applied sciences

MDPI

Article

Impersonation Attack Detection in Mobile Edge Computing by Levering SARSA Technique in Physical Layer Security

Xiaodan Yan [1], Ke Yan [2,*], Meezan Ur Rehman [3] and Sami Ullah [4]

[1] School of Cyber Science and Technology, Beihang University, Beijing 100191, China
[2] Department of the Built Environment, National University of Singapore, Singapore 117566, Singapore
[3] Department of Electrical Engineering, CECOS University, Peshawar 25000, Pakistan
[4] Department of Computer Science, Shaheed Benazir Bhutto University, Sheringal 18050, Pakistan
* Correspondence: yanke@nus.edu.sg

Abstract: Smart health systems typically integrate sensor technology with the Internet of Things, enabling healthcare systems to monitor patients. These biomedical applications collect healthcare data through remote sensors and transfer the data to a centralized system for analysis. However, the communication between the edge node and the mobile user is susceptible to impersonation attacks in mobile edge computing (MEC) for the biomedical application. For this purpose, we propose a detection mechanism for medical and healthcare services, i.e., reinforcement learning for impersonation attacks. We construct a system model of MEC, a key generation model (KGM), and an impersonation attack model (IAM). In addition, we also design an impersonation attack detection algorithm based on the SARSA technique under the IAM. In our proposed work, the SARSA-based method outplays the detection of impersonation attacks in the dynamic environment compared to the traditional Q-learning technique. Finally, we evaluate the false alarm rate (FAR), miss detection rate (MDR), and average error rate (AER) in the hypothesis tests to compare the performance of our proposed method with the traditional Q-learning. In comparison to the classic Q-learning based technique, simulation experiments show that the suggested approach can avoid impersonation attacks in a dynamic environment for medical and healthcare services. The results also indicate that the SARSA technique has a high detection accuracy and low average error rate compared to the conventional Q-learning based approach.

Keywords: network security; encrypted traffic identification; deep learning; convolutional neural network

Citation: Yan, X.; Yan, K.; Rehman, M.U.; Ullah, S. Impersonation Attack Detection in Mobile Edge Computing by Levering SARSA Technique in Physical Layer Security. *Appl. Sci.* **2022**, *12*, 10225. https://doi.org/10.3390/app122010225

Academic Editors: Thi-Thu-Huong Le and Howon Kim

Received: 27 September 2022
Accepted: 10 October 2022
Published: 11 October 2022

1. Introduction

With the drastic advancement of biomedical applications and digital healthcare, many secure wireless communication problems arise while enhancing the information processing rate. Distributed systems in healthcare services increasingly require increasing demand for high data rates and real-time communication. Moreover, cloud computing can no longer meet heterogeneity, low delay, and other network requirements [1–3]. With this context, mobile edge computing (MEC) came into being that extends cloud computing to the edge of the network [4–6]. It is used to enhance the system efficiency and solve the problems of weak geographic information perception ability, mobility, and provide low delay [7].

Nevertheless, MEC is more vulnerable to impersonation attacks by malicious users. In an impersonation attack, an illegal edge node or mobile user pretends to be the legitimate node and sends data to other network nodes for illicit gains. Traditional security solutions, such as access control, key management, digital signature, and identity authentication, require high complexity, maintenance, and management. The security of traditional authentication technology mainly depends on the mathematical computation complexity of an encryption algorithm, but with the development of quantum computers, this technology is facing the risk of being breached. Moreover, the traditional authentication technology

has great limitations in the MEC environment. The computing resources of each node are limited and cannot satisfy the complex encryption and decryption algorithms.

Several mobile users in the fields of healthcare and biomedical applications communicate with edge nodes through wireless channels in the MEC environment. However, the open characteristic of wireless network makes it highly vulnerable to impersonation attack. Malicious users forge legitimate user identities to obtain administrative frame or control frame messages, and then launch other malicious attacks. Due to wireless media, the communication between edge nodes and mobile users can face numerous security challenges such as a man-in-the-middle attack, impersonation, hijacking, and eavesdropping attack. As the wireless network environment becomes dynamically complex and unknown, it is difficult to obtain the channel model or parameters, and it is difficult for the receiver to select an appropriate detection threshold for sender identity authentication. The impersonation attack detection method based on unsupervised learning in a physical layer can learn the optimal detection threshold in an unknown communication environment and improve detection performance, so it is more suitable for the development of wireless communication in the future. By introducing a reinforcement learning algorithm, the best detection threshold can be obtained by adaptive spoofing attack detection using current perception information in a dynamic unknown environment. The main advantages of Q-learning-based methods over other algorithms are optimality effects and generality. At present, most studies on impersonation attack detection based on deep reinforcement learning are based on discrete detection threshold analysis, while the continuous control of the detection threshold based on deep reinforcement learning lacks relevant literature research, and the related detection performance needs to be analyzed. In an MEC environment, physical layer authentication is usually used to protect the secure communication of mobile users. The authentication scheme mainly includes device-based features and channel-based features.

In recent years, researchers have taken advantage of physical layer characteristics and have relied on signal processing and encoding modulation to secure communication. This can help us to establish a secure transmission between mobile users and edge nodes. Remarkably, the physical layer key generation technology can directly realize the key generation between two communicating parties, eliminating key management center and key distribution procedures. The physical layer security (PLS) utilizes code modulation, key generation, key consistency, and encryption/decryption to ensure secure information by enhancing the security performance of MEC. Based on the PLS, the impersonation attack detection (IAM) algorithm is proposed in this work. In summary, the main contributions of this paper are given as follows:

- An impersonation attack model (IAM) in MEC is constructed. The model assumes that attackers (edge nodes or mobile users) can perform active impersonation attacks, meeting the subsequent experiments' requirements.
- According to the established model and hypothesis test, an IAM is proposed to optimize the threshold and achieve impersonation detection in a dynamic environment. We propose a SARSA algorithm based on reinforcement learning that realizes the detection action and can defend against the impersonation attack in a dynamic environment.
- We analyze the miss detection rate (MDR), false alarm rate (FAR), and an average error rate (AER) with the SARSA algorithm and compare to the traditional Q-learning method.
- The experimental results show that our proposed scheme has a higher detection accuracy and lower AER, effectively preventing the impersonation attack compared to the traditional Q-learning-based approach.

The rest of the work is structured as follows. Following the introduction, we outline the related work in the next section. After the related work, the system model is illustrated, which is further subdivided into the impersonation attack model and security model. The system model is followed by the hypothesis test and detection method, which includes

the hypothesis test and impersonation detection method. Afterward, we describe the simulation experiments that include the receiver utility, receiver's test threshold, analysis of MDR, analysis of FAR, and analysis of AER. In final section, the paper is concluded.

2. Related Works

MEC can provide mobile users with a variety of services, including storage, computing, and data transmission. However, when sending data using an open wireless network, there will be a denial of service (DoS), interference, eavesdropping, data leakage, and other security issues. To address these issues, researchers have proposed a variety of solutions. For instance, in [8], the authors proposed an edge computing data protection model, enabling edge devices in other regions to share and securely access the resources. However, the computational complexity is higher in the proposed method. Ullah et al. [9] suggested an access tree design and attribute-based signature technology to achieve ciphertext updating and computing outsourcing in MEC. Nevertheless, there are still security vulnerabilities, and attackers can effortlessly impersonate legitimate mobile users or edge nodes. Alrawais et al. [10] proposed an encrypted key exchange (EKE) protocol that combined the CP-ABE method and digital signature technology to achieve security goals such as identity verification, confidentiality, and access control. Similarly, Paharia et al. [11,12] investigated a defense architecture including an edge layer to preclude DoS attacks between mobile users and cloud computing, improving network performance. However, the authors fail to consider the security threat between mobile users and the edge layer.

Artificial intelligence technology can adapt to attack detection in dynamic unknown environments by using current perceptual information for trial-and-error learning. The detection method based on reinforcement learning algorithms can obtain higher detection performances in an unknown communication environment, so it is more suitable for the development of future wireless communication. Pan et al. [13] proposed a physical layer authentication method based on a machine learning and channel-based scheme. Only in specific industrial scenarios can the scheme show strong security and low computing energy consumption. Xie et al. [14] comprehensively introduced the technologies and theories related to physical layer authentication in wireless communication.

In various wireless systems, malicious users pose security threats to other users through disguised attacks. He pointed out that wireless communication leads to more security vulnerabilities due to its open nature and put forward some suggestions for current wireless systems. There are two main types of physical-layer-oriented authentication technologies. The first is to embed the identity information into the transmission signal, through the identity information for authentication [15–17]. Another type of approach is to take advantage of physical layer properties inherent in wireless transport for device authentication [18,19]. General wireless security schemes generally rely on encryption algorithms and protocol layer authentication, but there is no feasible scheme for security threats of physical layer vulnerabilities and interfaces. Therefore, Hou et al. [19] used the time-varying carrier frequency offset feature in the physical layer authentication process, and designed an adaptive device identification method.

The physical layer key generation (PKG) method can generate a key having characteristics of reciprocity, uniqueness, and randomness of channel state information (CSI). Both the transmitter and receiver use a common key to send and receive information. The works in [20–23] presented a key PKG method, which is an important research content in wireless system security, achieving secure communication. For instance, Eberz et al. proposed a secret key generation (SKG) scheme in [24], which can prevent man-in-the-middle (MITM) attacks by encrypting and authenticating the wireless channel. Chen et al. [25] proposed an effective channel information extraction algorithm to avoid interference and noise that could reduce the key disagreement rate (KDR), achieving the goal of improving SKG. Sudarsono et al. [26] studied the key generation scheme and extracted a secret key bits stream from the received signal strength (RSS) in wireless network communication, which caused the KDR to lower. However, the environment in MEC is dynamic. The mobile

user's constant movement leads to changes in the channel, affecting the SKG. Hence, it is difficult for the receiver to determine the legality of the signal in a dynamic environment. To that end, this paper proposes a reinforcement learning algorithm for impersonation attack detection (IAD) to obtain the optimal strategy in a dynamic environment.

3. System Model

To study the IAD problem, we design an IAM between the edge and mobile user layers. A security model based on the CSI is considered a reference for the detection method. In the proposed scheme, we begin by demonstrating the concept of IAM and then move on to discuss the security model.

3.1. Impersonation Attack Model (IAM)

The IAM proposed in our work is illustrated in Figure 1. This model is composed of several mobile users and edge nodes. In a MEC environment, the users (nodes) communicate with each other through wireless networks. We consider x legitimate nodes (edge node or mobile user) and y illegitimate nodes (edge node or mobile user). In our proposed network, the total number of nodes (legitimate + illegitimate), including mobile users and edge nodes, is represented by N^*. However, N^* is further divided into the set of X and Y to differentiate legitimate and illegitimate mobile users and edge nodes, respectively. Hence, the set of all the legitimate nodes including edge nodes and mobile users is represented by X, which is given by

$$X = \{1, 2, 3, \ldots, x\}, x \in N^*, \tag{1}$$

and the set of all the illegitimate nodes, including edge nodes and mobile users, is represented by Y, which is given by

$$Y = \{1, 2, 3, \ldots, y\}, y \in N^*. \tag{2}$$

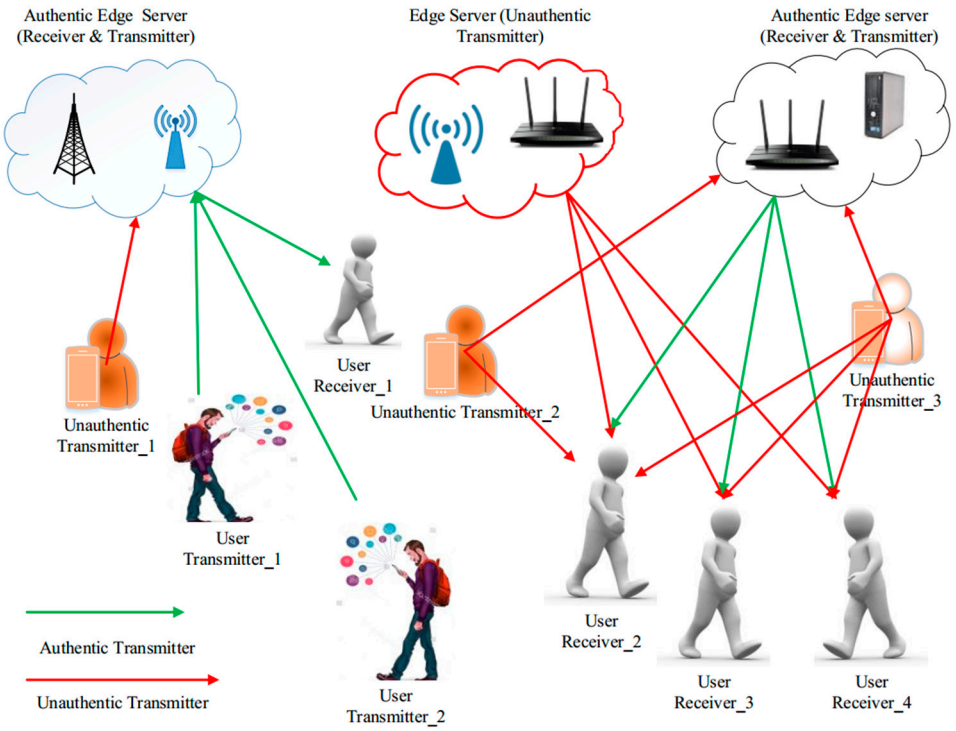

Figure 1. IAD between MEC servers and mobile users to detect authentic and unauthentic transmitter.

In summary, the illegitimate users/nodes are those users/nodes that pretend to be legal users/nodes and send data to legitimate nodes/users. Hence, the impersonation attack is assumed by such behavior of illegitimate nodes.

3.2. Security Model

As shown in Figure 2, in order to ensure the wireless communication security between edge nodes and the end-user, the PKG method is used to generate the secret keys [20–23]. This model assumes that the edge node and the end-user node use a time-division duplex (TDD) for communication purposes. Initially, the nodes send multiple training signals to each other in a particular time slot and obtain channel characteristics such as CSI during the channel estimation phase. It is necessary to determine whether the received CSI is legal or not before quantization. Thus, IAD is introduced. If it is an impersonation attack signal, the proposed model will estimate the channel and obtain CSI. Otherwise, it will conduct de-correlation and re-sampling on the legitimate CSI. Afterward, it performs the quantization step to quantify the sample value into the binary initial key bit sequence.

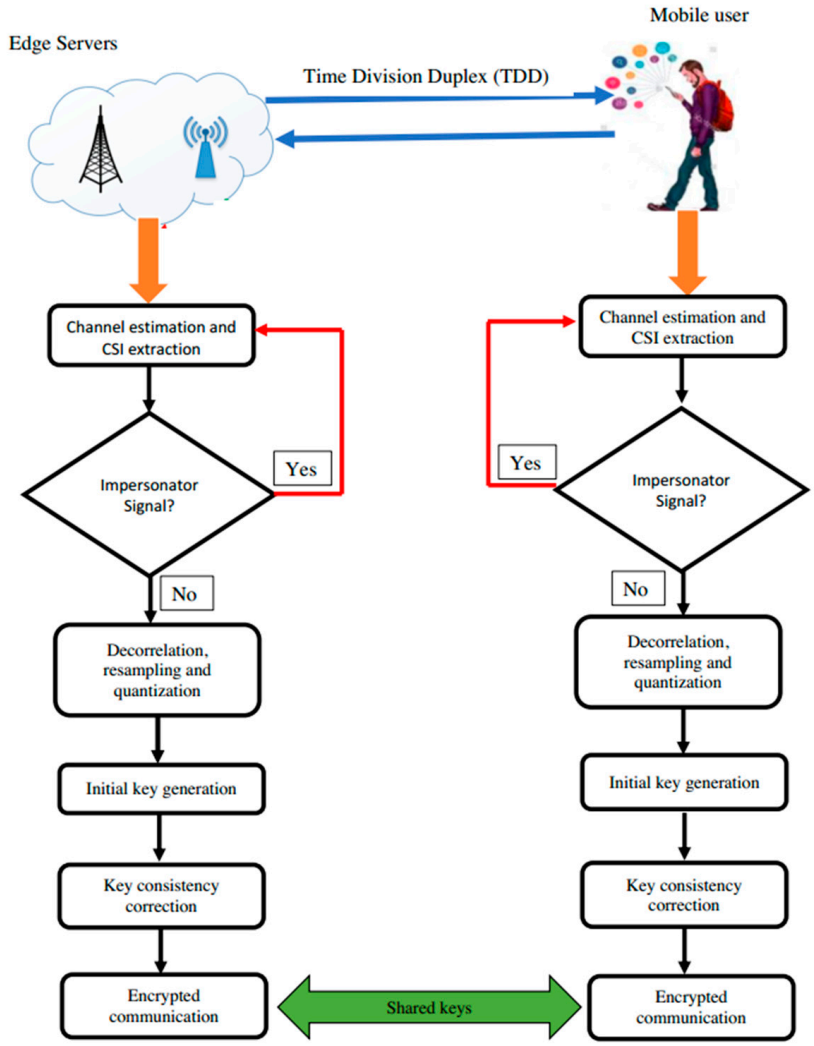

Figure 2. The key generation model.

The physical layer key generation method includes the following steps:

- The transmitting and receiving nodes send training signals to each other for channel estimation to obtain CSI and determine whether they are impersonator signals or authentic signals.
- The legitimate signals received are quantified.
- A part of the quantized data are used as the initial key.
- The two parties (sender and receiver) perform key consistency negotiation to correct the quantization-mismatched output bits, thereby obtaining the same key to communicate.

The filtered signals are transferred to the upper layer because the impersonation attack is roughly detected at the physical layer by building the security model. Hence, the communication and computation overhead is significantly reduced. Consequently, the overall system performance is improved in our proposed method.

4. Hypothesis Test and Detection Method

By building the IAM in the previous section, we need a hypothesis test to be established at the receiving end. Consequently, the detection method based on the SARSA algorithm is explained in the following subsections.

4.1. Hypothesis Test

From IAM, it is known that the edge node and mobile users estimate their CSI and extract the channel vector after receiving the training signals in the channel estimation stage. Therefore, the transmitted training signals received by the receiver is known as the channel vector and is denoted by β_a^b. Furthermore, each channel vector received by the receiver is recorded as a channel record donated by γ_a^b. In addition, the $W\left(\beta_a^b\right)$ is the channel gain of the nel vector of the training signal sent by the legal nodes, and H_1 represents the channel vector of the training signal sent by the illegitimate nodes. Thus, the hypothesis test is established as follows.

$$H_0 : W\left(\beta_a^b\right) \geq Z_{(E,F)}, \tag{3}$$

$$H_1 : W\left(\beta_a^b\right) \geq Z_{(E|F,I)}, \tag{4}$$

where E, F, and I represent the edge node, end-user, and impersonation attacker, respectively. $Z_{(E,F)}$ denotes the estimated channel gain of the legitimate mobile user, and $Z_{(E|F,I)}$ is the estimated channel gain of the potential impersonation attacker. CSI is unique and represents the channel's state. Therefore, β_a^b and γ_a^b are also uniquely determined. On this basis, the statistic of the hypothesis test is

$$D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right) = \frac{\| \left(\beta_a^b\right) - \left(\gamma_a^b\right) \|^2}{\| \left(\gamma_a^b\right) \|^2}. \tag{5}$$

where $\| \cdot \|$ is the Frobenius norm and D is the normalized Euclidean distance between β_a^b and γ_a^b. We compare $D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right)$ with the test threshold, denoted by λ. If $D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right) < \lambda$, it is considered as a legitimate user; otherwise, it is considered to be an attacker. The hypothesis test from Equations (3) and (4) can be expressed by

$$D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right) < \lambda \Rightarrow H_0, \tag{6}$$

and

$$D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right) > \lambda \Rightarrow H_1. \tag{7}$$

By establishing the above hypothesis test to detect impersonation attack and comparing $D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right)$ with the test threshold λ, we can judge whether each training signal received by the receiver is legitimate. It is assumed that the accuracy of the IAD is related to the threshold value. The larger the threshold value, the higher the MDR, whereas the smaller the threshold value, the higher the FAR. Therefore, the impersonation detection method is introduced in the following subsection.

4.2. Impersonation Detection Method

The Q-learning algorithm is an off-policy algorithm, whereas the SARSA algorithm is an on-policy algorithm [27]. Both of them are reinforcement learning algorithms that optimize the IAD strategy to a certain extent in a dynamic environment. Using continuous

learning, the proposed method selects an optimal test threshold to detect impersonation attacks. According to the SARSA algorithm, the receiving ends choose specific suboptimal actions on each state with the probability of ε based on the ε-greedy strategy. In the ε-greedy strategy, the receiver perceives the state in each time slot, which is denoted by L, and detects the legitimate and illegitimate nodes in each state. Upon perceiving the state, it performs action randomly with a probability of ε and decides the optimal action with a $1-\varepsilon$. In short, the probability of choosing a more significant action value is $1-\varepsilon$. Hence, the probability is given as

$$P(\lambda) = \begin{cases} 1-\varepsilon, \lambda = \lambda^*, \\ \frac{\varepsilon}{L}, \lambda \in \{l/L\}, \lambda \leq l \leq L, \lambda \neq \lambda^*, \end{cases} \tag{8}$$

where λ represents the test threshold value and is selected from $L+1$ states, i.e., $\lambda \in \{l/L\}$. In each state, regardless of which action is chosen, the immediate reward can be obtained as

$$R = \sum_{t=(\tau-1)T+1}^{\tau T} U_N^t(\lambda, \kappa), \tag{9}$$

where T shows the number of transmitted signals in a time slot, U_N^t is the immediate utility function, and κ is a set of training signals transmitted by illegitimate nodes. Similar to [27], the Bayesian risk of impersonation detection, $I(\lambda, \kappa)$, can be expressed as

$$\begin{aligned} I(\lambda,\kappa) = (g_0(1-P_1(\lambda)) - c_0 P_1(\lambda))(1-\textstyle\sum_{i=1}^{F} p_i) + \\ (g_1(1-P_2(\lambda)) - c_1 P_2(\lambda))\Big(\textstyle\sum_{i=1}^{F} p_i\Big). \end{aligned} \tag{10}$$

The first part of Equation (10) represents the gain of legal users, while the second part represents the gain of illegitimate users. Here, g_0, g_1 represent the gain of accepting the legal packet and the gain of rejecting the illegal packet, respectively. Likewise, c_0, c_1 show the cost of rejecting the legal packet and the cost of receiving the illegal packet, respectively. If R_0 is the immediate utility gained by accepting the legal packet and R_1 is the primary utility gained by rejecting the illegal packet, then Equation (10) can be rewritten as

$$\begin{aligned} R_0(\lambda,\kappa) = (g_1 - g_0) \quad & \textstyle\sum_{i=1}^{F} p_i - (g_1+g_0) P_2(\lambda) \textstyle\sum_{i=1}^{F} p_i - \\ & (g_1 + c_0) P_1(\lambda)(1-\textstyle\sum_{i=1}^{F} p_i) + g_0, \end{aligned} \tag{11}$$

and

$$\begin{aligned} R_1(\lambda,\kappa) = (g_0 - g_1\ ) \quad & \textstyle\sum_{i=1}^{F} p_i + (g_1+g_0) P_2(\lambda) \textstyle\sum_{i=1}^{F} p_i + \\ & (g_1 + c_0) P_1(\lambda)(1-\textstyle\sum_{i=1}^{F} p_i) - g_0 \end{aligned} \tag{12}$$

In the SARSA algorithm, the discount rate $\eta, \eta \in (0,1)$ determines the importance of future rewards, whereas the learning rate, $\alpha, \alpha \in (0,1)$, means the retention of previous learning outcomes. Hence, the updated value at the receiver side can be obtained as follows.

$$Q(s_t, \lambda) = (1-\alpha) Q(s_t, \lambda) + \alpha\big(R + \eta Q\big(s_{t+1}, \lambda'\big)\big) \tag{13}$$

and

$$V(s_t) = \mathrm{m}\ Q(s_t, \lambda), \lambda \in \{l/L\}, \lambda \leq l \leq L \tag{14}$$

The optimal test threshold is:

$$\lambda^* = \arg\mathrm{m}\ Q(s_t, \lambda), \lambda \in \{l/L\}, \lambda \leq l \leq L \tag{15}$$

The SARSA algorithm with the IAD method is detailed in Algorithm 1.

Algorithm 1 SARSA-enabled impersonation attack detection (IAD) method

(1) **Initialize** $\varepsilon, \eta, \alpha, Q(s_t, \lambda) = 0$
(2) **Select** λ in the current state $s_t, t = 1, 2 \ldots$
(3) Obtain the channel state information, CSI of the training signals, channel record, channel vector, and calculate the estimated channel gains $Z_{E,F}$ and $Z_{E|F,A}$
(4) Calculate Euclidean distance $D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right)$. If $D\left(\left(\beta_a^b\right), \left(\gamma_a^b\right)\right) \leq \lambda$, then $\gamma_a^b \leftarrow \beta_a^b$, and accept the training signal,
(5) **else**, reject the signal;
(6) **Repeat** steps 3 and 4 until the receiver can handle T training signals received in a given time slot;
(7) Go to the next state s_{t+1}, compute R using Equation (9), and update $Q(s_t, \lambda)$ using Equation (13);
(8) Return step 2 and repeat step 2–6 until λ is optimal.

5. Simulation Experiment

In this section, we investigate the performance of our proposed SARSA algorithm. The receiver end achieves the analysis by analyzing the receiver utility, FAR, MDR, and AER.

5.1. Receiver's Test Threshold

Figure 3 illustrates the change in test threshold as a function of the number of experiments. It can be observed from the figure that the test threshold of IAD changes with the increasing number of experiments. However, to realize Figure 3, we notice that in the first 100 experiments, the test threshold of IAD changes rapidly, and then it stabilizes (same pattern) after 100 runs.

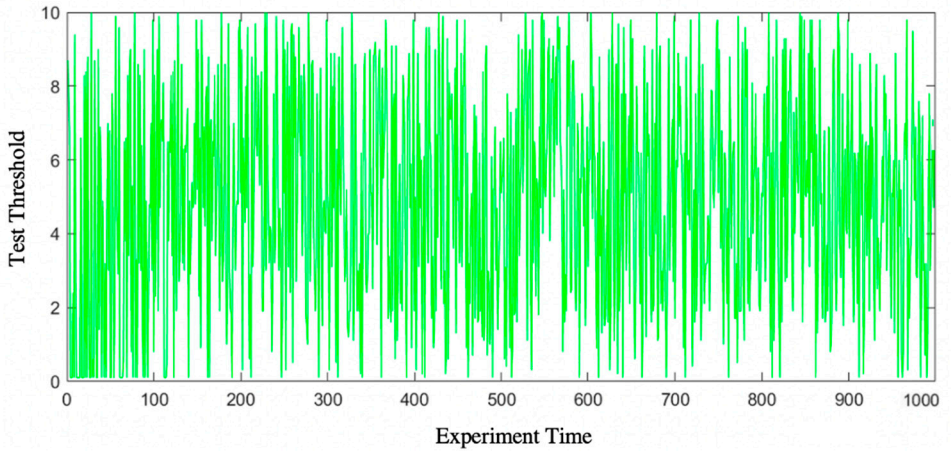

Figure 3. Test threshold of IAD changes with the increasing number of experiments.

5.2. Receiver Utility

In several experiments, it can be seen in Figure 4 that in a scenario where the legitimate user's channel gain ratio k is the same as the attacker's channel gain, the greater is the average utility of the receiving end for smaller legal training sample ρ. The figure reveals that for 1000 experiments, the average return value at $\rho = 10$ is 2.4238, which is a 17% increase compared to the average utility value at $\rho = 20$. The result depicts that a smaller SINR value of the training sample yields a better experimental result. The accuracy of IAD is improved by taking into account the SINR of legal training samples. According to the hypothesis test, λ influences the detection accuracy, i.e., FAR and MDR. Thus, the FAR and MDR can be expressed as

$$P_1(\lambda) = 1 - \mathcal{F}^2_{\chi^2_{2\mathcal{M}}}\left(\frac{2\theta\lambda}{2\theta^2 + \omega\theta\delta^2}\right), \tag{16}$$

and

$$P_2(\lambda) = \mathcal{F}^2_{\chi^2_{2\mathcal{M}}}\left(\frac{2\theta\lambda}{2\theta^2 + (1+q)\theta\delta^2}\right), \tag{17}$$

where $\mathcal{F}^2_{\chi^2_{2\mathcal{M}}}$ denotes the cumulative distribution function having $2\mathcal{M}$ degrees of freedom. ω is the relative variation rate in the channel gain and δ^2 is the average power gain of the received signal. θ is the SINR of the legal training sequence and q is the ratio of the channel gain between authentic and unauthentic nodes. Furthermore, based on Equations (16) and (17), the computation of AER is expressed as

$$P_{AER} = P_1(\lambda) + P_2(\lambda). \tag{18}$$

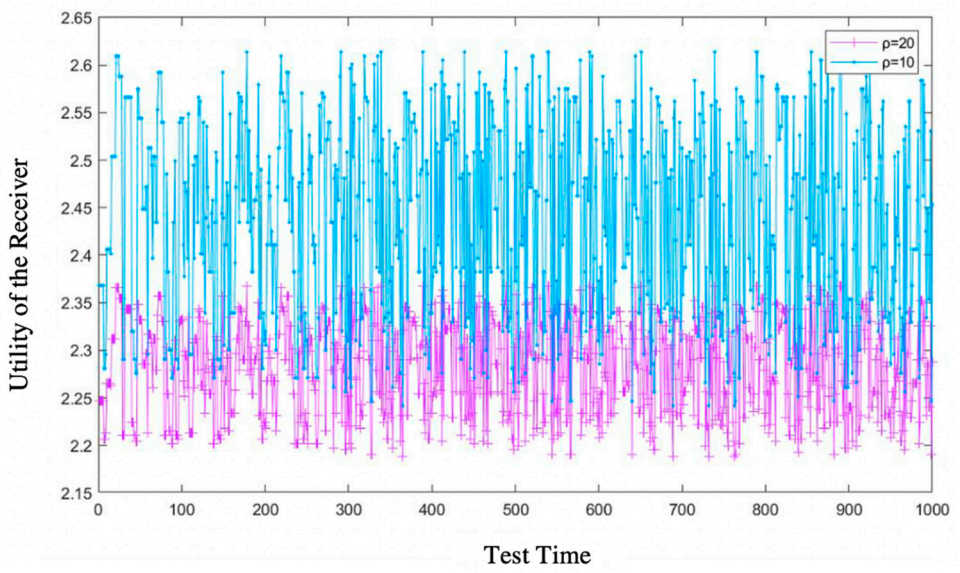

Figure 4. When $k = -3$, the change in the receiver gain is tested with the increase in the number of experiments at $\rho = 20$ and $\rho = 10$, respectively, and the experiment is tested 1000 times.

In our proposed SARSA-based method, we aim to evaluate the performance of AER, MDR, and FAR. The Q-learning algorithm is also compared with the impersonation attack model based on the SARSA algorithm. Therefore, we need to analyze the AER, MDR, and FAR in comparison to the traditional Q-learning algorithm. We consider a simulation area of 20×20 square meters. We assume that the number of nodes in the simulation area is distributed randomly. Similar to [3], all channel gains follow the normal distribution $\varsigma(0, 1)$. Moreover, we assume that T consists of 20 training signals in one time slot. For the analysis of FAR, MDR, and AER, the initial values in the simulation environment are set to be $\theta = 10$ dB, $\omega = 3$ dB, $q = 0.2$, and $\delta^2 = 5$.

In our recommended SARSA-based technique, we intend to investigate the FAR, MDR, and AER. The Q-learning method is evaluated compared to the IAM based on the SARSA algorithm. As a result, we need to examine the FAR, MDR, and AER compared to the conventional Q-learning approach. The simulation environment is set-up with an area of 20×20 square meters, and the number of nodes is randomly scattered. All channel gains follow the normal distribution $\varsigma(0, 1)$ [13]. It is assumed that there are $T = 20$ training signals in one slot, and the center frequency of the signals is $f_0 = 2.4$ GHz, $g_0 = 6$, $g_1 = 9, c_0 = 2$, and $c_1 = 4$. The parameter values used in the SARSA algorithm are $\varepsilon = 0.1$, $\alpha = 0.6$, and $\eta = 0.8$. The initial values are given to calculate FAR, MDR, and AER are $\theta = 10$ dB, $b = 3$ dB, $q = 0.2$, and $\delta^2 = 5$.

5.3. Analysis of FAR

Figure 5 shows that FAR will yield a lower value for the smaller legal training sample ρ in a scenario where the legitimate user's channel gain ratio k is the same as the attacker's

channel gain. The reason for this is that as the SINR decreases, the channel estimation error (CER) decreases. The CER has a direct impact on the signals' quantization quality, which seriously influences the subsequent calculation of the FAR. It can be observed from the two curves in Figure 5 that with different values of ρ, the lowest value is present at $x = 9$. This is due to the fact that the Q-learning algorithm slowly optimizes the threshold in order to find the optimal threshold. As the threshold value changes from $x = 1$ to 9, the FAR changes accordingly. The slope of the FAR progressively becomes smaller when the threshold approaches the optimal threshold. However, the slope of the FAR tends to stabilize after 400 training sessions. For $\rho = 20$, the FAR remains stable at about 0.225%, which is 13% greater than $\rho = 10$.

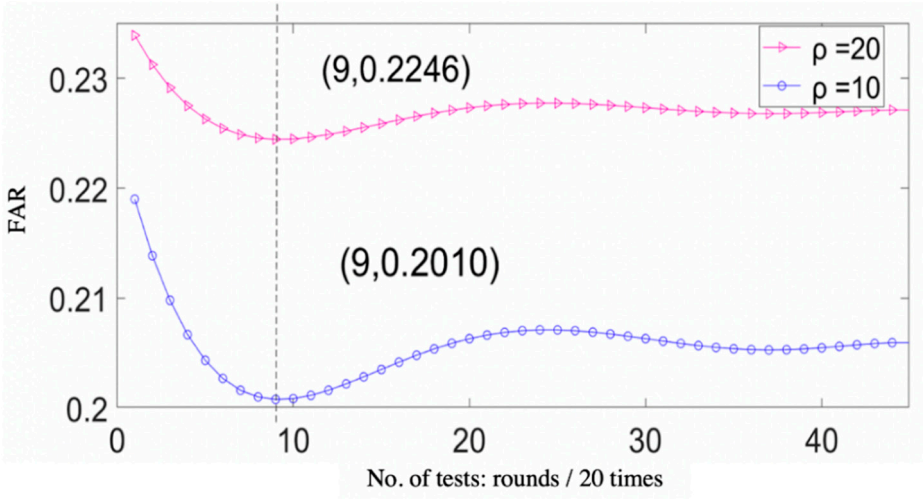

Figure 5. When $k = -3$, FAR at $\rho = 20$ and $\rho = 10$ is tested as the number of experimental rounds increases, and the Q table is updated 20 times per round.

Similarly, the FAR is analyzed between the SARSA-based detection and Q-learning method. Figure 6 depicts that the threshold changes constantly with the increase in the number of training. Initially, the receiver has no information about the system environment; therefore, the FAR decreases according to Equation (16). Nevertheless, during the training phase, the receiver employs past experience detecting impersonation attacks to choose the optimal threshold by continuously testing the environment. Thus, the optimal threshold gradually stabilizes after reaching a number of experiments. Moreover, the dynamic wireless channel between mobile users and edge nodes causes signal fading, resulting in some differences in the channel vector. As a result, the FAR slightly varies but remains stable at 20.0–20.6%. Mostly, the FAR of Q-learning algorithms is higher than that of the SARSA algorithm. This is due to the SARSA conservative nature that finds a suitable threshold via a continuous trial-and-error process. Hence, when detecting an impersonation attack, the SARSA algorithm has a lower chance of misinterpreting a legal signal as an illegal signal, resulting in a lower rejection. As a result, it accepts more legal signals at the receiving side and improves the SKG between the communicating parties.

5.4. Analysis of MDR

As illustrated in Figure 7, when the attacker's channel gain is equal to the legitimate user's channel gain, the value of MDR is lower due to higher training samples ρ of the legal training sample ρ. Conversely, for different values of the attacker and legitimate user's channel gain, ρ has the highest points $(9, 0.2916)$ and $(9, 0.2735)$, respectively. This is due to the fact that a smaller SINR results in smaller CERs. Thus, it has a direct impact on the signal's quantization accuracy and poses diverse effects on the subsequent calculation of MDR. It can be observed from the two curves in Figure 7 that with different values of ρ, the highest value is presented at $x = 9$. This is due to the fact that the Q-learning algorithm slowly adjusts the threshold in order to find the optimal threshold. As the threshold value

rapidly changes during $x = 1$ to 9, the MDR changes accordingly. The slope of the MDR progressively becomes smaller and smaller when the threshold approaches the optimal threshold. However, the slope of MDR tends to stabilize after 400 training sessions. For $\rho = 20$, the MDR remains stable at about 0.2735%, which is 6.2% lower than that for $\rho = 10$.

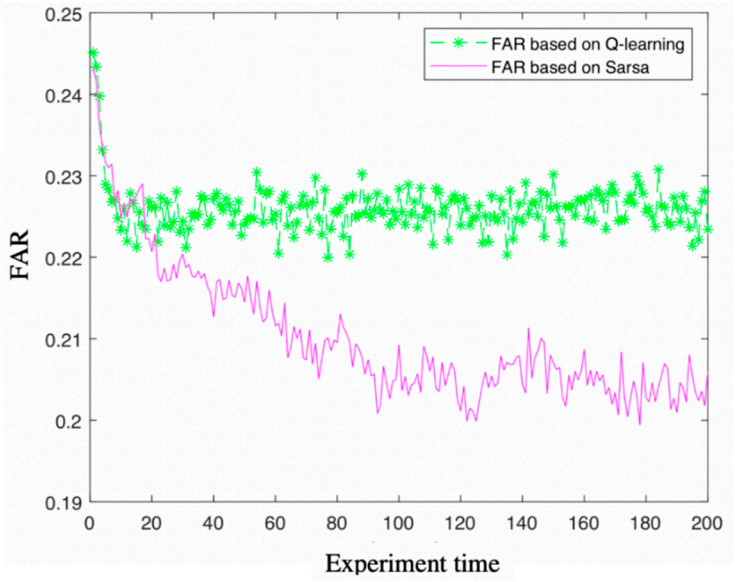

Figure 6. The comparison of FAR based on SARSA and Q-learning.

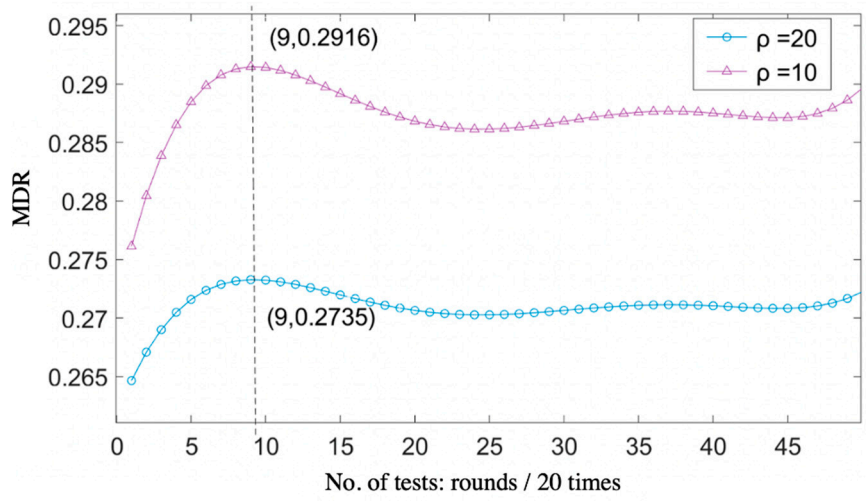

Figure 7. When $k = -3$, the MDR at $\rho = 20$ and $\rho = 10$ is tested separately. The Q table is updated 20 times per round as the number of experimental rounds increases.

Moreover, the impersonation attack in the simulation environment is unknown at the beginning, and the chosen threshold at the receiver end is small. However, as the environment is explored and the number of training sessions increases, the test threshold gradually rises. According to Equation (17), MDR is increased. This is due to the acquired experience from past learning. In this manner, the increasing number of training gradually stabilizes the optimal threshold. Nonetheless, the channel vector varies from experiment to experiment. As a result, the MDR ranges from 28.5% to 29.0%. Figure 8 reveals that the SARSA algorithm outperforms Q-learning by 4%–5% in terms of MDR. This is because the SARSA algorithm is more vigilant than the Q-learning algorithm. There is also a 10% chance of selecting the detection threshold randomly, which means that some illegal signals may pass undetected. Consequently, the MDR of the SARSA algorithm is slightly higher than Q-learning.

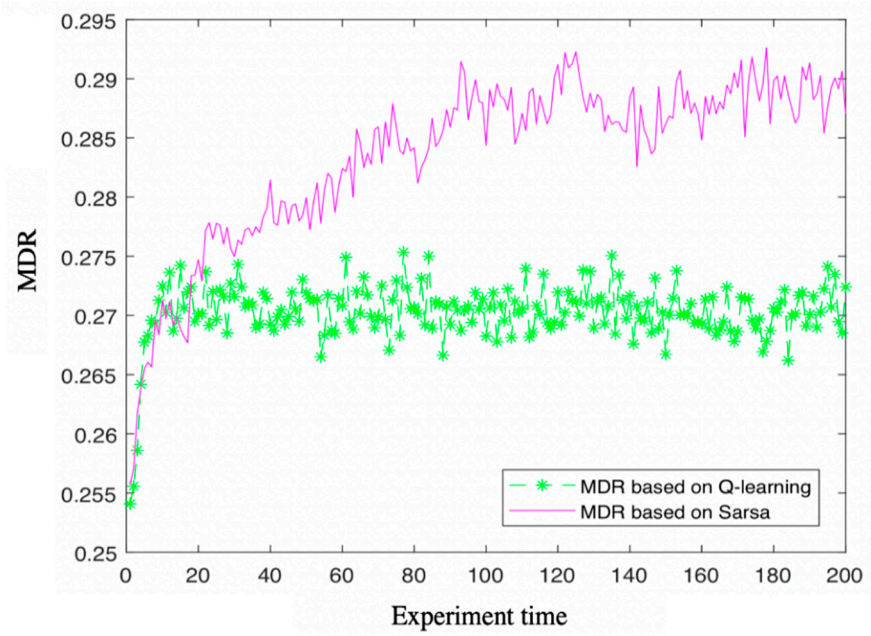

Figure 8. The comparison of MDR based on SARSA and Q-learning.

5.5. Analysis of AER

Figure 9 depicts that a lower SINR can result in a lower AER. After 400 experiments, the AER is stabilized at around 0.4923% when $k = -3$ and $\rho = 10$. The reason for the high AER is that detailed simulation in the signal generation time is not performed in the initial stage of the simulation experiment. However, a random noise generation source and a simple random signal generation source are employed, and the subsequent signal processing is used. Furthermore, the signal that needs to be analyzed is not sufficiently pre-processed, resulting in a higher AER.

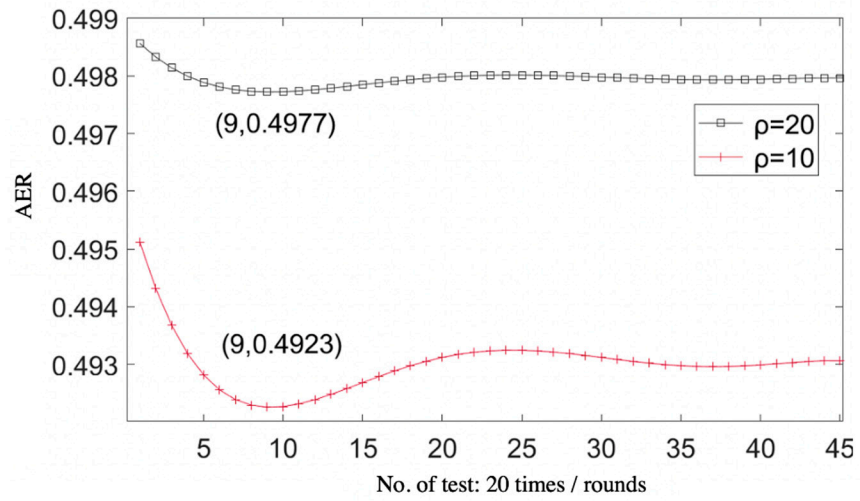

Figure 9. When $k = -3$, the AER at $\rho = 20$ and $\rho = 10$ are tested separately. The Q table is updated 20 times per round as the number of experimental rounds increases.

Similarly, as the number of experiments grows, the training signals are assessed and constantly received at the receiver, allowing the optimal threshold for detecting attacks with rich experience to be determined. Finally, the optimal threshold stabilizes at a particular stage, and the AER remains essentially constant. Figure 10 depicts a comparison of AER based on the two algorithms. It can be observed from the figure that AER based on the SARSA algorithm is lower than that of Q-learning, implying that the SARSA algorithm

detects legal and illegal signals at the receiving end. Hence, we can summarize that the SARSA-based method can improve detection accuracy while reducing the risk of impersonation attacks.

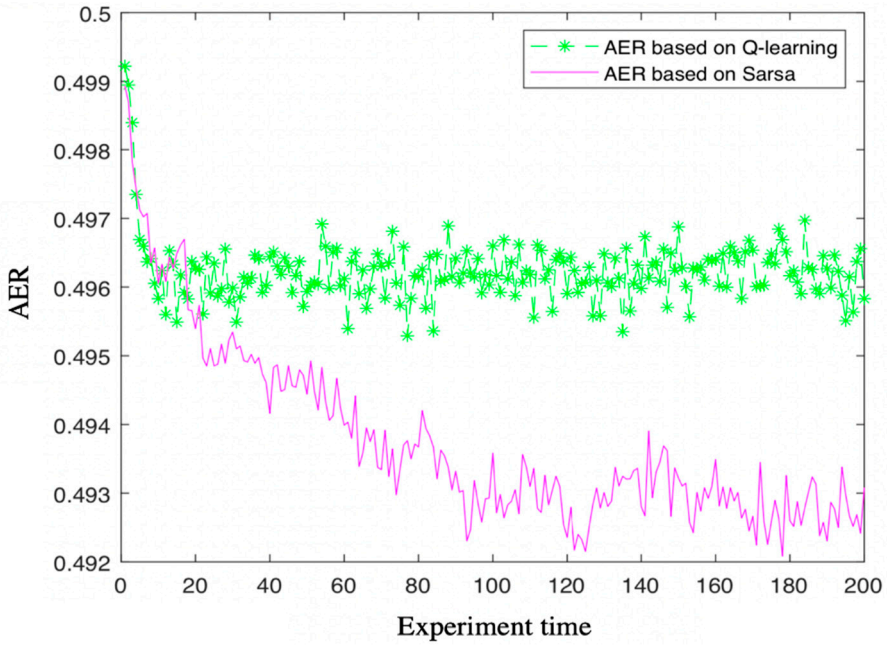

Figure 10. The comparison of AER based on SARSA and Q-learning.

In summary, we compared the result of our proposed SARSA algorithm with the traditional Q-learning technique. It was found that the proposed method outperforms the traditional algorithms in terms of FAR, MDR, and AER. For instance, the SARSA-based algorithm achieves a 3–5% lower FAR as compared to the traditional algorithm after 200 rounds of experiments. Furthermore, the algorithm detection accuracy of SARSA is approximately 2–3% higher than the traditional Q-learning-based technique. Consequently, the AER based on the SARSA approach is lower than that based on the Q-learning technique. This is due to the conservative nature of SARSA that finds a suitable threshold via a continuous trial-and-error process. Therefore, the SARSA algorithm has a lower chance of misinterpreting the training signal from legal or illegal users that results in a lower rejection.

We conducted a series of simulation experiments that included receiver utility, receiver's test threshold, analysis of MDR, analysis of FAR, and analysis of AER. Compared with traditional reinforcement learning, we can draw the following conclusions. The suggested scheme can avoid impersonation attacks in a dynamic environment. The experimental results show that our proposed scheme has a higher detection accuracy and lower average error rate, effectively preventing the impersonation attack compared to the traditional reinforcement learning approach.

6. Conclusions

To study the impersonation problem between edge nodes and mobile users in edge computing environments, an IAD method is proposed in this paper with physical layer security technology and a reinforcement learning algorithm. By establishing the IAM in an edge environment, a detection method based on the SARSA algorithm is designed under the IAM, detecting impersonation attacks in a dynamic environment. The experimental results show that the impersonation detection method based on the SARSA algorithm is slightly higher than Q-learning in MDR, but FAR and AER are lower. In this way, communication security is better protected between edge nodes and mobile users in an edge computing environment with higher accuracy.

Author Contributions: Formal analysis, X.Y.; investigation, X.Y. and K.Y.; writing—original draft preparation, X.Y. and K.Y.; writing—review and editing, X.Y., M.U.R. and S.U. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China Postdoctoral Science Foundation (Grant No. 2021M690305). This work was also partially supported by the Singapore Ministry of Education (MoE) AcRF Tier 1 research grants under grant numbers A-0008299-00-00 and A-0008552-01-00.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sulaiman, M.; Halim, Z.; Lebbah, M.; Waqas, M.; Tu, S. An evolutionary computing-based efficient hybrid task scheduling approach for heterogeneous computing environment. *J. Grid Comput.* **2021**, *19*, 1–31. [CrossRef]
2. Abbas, N.; Zhang, Y.; Taherkordi, A.; Skeie, T. Mobile edge computing: A survey. *IEEE Internet Things J.* **2017**, *5*, 450–465. [CrossRef]
3. Lai, X.; Fan, L.; Lei, X.; Deng, Y.; Karagiannidis, G.K.; Nallanathan, A. Secure mobile edge computing networks in the presence of multiple eavesdroppers. *IEEE Trans. Commun.* **2021**, *70*, 500–513. [CrossRef]
4. Ren, J.; Zhang, D.; He, S.; Zhang, Y.; Li, T. A survey on end-edge-cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–36. [CrossRef]
5. Sarkar, S.; Misra, S. Theoretical modelling of fog computing: A green computing paradigm to support iot applications. *IET Netw.* **2016**, *5*, 23–29. [CrossRef]
6. Siriwardhana, Y.; Porambage, P.; Liyanage, M.; Ylianttila, M. A survey on mobile augmented reality with 5g mobile edge computing: Architectures, applications, and technical aspects. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 1160–1192. [CrossRef]
7. Fernando, N.; Loke, S.W.; Rahayu, W. Mobile cloud computing: A survey. *Future Gener. Comput. Syst.* **2013**, *29*, 84–106. [CrossRef]
8. Dang, T.D.; Hoang, D. A data protection model for fog computing. In Proceedings of the 2017 Second International Conference on Fog and Mobile Edge Computing (FMEC), Valencia, Spain, 8–11 May 2017; pp. 32–38.
9. Ullah, A.; Azeem, M.; Ashraf, H.; Jhanjhi, N.; Nkenyereye, L.; Humayun, M. Secure critical data reclamation scheme for isolated clusters in iot-enabled wsn. *IEEE Internet Things J.* **2021**, *9*, 2669–2677. [CrossRef]
10. Alrawais, A.; Alhothaily, A.; Hu, C.; Xing, X.; Cheng, X. An attribute-based encryption scheme to secure fog communications. *IEEE Access* **2017**, *5*, 9131–9138. [CrossRef]
11. Paharia, B.; Bhushan, K. Fog computing as a defensive approach against distributed denial of service (ddos): A proposed architecture. In Proceedings of the 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 10–12 July 2018; pp. 1–7.
12. De Donno, M.; Felipe, J.M.D.; Dragoni, N. Antibiotic 2.0: A fog-based anti-malware for internet of things. In Proceedings of the 2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), Stockholm, Sweden, 17–19 June 2019; pp. 11–20.
13. Tu, S.; Waqas, M.; Rehman, S.U.; Mir, T.; Abbas, G.; Abbas, Z.H.; Halim, Z.; Ahmad, I. Reinforcement Learning Assisted Impersonation Attack Detection in Device-to-Device Communications. *IEEE Trans. Veh. Technol.* **2021**, *70*, 1474–1479. [CrossRef]
14. Xie, N.; Li, Z.; Tan, H. A survey of physical-layer authentication in wireless communications. *IEEE Commun. Surv. Tutor.* **2020**, *23*, 282–310. [CrossRef]
15. Paul, L.Y.; Baras, J.S.; Sadler, B.M. Physical-layer authentication. *IEEE Trans. Inf. Forensics Secur.* **2008**, *3*, 38–51.
16. Wang, X.; Liu, F.J.; Fan, D.; Tang, H.; Mason, P.C. Continuous physical layer authentication using a novel adaptive ofdm system. In Proceedings of the 2011 IEEE International Conference on Communications (ICC), Kyoto, Japan, 5–9 June 2011; pp. 1–5.
17. Chatterjee, B.; Das, D.; Maity, S.; Sen, S. Rf-puf: Enhancing iot security through authentication of wireless nodes using in-situ machine learning. *IEEE Internet Things J.* **2018**, *6*, 388–398. [CrossRef]
18. Mathur, S.; Reznik, A.; Ye, C.; Mukherjee, R.; Rahman, A.; Shah, Y.; Trappe, W.; Mandayam, N. Exploiting the physical layer for enhanced security [security and privacy in emerging wireless networks]. *IEEE Wirel. Commun.* **2010**, *17*, 63–70. [CrossRef]
19. Hou, W.; Wang, X.; Chouinard, J.-Y.; Refaey, A. Physical layer authentication for mobile systems with time-varying carrier frequency offsets. *IEEE Trans. Commun.* **2014**, *62*, 1658–1667. [CrossRef]
20. Gopinath, S.; Guillaume, R.; Duplys, P.; Czylwik, A. Reciprocity enhancement and decorrelation schemes for PHY-based key generation. In Proceedings of the 2014 IEEE Globecom Workshops (GC Wkshps), Austin, TX, USA, 8–12 December 2014; p. 13671372.
21. Waqas, M.; Ahmed, M.; Li, Y.; Jin, D.; Chen, S. Social-aware secret key generation for secure device-todevice communication via trusted and non-trusted relays. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 3918–3930. [CrossRef]

22. Zeng, K. Physical layer key generation in wireless networks: Challenges and opportunities. *IEEE Commun. Mag.* **2015**, *53*, 33–39. [CrossRef]
23. Aldaghri, N.; Mahdavifar, H. Physical layer secret key generation in static environments. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 2692–2705. [CrossRef]
24. Eberz, S.; Strohmeier, M.; Wilhelm, M.; Martinovic, I. A practical man-in-the-middle attack on signalbased key generation protocols. In *European Symposium on Research in Computer Security*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 235–252.
25. Chen, D.; Qin, Z.; Mao, X.; Yang, P.; Qin, Z.; Wang, R. Smokegrenade: An efficient key generation protocol with artificial interference. *IEEE Trans. Inf. Forensics Secur.* **2013**, *8*, 1731–1745. [CrossRef]
26. Sudarsono, A.; Yuliana, M.; Kristalina, P. A reciprocity approach for shared secret key generation extracted from received signal strength in the wireless networks. In Proceedings of the 2018 International Electronics Symposium on Engineering Technology and Applications (IES-ETA), Bali, Indonesia, 29–30 October 2018; pp. 170–175.
27. Ma, Z.; Teschendorff, A.E.; Leijon, A.; Qiao, Y.; Zhang, H.; Guo, J. Variational bayesian matrix factorization for bounded support data. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *37*, 876–889. [CrossRef] [PubMed]