nodes. Specifically, after receiving identity information from each node, the industrial edge computing node makes a preliminary assessment based on the received information. Clone attacks may occur if some nodes have the same ID. In this case, the industrial edge computing node uses the pilot signal to estimate the channel and generate the CSI $\hat{H}_k(n)$, which can be given by

$$\hat{H}_k(n) = [\hat{H}_k(n,1), \hat{H}_k(n,2), \ldots, \hat{H}_k(n,M)]^T. \tag{6}$$

Assuming that an industrial edge computing node obtains the CSI $\hat{H}_k$, at the receiver side of nodes to be tested at t_1, the CSI $\hat{H}_{k+1}$ at t_2 is compared with $\hat{H}_k$. Then, it is possible to determine whether the channel information received from the different time comes from the same node. The clone attack detection problem can be considered as a binary hypothesis test. In the null hypothesis $\mathcal{H}_0$, the message of the transmitting node comes from the same location, thus there is no clone attack. Otherwise, the alternative hypothesis $\mathcal{H}_1$, is that the transmitting node comes from different geographical locations, and there are clone attacks on the node. The binary hypothesis testing can be described as

$$\begin{cases} \mathcal{H}_0 : \hat{H}_k \longrightarrow \hat{H}_{k+1}, & \text{No existing clone attack} \\ \mathcal{H}_1 : \hat{H}_k \longmapsto \hat{H}_{k+1}, & \text{Existing clone attack} \end{cases}. \tag{7}$$

where $\mathcal{H}_0$ indicates no clone attacks occurring, while $\mathcal{H}_1$ means clone attacks exist. To determine the similarity of the CSI $\hat{H}$, it is essential to make comparisons of the difference and threshold of the channel matrices of two consecutive data frames $\hat{H}_k, \hat{H}_{k+1}$. We adopted the Euclidean distance as the test statistic, which can be given by

$$T_{\text{clone}} = \text{diff}(\hat{H}_{k+1}, \hat{H}_k) = \frac{\left\|\hat{H}_{k+1} - \hat{H}_k\right\|_2}{\left\|\hat{H}_k\right\|_2} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \eta_1 \tag{8}$$

where $\hat{H}_k$ is the CSI of a legitimate node, $\hat{H}_{k+1}$ is the CSI of a node to be tested. When the CSI difference is less than the clone threshold η_1, the channel matrices are very similar, which determines that the current transmitter is legitimate. Contrarily, when the CSI difference is greater than the clone threshold η_1, the channel matrices are considered to be different, which determines that the current transmitter is a clone node with the same ID in different locations.

2) *Channel Identification for Sybil Detection:*

In Sybil attack, malicious nodes pretend to be multiple other nodes and have multiple IDs in the IWEN. When the industrial edge computing node communicates and interacts with terminal nodes, the channel identification algorithms for Sybil attacks are launched. Specifically, the industrial edge computing node, according to the received response signals R from each node, determines that Sybil attacks may occur if some terminal nodes have different IDs. In this context, the industrial edge computing node uses the pilot signal to estimate the channel and generate the CSI $\hat{H}_k(n)$, which can be given by

$$\hat{H}_k(n) = [\hat{H}_k(n,1), \hat{H}_k(n,2), \ldots, \hat{H}_k(n,M)]^T. \tag{9}$$

Assuming that the industrial edge computing node obtains the CSI $\hat{H}_k$ at the receiver side of nodes to be tested at t_1, the CSI $\hat{H}_{k+1}$ at t_2 is compared with $\hat{H}_k$. Then it is possible to determine whether the channel information received from the different time comes from the same node. The Sybil attack detection problem can be considered as a binary hypothesis test. In the null hypothesis $\mathcal{H}_0$, the message of the transmitting node comes from different locations, thus there is no Sybil attack. Otherwise, the alternative hypothesis $\mathcal{H}_1$ indicates that the transmission nodes are from the same location and Sybil attacks may exist. Hypothesis testing can be expressed as

$$\begin{cases} \mathcal{H}_0 : \hat{H}_k \longmapsto \hat{H}_{k+1}, & \text{No existing sybil attack} \\ \mathcal{H}_1 : \hat{H}_k \longrightarrow \hat{H}_{k+1}, & \text{Existing sybil attack} \end{cases}. \tag{10}$$

To determine the similarity of the channel matrices, it is essential to make the comparison of the CSI difference and threshold of the channel matrices of two consecutive data frames. Again we adopted the Euclidean distance as the test statistic

$$T_{\text{Sybil}} = \text{diff}(\hat{H}_{k+1}, \hat{H}_k) = \frac{\left\|\hat{H}_{k+1} - \hat{H}_k\right\|_2}{\left\|\hat{H}_k\right\|_2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\lessgtr}} \eta_2 \tag{11}$$

where $\hat{H}_k$ is the CSI of a node to be tested at the time of k, and $\hat{H}_{k+1}$ is the CSI of a node to be tested at the time of $k+1$. When the CSI difference between the two is less than the Sybil threshold η_2, it indicates that a Sybil attack occurs. When the difference is less than the Sybil threshold η_2, the channel matrices are very similar, and the nodes have different ID in the same location. Thus, it is determined that the current transmitter is a Sybil attack node.

D. *Support Vector Machine (SVM)*

In machine learning, SVM is functioned as a supervised learning model with associated learning algorithms that analyze data used for classification and regression analysis. We consider adopting an SVM as our attack detection algorithms for two main reasons. First, the attack detection is a two-class problem that only has two results. One is that attacks exist, the other indicates no attacks. Second, due to the limited sample size of open datasets provided by NIST, we need to consider that machine learning algorithms should probably gain better classification results with small offline training sample sets. The optimization model is given by

$$\begin{aligned} \max_{\boldsymbol{\alpha}} \quad & \sum_{n=1}^{N} \alpha_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m k(\boldsymbol{x_n}, \boldsymbol{x_m}) \\ \text{s.t.} \quad & \sum_{n=1}^{N} \alpha_n y_n = 0, \\ & \alpha_n \geq 0, \quad (n = 1, \ldots, N). \end{aligned} \tag{12}$$

where $k(\boldsymbol{x}, \boldsymbol{x'}) = \phi(\boldsymbol{x})^T \phi(\boldsymbol{x'})$ is the kernel function. In this article, the offline training sample set S comprises n input vectors $\hat{H}_1, \ldots, \hat{H}_n$ with corresponding target labels $y_1, \ldots, y_n$, where $y_i \in \{-1, 1\}$. S can be given by

$$S = \{(\boldsymbol{x_1} = \hat{H}_1, y_1), \ldots, (\boldsymbol{x_n} = \hat{H}_n, y_n)\}. \tag{13}$$

Algorithm 1: Automated Labeling for Physical-Layer Authentication.

```
Input: Clone threshold $\eta_1$, Sybil threshold $\eta_2$, the feedback
    signal of $i$ node $R_i = \{\text{Pilot}, \text{ID}_i\}$, the feedback signal of
    $j$ node $R_j = \{\text{Pilot}, \text{ID}_j\}$.
Output: Offline training samples.
 1:  If $\text{ID}_i = \text{ID}_j$
 2:  Calculate CSI and channel difference $T_{\text{clone}}$ between $i$
     and $j$ nodes, via (6) and (8);
 3:    If $T_{\text{clone}} > \eta_1$
 4:      Insert clone attack labels $y_k^j = -1$ into channel
 5:      samples of $j$ node;
 6:      Issue a warning that clone attack occurs;
 7:      Else insert legitimate node labels $y_k^j = +1$ into
 8:      channel samples of $j$ node;
 9:    End if.
10:  Else if $ID_i \neq ID_j$
11:  Calculate CSI and channel difference $T_{\text{Sybil}}$ between $i$
     and $j$ nodes, via (9) and (11);
12:    If $T_{\text{Sybil}} < \eta_2$
13:      Insert Sybil attack labels $y_k^j = -1$ into channel
14:      samples of $j$ node;
15:      Warn Sybil attack occur;
16:      Else insert legitimate node label $y_k^j = +1$ for
17:      channel samples of $j$ node;
18:    End if.
19:  End if.
20:  Output offline training samples including clone attack
     samples, Sybil attack samples, and legitimate node
     samples.
```

The offline training sample set S is utilized to obtain a value for $\boldsymbol{\alpha}$, and $\boldsymbol{w}^* = \sum_{n=1}^{N} \alpha_n^* y_n \phi(\boldsymbol{x_n})$ via (12) [21]. We can then determine the b^* by any support vector $\boldsymbol{x_n}$ satisfying $y_n(\sum_{m\in S} \alpha_m^* y_m k(\boldsymbol{x_n}, \boldsymbol{x_m}) + b^*) = 1,\ y_n \in \{-1, 1\}$, where S denotes the set of indices of the support vectors [21]. By multiplying y_n and making use of $y_n^2 = 1$ and all support vectors, we obtain a more stable solution b^* given by [21]

$$b^* = \frac{1}{N_s} \sum_{n\in S} \left(y_n - \sum_{m\in S} \alpha_m^* y_m k\left(\boldsymbol{x_n}, \boldsymbol{x_m}\right) \right) \tag{14}$$

where N_S is the total number of support vectors. According to w^* and b^*, the final SVM offline model is given by

$$f(\boldsymbol{x}) = \text{sign}\left(\sum_{n=1}^{N} \alpha_n^* y_n k(\boldsymbol{x}, \boldsymbol{x_n}) + b^* \right). \tag{15}$$

E. Channel-Based Machine Learning Scheme

We hypothesize that attackers launched Sybil attacks and clone attacks in the IWEN. As a result, there exist three kinds of nodes: legitimate nodes, clone nodes, and Sybil nodes. Meanwhile, each node has its own ID. The industrial edge computing node can deal with automated labeling for physical layer authentication via Algorithm 1. The industrial edge computing node classifies the nodes according to their ID. If nodes have the same identity, the industrial edge computing node needs to consider whether these nodes are legitimate nodes or clone ones, and then, conducts the channel authentication. The industrial edge computing node extracts CSI of each node to obtain the channel difference T_{clone}, and compares it with the clone threshold value η_1 to determine whether a clone attack occurs. If a clone attack occurs $T_{\text{clone}} > \eta_1$, the industrial edge computing node injects the clone attack label $y_k^i = -1$ into channel response vectors of the clone attack node, $[\hat{H}_k^i, y_k^i = -1]$, to form a clone attack node sample set, whereas the legitimate node injects the legitimate label $y_k^i = +1$ to form a legitimate node sample set. Offline training samples sets are given by $S = \{[\hat{H}_k^i, y_k^i = -1], \ldots, [\hat{H}_k^j, y_k^j = +1]\}$. In another situation, if these nodes have different IDs, the industrial edge computing node begins to consider whether these nodes are legitimate nodes or Sybil nodes, and then, also performs the physical layer authentication process. The industrial edge computing node extracts the CSI of the nodes, respectively, and compare their channel differences at different times. If the difference T_{Sybil} is less than the Sybil threshold η_2, the industrial edge computing node can correctly detect whether a Sybil attack has occurred in this node. If a Sybil attack occurs $T_{\text{Sybil}} < \eta_2$, the industrial edge computing node injects Sybil attack labels $y_k^j = -1$ into the Sybil attack node to form a sample set of Sybil attack nodes. On the contrary, the industrial edge computing node injects legitimate labels $y_k^j = +1$ to form a sample set of legitimate nodes. Offline training sample sets are given by $S = \{[\hat{H}_k^i, y_k^i = -1], \ldots, [\hat{H}_k^j, y_k^j = +1]\}$. Subsequently, the industrial edge computing node uses the labeled node channel information samples to form the training and testing sample sets S required by the machine learning algorithm. According to Algorithm 2, the industrial edge computing node implements the CSI-based machine learning algorithm, and optimizes the model through the sample sets S to achieve the target authentication accuracy rate G. Otherwise, the optimization model will continue iteratively returning to step 2 until target G achieved. Finally, an online attack detection model is generated to detect clone attacks and Sybil attack nodes.

V. Simulations on Real Industrial Open Dataset

A. Measurement Setup

The industrial environment chosen for this measurement is an automotive factory of 400 × 400 × 12 m, which is full of metallic equipment and obstacles. The channel information used in this work is based on the NIST dataset that are acquired by the channel sounder system composed of NIST TX and RX channel sounder [18]. The transmitter repeatedly sends the pseudonoise code serial number of a set of digital symbols, which are modulated by the signal of binary phase shift keying and converted up to the radio frequency carrier frequency. Though the power amplifier, the signals propagate through the factory to the TX and RX channel sounder. The receiver converts, digitizes, and stores the received signal locally. Balancing steps are performed during the postprocessing to eliminate hardware damage. The channel measurement equipment keeps moving in during the

Algorithm 2: Machine Learning for Physical Layer Authentication.

```
Input: The labeled node channel information samples $S$ can
    be given by, $S = [\hat{H}_k^i, y_k^i = -1], \ldots, [\hat{H}_k^j, y_k^j = +1]$,
    target authentication accuracy rate $G$.
Output: Online authentication model.
1:  Divide $S$ into training sample set $S_1$ and testing
    sample set $S_2$;
2:  Calculate offline authentication model based on SVM,
    via $S_1$ and (12) and (14). Specifically, $\hat{H}_k^i$ is put into
    the kernel function $K(\cdot)$, and $y_k^i$ is put into (14);
3:  Calculate authentication accuracy rate $A$ of offline
    authentication model based on the SVM, via $S_2$ and
    (15) and (18);
4:  If $A > G$
5:      Return online authentication model = offline
6:      Authentication model.
7:  Else return to step 2 for optimizing the offline
    authentication model;
8:  End if.
9:  Output online authentication model.
```

measurements, following a loop with non-line-of-sight path with rich multipath. The parameter configuration of the channel measurement system includes center frequency, antenna, and power. The center frequency is 2.245 GHz, an omnidirectional antenna is applied for both the receiving and transmitting, and the polarization modes are Cross pol and V pol, respectively. The receiving antenna gain is −4.2 dBi, the transmitting antenna gain is 2.9 dBi, and the transmitting power is 1.5 W. The sampling rate is 80 MHz. In this scenario, 106 channel measurement positions were measured and 300 records are for each measurement position [18].

B. Performance Metrics

The receiver operating characteristic (ROC) curve is the measurement standard of the effect of strategy on attack detection, consisting of true positives rate (TPR) and false positive rate (FPR). TPR is the proportion of the number of attacks being detected in the total number of all attacks, given by

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{16}$$

where true positive (TP) is the number of attacks being detected and false negative (FN) is the number of attacks launched without being detected. The sum of the two is the total number of attacks. The FPR is the proportion of the number of attacks not being detected in the total number of all attacks given by

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{17}$$

where false positive (FP) is the number of attacks being not launched but detected. True negative (TN) represents the number of attacks being not launched but correctly detected. The sum of the two is the total number of attacks that do not occur. In addition, the authentication accuracy rate (AC) is a key performance metric, which is given by

$$\mathrm{AC} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}. \tag{18}$$

Table I illustrates the relationship between TP, TN, FP, and FN.

TABLE I
PERFORMANCE METRICS

		Actuality	
		Existing attack	No existing attack
Detection	Existing attack	TP	FP
	No existing attack	FN	TN

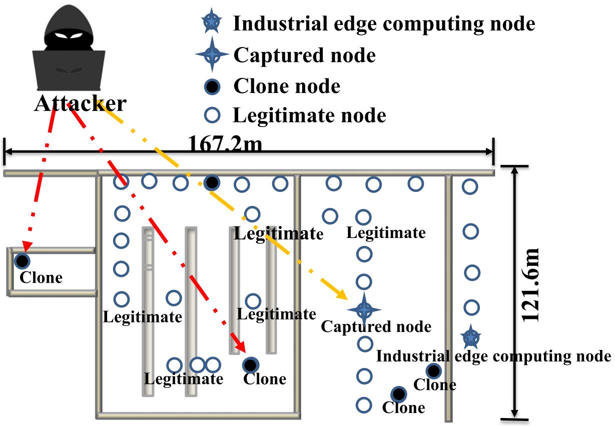


Fig. 4. Simulation experiment of clone attack under automotive assembly, utilizing open dataset from NIST.

C. Simulation of Clone Attack Scenario

We utilize the channel information dataset provided by NIST in the automotive factory to simulate clone attacks. To make use of multiple mobile measurement locations in the automative factory, we set up legitimate nodes and an industrial edge computing node. Assuming that an attacker captures a legitimate node, then multiple cloned nodes are deployed in different locations of the industrial network. As shown in Fig. 4, in the simulation, the nodes deployed are 167.2-m long and 121.6-m wide. The number of nodes in three simulation are 30, 60, and 90, respectively, where we guarantee that clone nodes occupy 20% of all nodes.

D. Performance of Clone Attack

In order to satisfy the condition that the number of clone nodes occupy 20% of all nodes, we assume a fixed channel measurement location is a legitimate node, and the other are clone nodes. We deploy 30, 60, and 90 nodes in the network. Therefore, the number of clone attack nodes in the three simulations is 6, 12, and 18, respectively. Fig. 5(a) shows the ROC curve of the clone threshold ranging from 0 to 1 with interval 0.1, representing the detection performance of the system. The simulation results show that the value of the TPR in the three curves can achieve 1 at a lower FPR, indicating the effectiveness of detecting clone attacks.

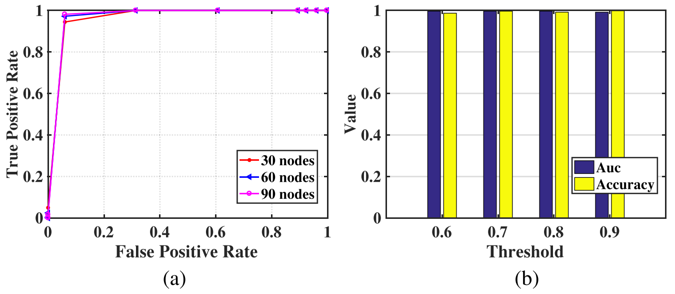


Fig. 5. (a) ROC of channel threshold method for detection clone attack under different nodes density. (b) Accuracy of channel-based SVM scheme for detection clone attack.

Another simulation is carried out to verify our proposed strategy. We utilize different clone thresholds to provide labeled offline training sample sets for the machine learning models. The experiment uses different clone thresholds to generate attack labels into node samples in different sizes. The node under attacks is marked as -1, and the other nodes without being attacked are marked as $+1$. The purpose is to form offline training samples for machine learning. Subsequently, the machine learning algorithm generates the final clone attack detection model to satisfy our set authentication accuracy rate through offline training sample sets. Finally, the online decision of the model is carried out to evaluate the performance of the system. The simulation results show that the offline training sample sets generated by different clone thresholds affect the clone attack detection model based on machine learning of CSI. The area under curve (AUC) and authentication accuracy rate are the indexes to measure the classification of machine learning. Fig. 5(b) shows that the offline training samples generated under four different cloning thresholds have no significant impact on the attack detection rate of the SVM based on CSI.

E. Simulation of Sybil Attack Scenario

We also adopt NIST channel information dataset in an automotive factory to simulate Sybil attacks. The legitimate nodes and an industrial edge computing node are set up at multiple mobile locations in this factory. If an attacker attacks a legitimate node, the node becomes a malicious node in an IWEN. The malicious node masquerades as another node or claims to fake IDs. As shown in Fig. 6, the black solid circles are Sybil nodes generated from the hexagonal malicious nodes, in which S1 and S2 are generated from M1; S3 and S4 are generated from M2. The true position of S1 and S2 is the position of M1. The true position of S3 and S4 is the position of M2. The area of this simulation is 167.2-m long and 121.6-m wide. The number of nodes in three simulation are 30, 60, and 90, respectively, where we guarantee that Sybil nodes occupy 20% of all nodes.

F. Performance of Sybil Attack

In order to satisfy this condition that Sybil nodes occupy 20% of all nodes, we hypothesizes that a fixed channel measurement location is a malicious node being attacked, and other different measurement locations are legitimate nodes being disguised by Sybil nodes or claimed to counterfeit IDs. We deploy 30, 60, and 90 nodes in the network, and keep the number of Sybil attack nodes as 6, 12, and 18, respectively. The ROC curves in the three cases are shown in Fig. 7(a). The ROC curve of the set Sybil threshold ranges from 0 to 1 with interval 0.1, representing the detection performance of the system. In the situation where the total number of nodes deployed is small scale, the simulation results show that the threshold attack detection performance based on CSI differences is better than that based on large scale nodes. Since the channel measurement information when devices are moving is different from that when devices keep static, the threshold method is a more appropriate choice when deploying a small-scale node in the industrial wireless network.

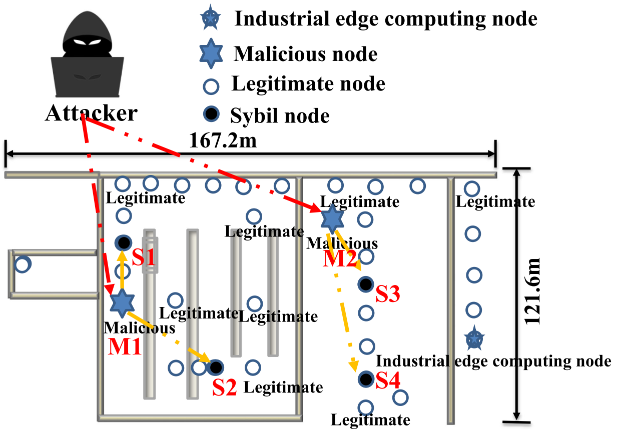


Fig. 6. Experiment of Sybil attacks under automotive assembly, utilizing open datasets from NIST. S1 and S2 are generated from M1 and the true position of them is at M1; S3 and S4 are generated from M2 and the true position of them is at M2.

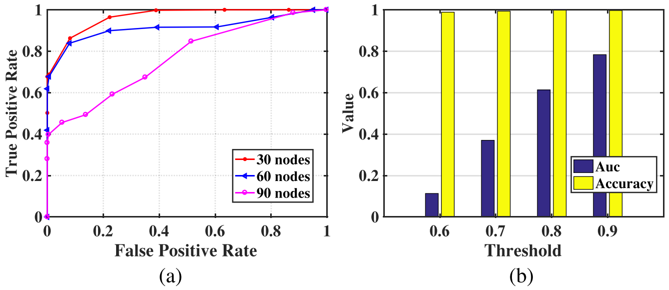
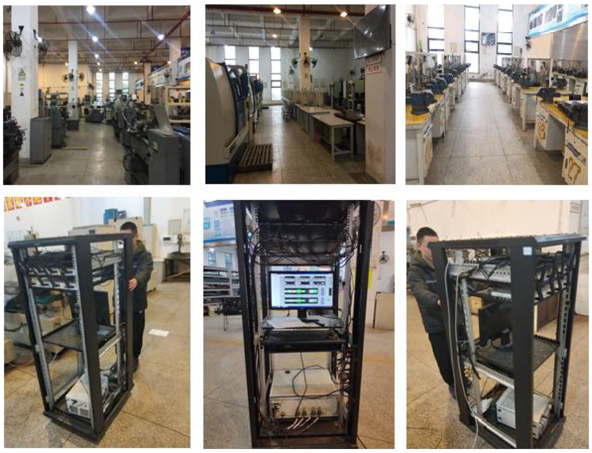


Fig. 7. (a) ROC of channel threshold method for detection of Sybil attack under different nodes density. (b) Accuracy of the CSI-SVM physical authentication scheme for detection of Sybil attack.

Another simulation is implemented to verify our proposed strategy. An SVM is applied for the attack detection method based on CSI machine learning. First, the Sybil threshold method is utilized for attack detection in the mobile environment so that it can provide label generation under small-scale samples for offline training. Furthermore, an SVM is suitable for classification detection under small sample conditions. As shown in Fig. 7(b), we use different Sybil thresholds, 0.6, 0.7, 0.8, and 0.9, respectively, to generate labels into the offline training sample set of the initial machine learning algorithm. If the node is under a Sybil attack, it is marked



(a) (b) (c)

Fig. 8. Engineering training center for real experiment environment and test platforms.

as -1. Otherwise, it is marked as $+1$. Then, the machine learning algorithm utilizes four different offline training sample sets to optimize the SVM model, and forms a Sybil attack detection model. This model satisfies the condition that the target authentication accuracy rate should be higher than 90%. Finally, a new sample set is employed to judge the model online and evaluate the performance of the model.

As shown in Fig. 7(b), the Sybil attacks detection model to implement four CSI-based SVM machine learning models is developed from offline training sample sets being generated at different Sybil thresholds. The four attack detection results show that the authentication accuracy rate of attack detection can reach 100%, and the AUC increases with the increase of the threshold value. The selection of the Sybil threshold determines the final performance of our proposed strategy, and also verifies the feasibility of our proposed scheme, thus reducing the complexity of the manual generation of learning labels.

VI. Experimental Verification

A. Experimental Setup for Verification

In order to verify the effects of the strategy in reality, we conducted attack detection experiments in an engineering training center. The engineering training center with a lot of metal equipment can truly reveal the reflection and scattering environment of the industrial wireless network, as shown in Fig. 8(a)–(c). The engineering training center is considered to be large, with outer dimensions of approximately 48.3 m × 38.8 m and a ceiling height of approximately 6.5 m. A map of the experimental site floor is shown in Fig. 9. A total of six Universal Software Radio Peripheral (USRP) platforms provided by National Instrument (NI) are employed. The equipment is numbered into six nodes, one of which is taken as an industrial edge computing node, and the rest are simulated terminals. The industrial edge computing is statically located at position 4, and other nodes are also statically placed at position 1, 2, 3, 5, and 6, respectively, as shown in Fig. 9.

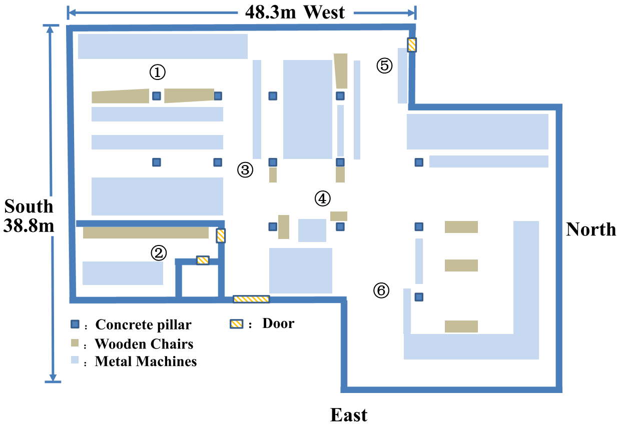
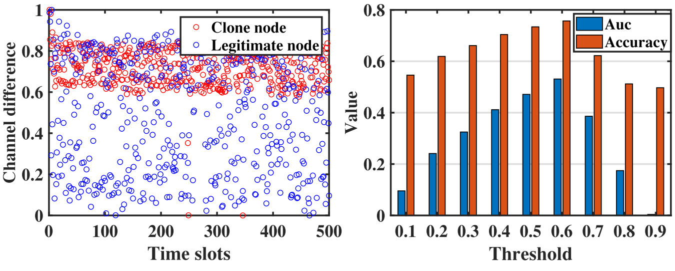


Fig. 9. Layout of experiment factory with area of 48.3 × 38.8 m where six nodes to be test are located.

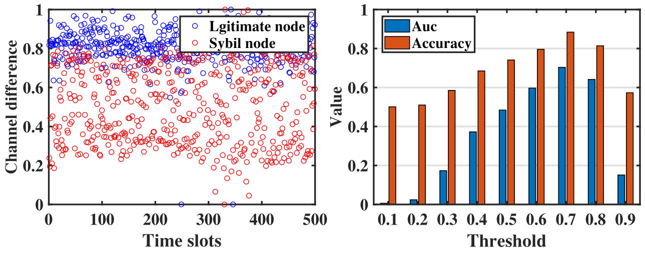


(a) (b)

Fig. 10. (a) Channel difference of clone and legitimate nodes. (b) Result of the SVM scheme for clone attack detection.

As shown in Fig. 8, each node is an 8 × 8 MIMO transceiver. Its specific configuration parameters are listed below. The center frequency is 3.5 GHz. The signal bandwidth is 2 MHz. The number of MIMO subcarriers is 128, and the transmission power is 15 dBm. The other nodes numbered 2, 3, 4, 5, and 6 can establish communication links with the industrial edge computing node. Subsequently, clone attack detections and Sybil attack detections are implemented.

1) Verification of Clone Attack Detection: The node 6 initiates clone attacks, and intercepts all information of node 5, including ID, key, and then, cheats the industrial edge computing node. The industrial edge computing node detects attacks and evaluates effect through automated labeling and learning scheme for the physical layer authentication. Fig. 10(a) shows the channel differences between the clone node and the legitimate node after normalization ranging from 0 to 1 according to (8). This figure shows that the channel difference of the clone node is greater than that of the legitimate node. The strategy generates a labeled offline training sample set in order to implement machine learning algorithms by setting different clone thresholds. The machine learning algorithm utilizes a linear SVM, where the kernel size is automatic and the cross validation is 5 k-fold. When the target authentication accuracy rate for the machine learning algorithm is larger than 90%, the model after offline training is the final authentication model. Then, the model is tested. The node 6 is taken as the clone attack node once again. It clones legitimate node 5. Meanwhile, it launches 1000 clone attacks



(a) (b)

Fig. 11. (a) Channel difference of Sybil and legitimate nodes. (b) Result of the SVM scheme for Sybil attack detection.

on the industrial edge computing node. The attack detection results of the authentication model generated under different clone threshold settings are shown in Fig. 10(b).

The results show that when the clone threshold is 0.6, AC can reach 75%. Furthermore, AUC is 0.531, which is higher than other clone thresholds. The results reflect the impact of accuracy of labeling in offline training on the machine learning algorithm. The accuracy of labeling in offline training set needs to be further improved.

2) Verification of Sybil Attack Detection: The node 3, with multiple identities including node 1, 2, and 6, initiates Sybil attack. The industrial edge computing node needs to judge whether Sybil attack occurs at node 3. Then, we implement the proposed Sybil attack detection strategy. Fig. 11(a) describes the channel differences between Sybil attack node and legitimate node after normalizing the differences from 0 to 1 through (11). The figure shows the channel characteristics of the Sybil node, and the channel difference at different times is smaller than that of the legitimate node. Subsequently, the automated labeling and learning process is implemented. Different Sybil thresholds are set to build a labeled offline training sample set. The machine learning algorithm also employs the linear SVM, where the kernel size is automatic and cross validation is 5 k-fold. When the authentication rate of the machine learning algorithm in offline training is larger than 90%, it is regarded as the final authentication model. Then, the authentication model is tested practically. Sybil attack was launched at node 3 once again, and 1000 attacks were carried simultaneously. Under different Sybil thresholds, attack detection results of the strategy are as shown in Fig. 11(b). When the Sybil threshold is set as 0.7, the attack detection effect of the generated authentication model is the best, reaching 84%, compared with other thresholds. This experiment shows the feasibility of the strategy. On the other hand, it shows again that the accuracy of the labeled offline training set is the key to obtain the high authentication accuracy rate of the machine learning model.

VII. Conclusion

In this article, we discussed physical layer channel information combined with the machine learning algorithm for detecting clone and Sybil attacks. In order to accomplish the goal that offline training sample sets were automatically labeled in the initial stage of the strategy, we proposed a channel difference threshold detection method to label learning samples. Thus, we solved the problem that the physical layer authentication method based on machine learning lacks learning samples. More importantly, we verified the feasibility of the strategy by conducting the simulation with the wireless network CSI dataset disclosed in the industrial environment and experiment in practical industrial environment.

For the future work, one direction is to improve the accuracy of offline training sample labels for machine learning through a better channel difference threshold method. Moreover, better channel characteristics are required to improve the accuracy of channel identification. Last but not least, we will follow with interest to the advanced persistent threat (APT) attack, which is a kind of sustained and effective attack and hard to be detected by the encryption-based scheme. CSI-based physical layer authentication will be explored for the APT attack detection, since long-term CSI information collection could be helpful for the comprehensive detection of the APT attack.

References

[1] F. Pan, Z. Pang, M. Luvisotto, M. Xiao, and H. Wen, “Physical-layer security for industrial wireless control systems: Basics and future directions,” *IEEE Ind. Electron. Mag.*, vol. 12, no. 4, pp. 18–27, Dec. 2018.

[2] H. Forssell, R. Thobaben, H. Al-Zubaidy, and J. Gross, “Physical layer authentication in mission-critical MTC networks: A security and delay performance analysis,” *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 795–808, Apr. 2019.

[3] J. Yang, Y. Chen, W. Trappe, and J. Cheng, “Detection and localization of multiple spoofing attackers in wireless networks,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 44–58, Jan. 2013.

[4] F. Pan, Z. Pang, M. Xiao, H. Wen, and R. Liao, “Clone detection based on physical layer reputation for proximity service,” *IEEE Access*, vol. 7, pp. 3948–3957, Dec. 2019.

[5] F. Pan *et al.*, “Authentication based on channel state information for industrial wireless communications,” in *Proc. 44th Annu. Conf. IEEE Ind. Electron. Soc.*, 2018, pp. 4125–4130.

[6] H. Wen, *Physical Layer Assisted Authentication for Wireless Sensor Networks*. New York, NY, USA: Springer, 2013, pp. 37–57.

[7] H. Wen, J. Luo, and L. Zhou, “Lightweight and effective detection scheme for node clone attack in wireless sensor networks,” *IET Wireless Sensor Syst.*, vol. 1, no. 3, pp. 137–143, Sep. 2011.

[8] L. Jiao, N. Wang, P. Wang, A. Alipour-Fanid, J. Tang, and K. Zeng, “Physical layer key generation in 5G wireless networks,” *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 48–54, Oct. 2019.

[9] S. Chen *et al.*, “Physical-layer channel authentication for 5G via machine learning algorithm,” *Wireless Commun. Mobile Comput.*, vol. 2018, 2018, Art. no. 6039878.

[10] F. Pan *et al.*, “Threshold-free physical layer authentication based on machine learning for industrial wireless CPS,” *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6481–6491, Dec. 2019.

[11] L. Chen, P. Zhou, L. Gao, and J. Xu, “Adaptive fog configuration for the industrial Internet of Things,” *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4656–4664, Oct. 2018.

[12] P. Pace, G. Aloi, R. Gravina, G. Caliciuri, G. Fortino, and A. Liotta, “An edge-based architecture to support efficient applications for healthcare Industry 4.0,” *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 481–489, Jan. 2019.

[13] C. Lin and J. Yang, “Cost-efficient deployment of fog computing systems at logistics centers in Industry 4.0,” *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4603–4611, Oct. 2018.

[14] W. Sun, J. Liu, Y. Yue, and H. Zhang, “Double auction-based resource allocation for mobile edge computing in industrial Internet of Things,” *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4692–4701, Oct. 2018.

[15] M. Aazam, K. A. Harras, and S. Zeadally, "Fog computing for 5G tactile industrial internet of things: QoE-aware resource allocation model," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3085–3092, May 2019.

[16] D. A. Chekired, L. Khoukhi, and H. T. Mouftah, "Industrial IoT data scheduling based on hierarchical fog computing: A key for enabling smart factory," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4590–4602, Oct. 2018.

[17] S. K. Mishra, D. Puthal, J. J. P. C. Rodrigues, B. Sahoo, and E. Dutkiewicz, "Sustainable service allocation using a metaheuristic technique in a fog server for industrial applications," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4497–4506, Oct. 2018.

[18] R. Candell, K. A. Remley, and N. Moaveri, "Radio frequency measurements for selected manufacturing and industrial environments," NIST-Tech. Rep., 2016. [Online]. Available: http://doi.org/10.18434/T44S3N

[19] 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 Netw.*, vol. 27, no. 3, pp. 34–39, May/Jun. 2013.

[20] H. Wen, Y. Wang, X. Zhu, J. Li, and L. Zhou, "Physical layer assist authentication technique for smart meter system," *IET Commun.*, vol. 7, no. 3, pp. 189–197, Feb. 2013.

[21] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2006.

Songlin Chen (Student Member, IEEE) is currently working toward the Ph.D. degree in communication and information system with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu, China.

His current main research interests include wireless communication system security, physical layer security, artificial intelligence, and industrial communications.

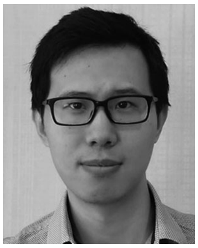
Zhibo Pang (Senior Member, IEEE) received the MBA degree from the University of Turku, Turku, Finland and the Ph.D. degree from the Royal Institute of Technology (KTH), Stockholm, Sweden.

He is currently a Senior Principal Scientist with ABB Corporate Research, Västerås, Sweden, an Adjunct Professor with the University of Sydney, Sydney, Australia, and an Affiliated Faculty and Ph.D. Supervisor with KTH. He has more than 60 patents and more than 60 refereed journal papers and more than 50 conference papers.

Dr. Pang is a Co-Chair of the Technical Committee on Industrial Informatics with the IEEE. He is also an Associate Editor for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, and IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN INDUSTRIAL ELECTRONICS, a Guest Editor for PROCEEDINGS OF THE IEEE, IEEE INTERNET OF THINGS JOURNAL, IEEE REVIEWS IN BIOMEDICAL ENGINEERING, etc. He was an Invited Speaker at the Gordon Research Conference on Advanced Health Informatics, and the General Chair of the 2017 IEEE International Conference on Enterprise Systems. He was the recipient of the 2016 Inventor of the Year Award and 2018 Inventor of the Year Award by ABB Corporate Research Sweden, which is the only award for individuals and only one winner per year out of 300 researchers.

Hong Wen (Senior Member, IEEE) received the M.Sc. degree from the Sichuan University, Chengdu, China, in 1997, and the Ph.D. degree from Southwest Jiaotong University, Chengdu, in 2004, and the second Ph.D. degree from the University of Waterloo, Waterloo, ON, Canada, in 2018, all in electrical and computer engineering.

From 2008 to 2009, she was a Visiting Scholar and a Postdoctoral Fellow with the Electrical and Computer Engineering Department, University of Waterloo. She is currently a Professor with the Department of Aeronautics and Astronautics, University of Electronic Science and Technology China, Chengdu. Her current research interests include communication systems and security.



Kan Yu (Member, IEEE) received the Bachelor of engineering degree in communication engineering from Beijing University of Posts and Telecommunications in China, in 2005, the Master of engineering degree in communication engineering from Chalmers University of Technology in Sweden, in 2010 and the Ph.D. degree in computer science from Malardalen University in Sweden, in 2014.

He is currently a Lecturer of Internet-of-Things (IoT) with La Trobe University, Bundoora, Australia. His current research interests include industrial Internet-of-Things, reliable and low latency industrial wireless communications, and real-time industrial communications.

Tengyue Zhang (Student Member, IEEE) is currently working toward the Ph.D. degree in communication and information system with the School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu, China. Her current research interests include physical layer security in mobile communications systems and multi antenna systems security.

Yueming Lu received the B. S. and M. S. degrees in computer science from Xi'an University of Architecture and Technology, Xi'an, China, in 1994 and 1997, respectively, and the Ph. D. degree in computer architecture from Xi'an Jiaotong University, Xi'an, in 2000.

He is currently a Professor with the Beijing University of Posts and Telecommunications, Beijing, China. His research interests include security control, evaluation, and data protection.