Physical Layer Authentication in the Internet of Vehicles based on Signal Propagation Attribute Prediction

Mubarak Umar[1,2,3,*], Jiandong Wang[1,2], Lei Liu[4], Zewei Guo[1,2], and Shuguang Wang[5]
[1]School of Computer Science and Technology, Xidian University, Xi'an 710071, China
[2]Xidian University Qingdao Institute of Computing Technology, Qingdao, China
[3]Department of Information Technology, Bayero University, Kano 700241, Nigeria
[4]Software College, Shandong University, China
[5]Shandong Institute of Standardization, No.146-6, Lishan Road, Jinan City, China

Physical layer authentication (PLA) has emerged as a promising alternative to complex cryptographic-based authentication schemes, especially for the Internet of Vehicles (IoV) scenarios with resource-limited onboard units (OBUs). However, the existing PLA schemes securing the IoV against GPS location spoofing/falsification attacks consider only insider attackers. Moreover, they cannot be used by mobile vehicles to validate GPS locations. To address these issues, this paper proposes a PLA scheme based on the Gaussian process (GP) path loss prediction, where channel state information (CSI) is used to track the variation of the channel characteristics and predict the next legitimate path loss (PL) of the signal from a transmitter for authentication. The key ideas in the proposed scheme are to first establish a mapping between the historical CSI attributes and PL features of the transmitter's signal and use this mapping to predict the next PL, which is then used to cross-verify the transmitter's reported location information. Extensive simulation experiments are conducted using generated radio channel characteristics from the quasideterministic radio channel generator (QuaDRiGa) to demonstrate the effectiveness of the proposed approach. The results of the experiments show that our system efficiently addressed the limitations of the existing works and improves the authentication performance in IoV environments.

Index Terms—**Internet of vehicles (IoV), Gaussian process (GP), machine learning (ML), physical layer authentication (PLA), path loss (PL).**

I. Introduction

The developments in wireless communication technologies witnessed in the past few decades have enabled the emergence of the Internet of Vehicles (IoV) in which vehicles exchange safety messages over open wireless channels [1], [2]. As an essential component of the future intelligent transportation system (ITS), IoV is a two-sided coin: while its deployment has the potential to improve road safety and traffic management [3], thereby saving lives, the broadcast nature of its communication through public wireless channels exposes the exchanged safety messages to unauthorized access and tampering, which could have life-threatening consequences [3]. Thus, it is vital to provide authentication solutions for secure IoV communications [4].

Traditionally, key-based cryptographic authentication approaches have been utilized to secure IoV against security attacks through encryption/decryption algorithms. For example, a three-level security framework called NOTSA is designed to secure onboard units (OBUs) communication [5]. Wazid *et al.* [6] designed an elliptic curve cryptography (ECC) based authentication scheme, where vehicles are directly verified by servers. A zero-knowledge proof (ZKP) technique is utilized in [7] to secure OBU communication and ensure secure toll payment processing. In [8], Ma *et al.* put forward a multicast service authentication scheme for 5G-based vehicle-to-everything (V2X). In [9], the authors proposed a certificateless authentication approach, which utilized the Chinese remainder theorem to effectively distribute keys and conduct dynamic wiretapping for anomaly detection. In [10], Yang *et al.* developed a cooperative-based authentication, where vehicles are verified by a small number of already authenticated vehicles. In [11], authentication keys are pre-distributed using the future location of vehicles predicted through the recurrent neural network (RNN). In [12], the authors presented a handover authentication approach for IoV communications. In [13], Zhang *et al.* put forward SMAKA, an effective authentication framework that secures vehicles to cloud server communications. Despite their proven security strength in that they cannot be broken by attackers without the decryption keys, these above schemes have high computation requirements making their adoption by the IoV applications with resource-constraints OBUs very difficult. Moreover, the complex key management procedures in these schemes incur unacceptable latency, especially for the delay-sensitive IoV applications. Thus, it is vital to provide lightweight authentication approaches for the IoV network.

A. Developments and Limitations of the Existing PLA

To address the limitations of the above cryptographic-based authentication approaches in the IoV, physical layer authentication (PLA) technology has emerged in recent years [14]. It is a type of authentication characterized by low computational complexity and low latency, making it suitable for the resource-limited OBUs in the IoV network. Specifically, PLA

Manuscript received December 14, 2022; revised February 7, 2023.
Corresponding author: Mubarak Umar (email: mubarakumar53@gmail.com) Email Addresses: Mubarak Umar (mubarakumar53@gmail.com), Jiandong Wang (wjdmail@126.com), Lei Liu (l.liu@sdu.edu.cn), Zewei Guo (guozewei1018@163.com), Shuguang Wang (2869377@qq.com)

utilizes machine learning (ML) [14] or hypothesis test [15] to facilitate device verification through physical layer channel features extracted from radio signals. The key idea in PLA is that the previously known signals declared by a legitimate vehicle should be close to the current signal received from that vehicle. Thus, authentication can be achieved by comparing the current signal attributes with the ones previously extracted from the legitimate vehicle. Commonly used channel attributes as the unique physical layer signatures for authentication include channel state information (CSI) [14], [16], received signal strength (RSS) [15], and angle of arrival (AoA) [17], which are both vehicle location and speed dependent. Based on the well-known principle that two legitimate devices measure similar channel characteristics, which are different from the channel characteristics of an attacker that is located half a wavelength [14], it is difficult for the channel between the legitimate transceivers to be forged by an attacker. As a result, CSI, as a fine-grained location/speed-specific channel attribute has been utilized as a unique fingerprint to verify legitimate vehicles.

Along this line, several works utilized CSI to design authentication approaches in IoV scenarios. In [18], Wang *et al.* set up a hypothesis test exploiting extended and unscented Kalman filters on RSS, speed, and distance between vehicles to achieve authentication. In Wang *et al.* [15], the statistical properties of noise in RSS and CSI are adjusted using the Sage-Husa adaptive Kalman filter for authentication in the IoV network. However, the thresholds for the hypothesis test in [15] and [18] may not be always available in dynamic IoV environment. Consequently, several works exploiting the internal features of the CSI through ML approaches have been proposed [19], [20], [21], [22]. Both Yin *et al.* [19] and Fang *et al.* [20] utilized multiple physical layers attributes jointly to provide robust ML-based authentication in the IoV environment. In [19], the attributes are chosen according to their past authentication performance while in [20], a kernel fusion machine is designed to deal with the multiple attributes without requiring the previous knowledge of their statistical properties. Moreover, Chen *et al.* [21] proposed a convolutional neural network-based authentication and data augmentation approach named triple pool network (TP-Net) for an edge computing system. In [22], a threshold-free authentication scheme is designed using a support vector machine (SVM), a decision tree, and ensemble learning. In [14], Wang and Fu put forward a Gaussian process (GP) based authentication scheme, where the channel of a legitimate vehicle is tracked and predicted using the previous relationship between the CSI and the geographical location of the vehicle. A CSI-based deep learning authentication scheme is designed in [16] to verify devices according to the CSI signature of their fixed locations.

Thanks to the integration of the ML techniques, the statistical distribution of CSI in the above ML-based schemes is not required and the thresholds are determined through training. However, most of the ML-based schemes are not secure against GPS location spoofing attack. Thus, authentication approaches have been designed recently to secure the IoV network from location spoofing/falsification attacks, which could have life-threatening consequences. In [23], an unscented Kalman filter is used on RSS data to design a misbehavior detection approach and protect the IoV against GPS location spoofing/falsification attacks. Three plausibility checks techniques are set up using RSS data to detect insider attackers in the IoV network [24]. In [17], the AoA of received signals is used to cross-check location information reported by vehicles and protect the IoV against location spoofing/falsifying attacks.

However, the above PLA approaches do not consider *outside attackers for GPS location spoofing/falsification attacks* (i.e., they only consider insider attackers). In a normal IoV setting, vehicles are most of the time moving and rely heavily on GPS technology to achieve fine-grained positioning. The commercial GPS receivers in most of these vehicles have been proven to have insecure designs exposing them to GSP signal spoofing attacks, where GPS receivers are tricked into accepting signals from attackers as though they are from legitimate GPS satellites. Although the protection against location spoofing/falsification attacks from insider attackers in the existing PLA schemes [23], [24] ensures that legitimate vehicles are not tricked into disseminating false locations that could lead to accidents, GPS location spoofing/falsification attacks from outside attackers could lead to the impersonation of legitimate vehicles by attackers or rejection of legitimate vehicles due to their spoofed locations. Thus, it is essential as well to prevent the IoV network against such attacks from outside attackers.

Another critical limitation of the existing PLA schemes is that *they cannot be used by mobile vehicles to validate reported location information* (i.e., they can only be used by stationary road-side units (RSUs) to validate reported locations of vehicles, not the other way around). In [17], for instance, only RSU can validate reported locations by vehicles during authentication. However, most of the time, the vehicles in the IoV network are mobile, making it difficult to adopt the existing schemes.

Overall, the performance of the existing PLA schemes could be severely affected due to the potential GSP location spoofing/falsification by outside attackers and the mobile nature of the vehicles in the IoV network. It is impractical to limit the prevention of these attacks from just insider attackers and by only static RSUs.

B. Our Solutions

To address the above limitations, we propose tracking and predicting the channel of legitimate vehicles based on the physical environment. In IoV applications, it is not out of place to assume that the trajectories of legitimate transmitters could be *priori* known by legitimate receivers. Our solution consists of two steps: First, we establish a mapping between historical CSI attributes and the corresponding path loss (PL) features and use this mapping to predict the next legitimate PL feature of a transmitter's signal. Second, the predicted PL feature is then used to cross-verify the transmitter's reported location information. The main innovation of our work is the verification of claimed locations of transmitters through PL measurements predicted using CSI attributes, which are significantly dependent on the transmitter's location and speed. This verification ensures that legitimate transmitters under location spoofing attack are not mistakenly rejected while impersonator vehicles are correctly rejected. The path loss of a signal is defined as the reduction of the signal's power as it travels through a medium. We propose to utilize a non-linear kernel function-based Gaussian process (GP) [25] to establish the relationship between the historical CSIs and the corresponding PL features of the legitimate transmitter's signals. Based on this established mapping, we predict the next legitimate PL feature of the transmitter's signal for authentication. The GP has in the last few decades received increasing attention especially for the prediction of vehicle trajectory thanks to its excellent performance even without a large amount of training data [25]. Furthermore, due to the absence of the attacker's PL estimates or its channel characteristics distribution, our scheme is a variant of one-class authentication which focuses on only the PL of legitimate vehicles for authentication decisions. After the next legitimate PL is predicted, we derive an acceptance region around the predicted PL to determine whether the received signal comes from a legitimate vehicle or an impersonator.

C. Contributions

Aiming to address the limitations of the existing PLA schemes, we propose a PLA approach based on signal propagation attribute prediction to secure IoV communications against GPS location spoofing and falsification attacks from outside attackers. Specifically, the main contributions of our work are summarized as follows.

1) First, we introduce PL information as a security parameter that is utilized as a solution to the problem of legitimate transmitters under location spoofing attack being rejected and that of outside attackers with stolen identities attempting to impersonate legitimate vehicles.
2) Second, we exploit the relationship between previously obtained CSI estimates and their corresponding PL measurements of the signals from legitimate transmitters and use the GP to predict the next PL of the transmitter's signal. The predicted PL is then used to cross-verify the claimed location information of the transmitters.
3) Finally, to demonstrate the effectiveness of our approach, we utilized a quasideterministic radio channel generator (QuaDRiGa) [26] simulation platform to generate realistic CSI measurements and conduct system-level simulation experiments. The results of the simulations show that the introduction of PL measurements has significantly addressed the challenges of the existing approaches.

Following the introduction in Section I, the rest of this article is structured as follows. Section II describes the system model in detail and introduces the GP regression. Section III describes the proposed signal propagation attribute prediction based authentication scheme. Section IV presents the results of the conducted simulation experiments and their analysis. Lastly, Section V gives the conclusions of our work.

Notations: Matrices and vectors are represented using bold uppercase and lowercase letters, respectively. Transpose, inverse, and determinant of a matrix are denoted by $(\cdot)^T$, $(\cdot)^{-1}$, and $det(\cdot)$. $\boldsymbol{I}$ denotes the identity matrix and $E(\cdot)$, $var(\cdot)$, $p(\cdot)$ represents expectation, variance, and probability operators. $g \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ denotes a zero mean Gaussian random vector g with covariance matrix $\mathbf{R}$ and zero mean.

II. System model and preliminaries

In this section, we present the considered system model and discuss some preliminaries.

A. System Model

As depicted in Fig. 1, we consider the well-known Alice-Bob-Eve model in the IoV environment. In this model, Alice is a legitimate vehicle that wishes to continuously exchange information with legitimate RSU referred to as Bob in a secure manner. Eve in our model is an attacker vehicle that continuously attempts to impersonate Alice or Bob. Bob is assumed to be situated in a fixed location, while Alice and Eve are assumed to be either stationary or mobile at any time. Both Alice and Bob are equipped with resource-limited sensors and OBUs, while Eve is assumed to have illegal OBUs and sensors that are used by Eve to launch security attacks. We consider two communication scenarios: vehicle-to-vehicle (V2V) and infrastructure-to-vehicle or vice versa (I2V/V2I). In our model, Bob builds up a table containing the $\hat{P}L_A(t-1)$ (i.e., the last path loss value of the last signal from Alice received by Bob) and the corresponding last location information of Alice. This stored $\hat{P}L_A(t-1)$ is the verified path loss value of Alice's legitimate message estimated during the channel estimation stage between Bob and Alice. The $\hat{P}L_A(t-1)$ value is critical in our scheme because it will be used by Bob to determine whether the next message received by Bob at a time t is from Alice or Eve depending on whether the predicted path loss value of the message $\tilde{P}L(t)$ at a time t is close to the last $\hat{P}L_A(t-1)$ of Alice's message.

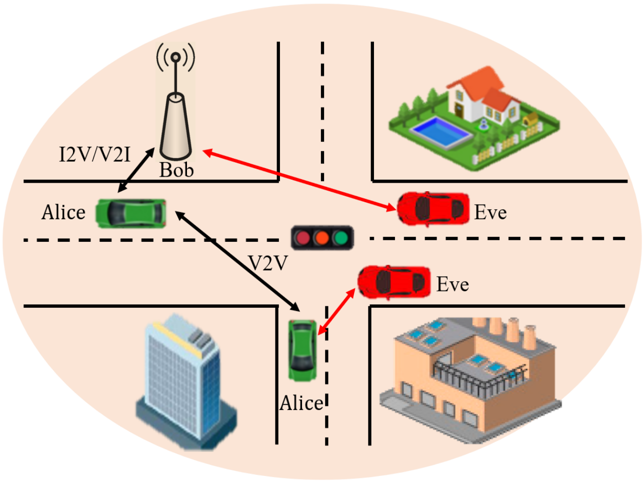


Fig. 1. System model.

In this model, the relationship between the CSI and their corresponding PL measurements in the past $t-1$ time slots is constructed and a GP model is trained offline to recognize this relationship and predict the next legitimate PL at time t for authentication. The CSIs and the PL measurements are estimated by exchanging known pilot sequences between Alice and Bob in the past $t-1$ times. Assuming an orthogonal frequency division multiplexing (OFDM) system, we extract channel frequency response (CFR) matrix from the estimated CSI of M subcarriers in the past $t-1$ times to obtain an M-dimensional CFR matrix which can be denoted as $\left\{\hat{\mathbf{h}}_A(v)\right\}, v = 1, 2, \ldots, t-1$, where $\hat{\mathbf{h}}_A(v) = \left(\hat{h}_A^{(1)}(v), \hat{h}_A^{(2)}(v), \ldots, \hat{h}_A^{(M)}(v)\right)^T$. During the online authentication when Alice sends Bob her claimed location information at time t, the CSI vector $\hat{\mathbf{h}}_A(t)$ of the received Alice's signal is inputted into the trained GP model to predict the $\tilde{PL}(t)$ of the signal from Alice. After that, Bob then obtains the $\hat{PL}_A(t-1)$ of the last Alice's signal from the table it keeps for Alice. Note that throughout this paper, we refer to the last PL of Alice's signal as expected path loss. Next, Bob formulates the problem of authentication as a binary hypothesis test as

$$\begin{aligned} \mathcal{H}_0 &: \tilde{PL}(t) = \hat{PL}_A(t-1) \\ \mathcal{H}_1 &: \tilde{PL}(t) \neq \hat{PL}_A(t-1), \end{aligned} \tag{1}$$

where $\mathcal{H}_0$ indicates that the received message is from Alice, whereas the $\mathcal{H}_1$ means that the received signal is from Eve.

Moreover, we assume a symmetric channel between Alice and Bob, and thus, our scheme can be used by both Alice and Bob to authenticate each other, thereby addressing the limitation of the existing schemes since they cannot be applied by moving vehicles to validate reported location information.

B. Attack Model

Eve in this model is assumed to be an *outsider attacker* capable of launching impersonation and GPS location spoofing attacks. During the impersonation attack, Eve fabricates the identities of Alice or Bob and uses them to impersonate Alice to Bob and vice versa. In the GPS spoofing attack, Eve tricks Alice into sending the wrong location information by interfering with the signals from GPS satellites meant for Alice's GPS receivers, which could lead to Alice's messages being rejected. Moreover, we assume that Eve is aware of the channel characteristics in the environment of Alice and Bob and can estimate the channel between itself and either Alice or Bob. Eve can also eavesdrop on the channels of Alice and Bob and inject false data. We also assume that Eve is located at least half a wavelength away from the locations of Alice and Bob and thus cannot be able to estimate the exact channel between Alice and Bob due to the channel decorrelation nature in time and space. Note that we do not consider jamming or denial-of-service (DoS) attacks in this work.

C. Gaussian Process Regression in Internet-of-Things Security

The Gaussian process (GP) regression is a powerful state-of-the-art tool for prediction and function approximation, which has recently been shown to excel in IoV security, especially for the prediction of vehicle trajectory [14]. As a kernel method of the Bayesian non-parametric system, the GP model approximates a distribution of training data by finding a solution to a series of hyperparameters so that corresponding output under random input can be obtained during prediction. In the following, we give a brief introduction to the GP and refer readers to [25] for a more detailed discussion.

Given a set of training data $\mathbf{H} = (\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_D)^T$ where $\mathbf{h}_n = (h_{n,1}, h_{n,2}, \ldots, h_{n,q})^T$ and the matching training labels $\mathbf{y} = (y_1, y_2, \ldots, y_D)^T$, where D represents the size of the training data and q represents the dimension of the training input. If we imagined the training data as a single point sampled from some multivariate Gaussian distribution, the training data can be partnered with a GP, which is completely determined by a mean $m(\mathbf{h})$ (assumed to be zero everywhere) and a covariance function $k(\mathbf{h}, \mathbf{h}^{'})$. The $k(\mathbf{h}, \mathbf{h}^{'})$ relates one sample in the training data to another and one of the popular choices for such a function is the squared exponential expressed as

$$k(\mathbf{h}, \mathbf{h}^{'}) \triangleq \sigma_f^2 \text{exp}\left(-\frac{1}{2}\frac{\left(\mathbf{h}-\mathbf{h}^{'}\right)^T\left(\mathbf{h}-\mathbf{h}^{'}\right)}{\sigma_l^2}\right) + \sigma_n^2 \delta_{h,h'}, \tag{2}$$

where σ_f is the signal variance, σ_l denotes the characteristic length scale, and σ_n is the variance of the input noise to signify the randomness attributes of wireless communication.

With the training data and training labels $\mathbf{H}$ and $\mathbf{y}$, respectively, the logarithmic marginal likelihood function can be derived as

$$\log p(\mathbf{y}|\mathbf{H}) = -\frac{1}{2}\mathbf{y}^T\mathbf{K}^{-1}\mathbf{y} - \frac{1}{2}\log|\mathbf{K}| - \frac{D}{2}\log 2\pi, \tag{3}$$

where $\mathbf{K}$ is the covariance matrix for the noisy target outputs $\mathbf{y}$, which consists of the covariance function among all the pairs of the training samples and is expressed as

$$K(\mathbf{H}, \mathbf{H}) \triangleq \begin{bmatrix} k(\mathbf{h}_1, \mathbf{h}_1) & k(\mathbf{h}_1, \mathbf{h}_2) & \cdots & k(\mathbf{h}_1, \mathbf{h}_N) \\ k(\mathbf{h}_2, \mathbf{h}_1) & k(\mathbf{h}_2, \mathbf{h}_2) & \cdots & k(\mathbf{h}_2, \mathbf{h}_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{h}_N, \mathbf{h}_1) & k(\mathbf{h}_N, \mathbf{h}_2) & \cdots & k(\mathbf{h}_N, \mathbf{h}_N) \end{bmatrix}. \tag{4}$$

The hyperparameters of the GP model are determined by maximizing the marginal log-likelihood function in (3). Given a single input $\mathbf{h}_*$ during prediction, its output y_* proven to also obey a Gaussian distribution [25] has a mean and variance as

$$\bar{y}_* = \mathbf{k}_*^T\mathbf{K}^{-1}\mathbf{y} \tag{5}$$

and

$$var(y_*) = \mathbf{K}_{**} - \mathbf{k}_*^T\mathbf{K}^{-1}\mathbf{k}_*, \tag{6}$$

respectively, where $\mathbf{k}_* = K(\mathbf{h}_*, \mathbf{H})$ and $\mathbf{K}_{**} = k(\mathbf{h}_*, \mathbf{h}_*)$. By adopting the GP model in our work for path loss prediction where the output is the PL at the next time slot, the training data $\mathbf{H}$ and labels $\mathbf{y}$ represent the previous CSIs and the corresponding PLs, respectively. After the GP model is trained, by putting a new CSI vector $\mathbf{h}_*$, the model predicts the mean and variance of y_*, i.e., the path loss.

III. The Proposed Signal Propagation Attribute Prediction Based Physical Layer Authentication Scheme

As depicted in Fig. 2, the proposed PLA approach is made up of two functional components: 1) offline training; 2) online authentication. Before the offline training, we collected raw CSI, 2-D location, PL data, and the identity of Alice. First, we discuss the offline training methodology and then describe the authentication procedure of the proposed PLA scheme.

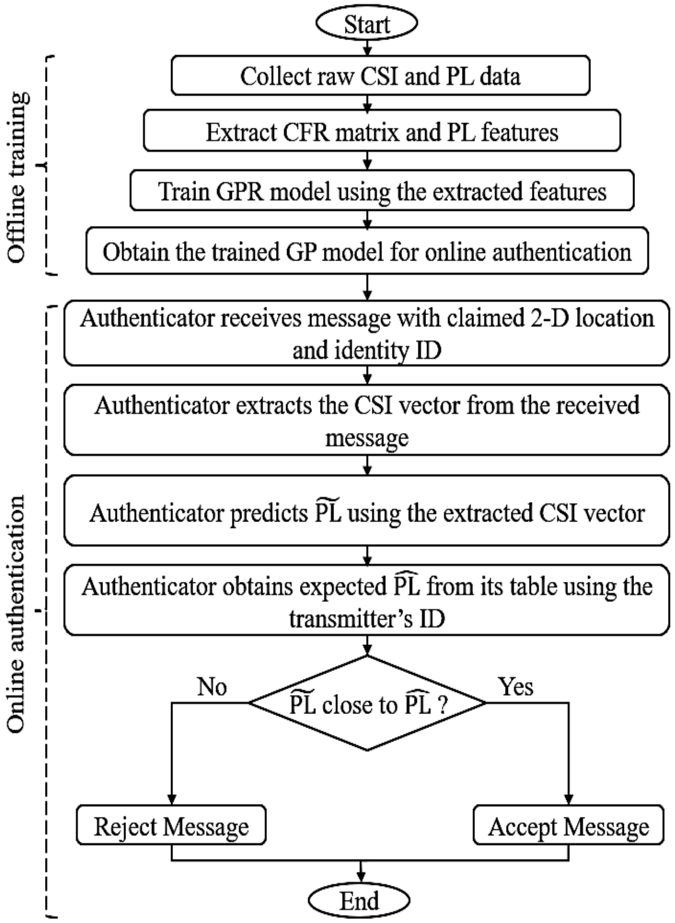

Start
Collect raw CSI and PL data
Extract CFR matrix and PL features
Train GPR model using the extracted features
Obtain the trained GP model for online authentication
Offline training
Authenticator receives message with claimed 2-D location and identity ID
Authenticator extracts the CSI vector from the received message
Authenticator predicts $\widetilde{\mathrm{PL}}$ using the extracted CSI vector
Authenticator obtains expected $\widehat{\mathrm{PL}}$ from its table using the transmitter's ID
$\widetilde{\mathrm{PL}}$ close to $\widehat{\mathrm{PL}}$?
No
Yes
Reject Message
Accept Message
End
Online authentication

Fig. 2. Flowchart of the proposed scheme.

A. *Offline Training*

The main elements of our proposed PLA scheme are the CSI and the PL acquired through the channel estimation procedure [20], [21]. In this section, we first describe the procedure for the CSI and PL data estimations and then discuss the training of the GP model.

1) *CSI Data Estimation*

The signal from Alice received by Bob during the CSI estimation is given as

$$y_\rho(v) = h_\rho \times x_\rho(v) + \omega(v), \tag{7}$$

where y_ρ is the received signal at time $v = 1, 2, \ldots, t-1$, where v denotes the time interval between every message, h_ρ denotes the channel matrix in the time domain containing the channel coefficients, x_ρ is the pilot signal known to both Alice and Bob used for the estimation of the CSI and the PL, and $\omega(v)$ represents the white Gaussian noise with variance σ^2. The time domain channel estimated by Bob though the channel estimation is given by

$$\begin{aligned} h(v) &= y_\rho(v) \times \frac{1}{x_\rho(v)} \\ &= (h_\rho \times x_\rho(v) + \omega(v)) \times \frac{1}{x_\rho(v)} \\ &= h_\rho + \omega(v) \times x_\rho^{-1}(v), \end{aligned} \tag{8}$$

where $x_\rho^{-1}(v)$ is the inversion of $x_\rho(v)$. Finally, Bob obtains the channel frequency response (CFR) matrix through the discrete Fourier transformation (DFT) of the above time-domain channel. The CFR matrix of Alice's signal is expressed as $\mathbf{H}_A = \left(\hat{\mathbf{h}}_A^{(1)}, \hat{\mathbf{h}}_A^{(2)}, \ldots, \hat{\mathbf{h}}_A^{(M)}\right)$, where $\hat{\mathbf{h}}_A(v) = \left(\hat{h}_A^{(1)}(v), \hat{h}_A^{(2)}(v), \ldots, \hat{h}_A^{(M)}(v)\right)^T$. For simplicity, we use single receiving/transmitting antennas, 64 subcarriers, and thus, for the entire duration of the CSI data collection, Bob collects $64 \times v$ raw CSI estimates, where $v = 1, 2, \ldots, t-1$ stands for the number of previous times Alice sends the pilot signal to the Bob.

2) *Path Loss Data Estimation*

Bob obtains the observed PL of the signals from Alice in decibel (dB) during the channel estimation procedure, which is expressed as

$$PL_A(v) = E \times \log_{10}(\mathscr{d}(v)) + F, \tag{9}$$

where $E = 28.5$ dB/log10(m) and $F = 38$ dB are path loss dependent distance between Alice and Bob and reference PL at 1 GHz/1 m distance between Alice and Bob, respectively. $\mathscr{d}$ is the current distance between Alice and Bob in meters.

It should be noted that the values of E and F are estimated from the Berlin survey, Germany (Berlin Uma), which are used in QuaDRiGa [26]. The obtained PL vector from Alice during the channel estimation is expressed as $\hat{\mathbf{y}}_A = \left(\hat{PL}_A(1), \hat{PL}_A(2), \ldots, \hat{PL}_A(v)\right)^T$. We assume that Alice and Bob are associated with anonymous identities ID, and they have GPS location information of themselves and each other through exchanging of known pilot signals. Each of the pilot signals exchanged between Alice and Bob during the channel estimation stage contains the 2-D location of Alice $(\hat{\ell}_{A_{v,1}}, \hat{\ell}_{A_{v,2}})$ and her identity ID_A. Using the expected PL, the IDs, and the 2-D location information collected through the channel estimation procedure, Bob builds up a table as shown in Table I containing the expected PL value (i.e., the path loss value of the last signal from Alice, which in this case is $\hat{PL}_A(t-1)$) and the corresponding 2-D location and ID of Alice contained in the last Alice's signal received by Bob. The main objective of Bob as an authenticator is to check the consistency between the expected PL of Alice's message and the predicted one by the GP model.

TABLE I. List of 2-D Location and Expected Path Loss Associated with Alice.

Alice's ID	2-D Location	Expected PL
ID_A	$(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$	$\hat{PL}_A(t-1)$

3) GPR Model Training

After the creation of the table containing the ID, 2-D location, and expected PL of Alice at $t-1$ time by Bob, the next stage is the training of the GP model using the estimated CFR matrix $\mathbf{H}_A$ as the training data set and the PL vector $\hat{\mathbf{y}}_A$ as the corresponding training label set. The generated training input and target sets are expressed as

$$\mathbf{D}_{train} = \left(\mathbf{H}_{train}, \mathbf{y}_{train}\right), \tag{10}$$

where

$$\mathbf{H}_{train} = \begin{bmatrix} \mathbf{H}_{A1} \\ \mathbf{H}_{A2} \\ \vdots \\ \mathbf{H}_{An} \end{bmatrix}, \tag{11}$$

$$\mathbf{y}_{train} = \begin{bmatrix} \hat{\mathbf{y}}_{A1} \\ \hat{\mathbf{y}}_{A2} \\ \vdots \\ \hat{\mathbf{y}}_{An} \end{bmatrix}. \tag{12}$$

The covariance matrix in (4) can be represented as $K(\mathbf{H}, \mathbf{H})$ consisting of covariance functions $k(\mathbf{h}_i, \mathbf{h}_j)$ over all pairs of the training sets and can be expressed as

$$\begin{aligned} k(\mathbf{h}_i, \mathbf{h}_j) &\triangleq (\sigma_f)^2 \text{exp}\left(-\frac{1}{2}\frac{\left(\mathbf{h}_i - \mathbf{h}_j\right)^T\left(\mathbf{h}_i - \mathbf{h}_j\right)}{(\sigma_l)^2}\right) \\ &+ (\sigma_n)^2 \delta_{h_i, h_j}, i, j = 1, 2, \dots, v, \end{aligned} \tag{13}$$

where σ_l^2 is used to control the variations scale of the CSI vectors of all the subcarriers. The hyperparameter set of the GP model is defined as $\boldsymbol{\theta} \triangleq \{\sigma_f, \sigma_l, \sigma_n\}$. Following the GPR process described in Section II, we derive the optimal solution of the hyperparameters through maximizing the marginal log-likelihood function in (3) by seeking the partial derivates of the marginal log-likelihood concerning the hyperparameters as follows:

$$\begin{aligned} \frac{\partial}{\partial \theta_j} \log p\left(\hat{\mathbf{y}}_A | \mathbf{H}_A\right) &= \frac{1}{2}\left(\hat{\mathbf{y}}_A\right)^T \mathbf{K}^{-1} \frac{\partial \mathbf{K}}{\partial \theta_j} \mathbf{K}^{-1}\left(\hat{\mathbf{y}}_A\right) \\ &- \frac{1}{2}\text{tr}\left(\mathbf{K}^{-1}\frac{\partial \mathbf{K}}{\partial \theta_j}\right) \\ &= \frac{1}{2}\text{tr}\left(\left(\boldsymbol{\alpha}\boldsymbol{\alpha}^T - \mathbf{K}^{-1}\right)\frac{\partial \mathbf{K}}{\partial \theta_j}\right) \end{aligned} \tag{14}$$

where $\boldsymbol{\alpha} \triangleq \mathbf{K}^{-1}\hat{\mathbf{y}}_A, \theta_j \in \boldsymbol{\theta}, j = 1, 2, 3$. We have

$$\begin{aligned} \frac{\partial \mathbf{K}}{\partial \theta_1} &= 2\sigma_f \mathbf{R} \\ \frac{\partial \mathbf{K}}{\partial \theta_2} &= \left(\sigma_f\right)^2 \mathbf{R} \odot \mathbf{P} \\ \frac{\partial \mathbf{K}}{\partial \theta_3} &= 2\sigma_n \mathbf{I} \end{aligned}$$

where $\mathbf{R}(i,j) \triangleq \exp\left(-(1/2)(\mathbf{h}_i - \mathbf{h}_j)^T(\mathbf{h}_i - \mathbf{h}_j)/(\sigma_l)^2\right)$, $\mathbf{P}(i,j) \triangleq \left[(h_{i,1} - h_{j,1})^2 + (h_{i,2} - h_{j,2})^2\right]/(\left[(\sigma_l)^3\right])$. After the prediction model is obtained through the optimization of the above hyperparameters, a CSI vector $\hat{\mathbf{h}}_A(t) = \left(\hat{h}_A^{(1)}(t), \hat{h}_A^{(2)}(t), \dots, \hat{h}_A^{(M)}(t)\right)^T$ of a signal from Alice at a time t is inputted into the trained model to predict the path loss $\tilde{P}L_A(t)$ referring to (5), where

$$\tilde{P}L_A(t) = \mathbf{k}^T\mathbf{K}^{-1}\hat{\mathbf{y}}_A, \tag{15}$$

and $\mathbf{k} \triangleq (k(\mathbf{h}_t, \mathbf{h}_1), k(\mathbf{h}_t, \mathbf{h}_2), \dots, k(\mathbf{h}_t, \mathbf{h}_N)))^T$, where $k(\mathbf{h}_t, \mathbf{h}_n)$ is as defined in (13). Due to the presence of noise in our communication model, the predicted path loss is not perfect and there will always be a prediction error. Thus, the predicted output can be expressed as

$$\tilde{P}L_A(t) = \hat{P}L_A(t-1) + \varepsilon(t), \tag{16}$$

where $\hat{P}L_A(t-1)$ denotes the expected PL from Alice in the table maintained by Bob and $\varepsilon(t) \sim \mathcal{N}(\mathbf{0}, \sigma_{pre}^2(t)\mathbf{I})$ represents the prediction error.

B. Online Authentication

In this section, we discuss the authentication procedure between Alice and Bob executed each time they want to communicate as illustrated in Fig. 2 and summarized in Algorithm 1.

First, in Algorithm 1, Alice begins by sending a message to Bob containing its claimed 2-D location $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ and identity ID_A at a time t (line 1). Upon receiving the message from Alice, Bob first obtains the observed CSI vector $\hat{\mathbf{h}}_A(t)$ of the signal and then inputs it into the trained GP model to predict the path loss $\tilde{P}L_A(t)$ of the signal from Alice (lines 2-3). Next, Bob checks his Table I, which contains the previously obtained path loss value of the last signal from Alice before time t, and uses Alice's identity ID_A to get the recently stored expected path loss value $\hat{P}L_A(t-1)$. Our reason for using the $\hat{P}L_A(t-1)$ at a time $t-1$ is that the current path loss value $\tilde{P}L_A(t)$ at t is closer to the $\hat{P}L_A(t-1)$ than it is to say $\hat{P}L_A(t-2)$ or $\hat{P}L_A(t-3)$. This is because, beyond a certain range of both geographical locations and time intervals, the PL values can be considered uncorrelated. Bob then uses the predicted $\tilde{P}L_A(t)$ and calculates a decision region to determine whether to accept or reject the message from Alice as follows (line 4).

$$\mathcal{D}_A^1(t) = \left[\tilde{P}L_A(t) - z\bar{\varepsilon}(t), \tilde{P}L_A(t) + z\bar{\varepsilon}(t)\right], \tag{17}$$

where $\bar{\varepsilon}(t)$ stands for the cumulative prediction error of the previous path loss values and z is a parameter controlling the performance of the false alarm and miss detection. The previous cumulative prediction error $\bar{\varepsilon}(t)$ is define as

$$\bar{\varepsilon}(t) = \frac{1}{t-1}\sum_{v=1}^{t-1} \varepsilon(v), \tag{18}$$

where $\varepsilon(t) = \tilde{P}L_A(t) - \hat{P}L_A(t-1)$ denotes the path loss prediction error at a time t.

Next, Bob checks if the predicted $\tilde{P}L_A(t)$ is close to the $\hat{P}L_A(t-1)$ (line 5). If the two path loss values are close, then Bob checks if the current claimed location $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ by Alice is also close to the previously stored location $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ (line 6). If they are close, Bob is sure that Alice is legitimate. Bob then accepts the message from Alice, discards the immediate past location $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$

and path loss $\hat{PL}_A(t-1)$ entries in Table I, and replaces them with the current $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ and $\tilde{PL}_A(t)$ (line 7). This step reduces storage cost on Bob by simply removing the entries that have less or no correlation with the newly obtained entries. If the two locations aren't close but the two path loss values are close, Bob still accepts the message from Alice and concludes that Alice is a legitimate vehicle under a location spoofing attack. Bob then replaces the previous path loss value $\hat{PL}_A(t-1)$ with $\tilde{PL}_A(t)$ in Table I and maintains the previous location of Alice $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ as the current one instead of the spoofed one $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ (lines 8-10). Otherwise, if both the two path loss values $\hat{PL}_A(t-1)$ and $\tilde{PL}_A(t)$ as well as the two locations $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ and $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ are not close, Bob concludes that Alice is an impersonator vehicle, rejects the message from Alice, and keeps $\hat{PL}_A(t-1)$ and $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ to be used for the next time slot (lines 11-13). Finally, Bob computes the new prediction error $\varepsilon(t)$ using $\tilde{PL}_A(t)$ and $\hat{PL}_A(t-1)$. Bob then updates the time to $t = t+1$ and computes the cumulative prediction error $\bar{\varepsilon}(t)$ for the path loss to be used by Bob in the next $t+1$ time slot (lines 14-15).

Algorithm 1: Summary of the Authentication at a time t in the Proposed Scheme

```
Input: $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$, $ID_A$;
Output: Accept or reject the received message;
Initialization: $z$, $\bar{\varepsilon}(t)$;
1  Alice sends its claimed $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ and $ID_A$ at a
     time $t$ to Bob;
2  Bob obtains the observed CSI vector $\hat{\mathbf{h}}_A(t)$ of the
     signal from Alice;
3  Bob inputs $\hat{\mathbf{h}}_A(t)$ into the trained GP model and
     predicts $\tilde{PL}_A(t)$;
4  Bob obtains the expected $\hat{PL}_A(t-1)$ from its Table I
     using Alice's identity $ID_A$ and uses $\tilde{PL}_A(t)$ to
     calculate the decision region $\mathcal{D}_A^1(t)$ by (17);
5  if $\hat{PL}_A(t-1)$ is within the decision region $\mathcal{D}_A^1(t)$ then
6      if $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ is close to $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ then
7          Bob accepts this message, replaces $\hat{PL}_A(t-1)$
             with $\tilde{PL}_A(t)$ and $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ with
             $(\hat{\ell}_{A_{t,1}}, \hat{\ell}_{A_{t,2}})$ in Table I (Go to line 14).
8      else
9          Bob accepts this message, replaces $\hat{PL}_A(t-1)$
             with $\tilde{PL}_A(t)$ and keeps $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ in
             Table I (Go to line 14).
10     end if
11 else
12     Bob rejects this message. Keeps $\hat{PL}_A(t-1)$ and
         $(\hat{\ell}_{A_{(t-1),1}}, \hat{\ell}_{A_{(t-1),2}})$ in Table I.
13 end if
14 Bob calculates $\varepsilon(t) = \tilde{PL}_A(t) - \hat{PL}_A(t-1)$, set
     $t = t+1$;
15 Bob calculates $\bar{\varepsilon}(t)$ through (18).
```

IV. Numerical Results

In this section, we study the performance of the proposed PLA scheme. To model both mobile and stationary IoV scenarios, we use QuaDRiGa to simulate multipath fading channels and obtain CSI and PL data. QuaDRiGa generates radio channel impulse response (CIR) and PL data according to realistic communication scenarios for system-level simulations. First, we explain the methodology for the dataset generation for the training and testing of the GP model. After that, we analyze the performance results of the proposed scheme.

A. Generation of Dataset and GPR Model Training

To train the GP model and evaluate the performance of the proposed PLA approach, we use one legitimate vehicle named Alice and one legitimate RSU named Bob through the time evaluation function in QuaDRiGa and customize their trajectories and generate continuous channel impulse responses (CIR) that vary when Alice moves along her trajectory. We sampled 3000 datasets each for Alice and Bob for the training. For the test datasets, we sampled 1500 datasets each for Alice and Bob. We also use one attacker vehicle named Eve, which is randomly located around Bob and Alice and we sampled 1500 datasets for Eve to test the attack mitigation performance of our approach. The settings of the parameters for the communication scenarios in the customized time evolution function of the QuaDRiGa are based on the terrestrial urban microcell parameters estimated in the Berlin survey, Germany (Berlin Uma) with 64 subcarriers and a carrier frequency of 2.53 GHz. After the training of the GP model, we test its prediction performance and plot the results in Fig. 3. We use mean square prediction error (MSE) expressed as

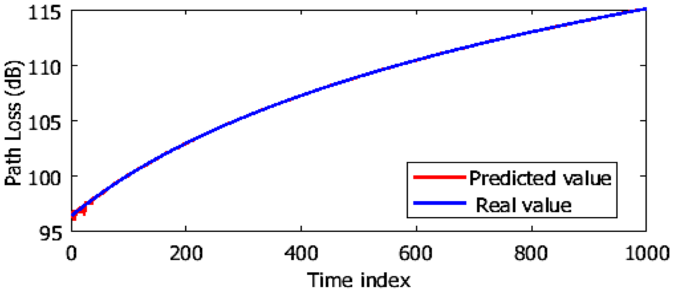


Fig. 3. Prediction performance.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\tilde{PL}_A(i) - \hat{PL}_A(i-1)\right), \tag{19}$$

where N is the number of the test sample to evaluate the prediction performance of the trained GP model and is about 0.01. The result of the prediction from Fig. 3 shows that our trained GP model can accurately capture the variation of the PL features of the channel and further improve the authentication process in the IoV environment.

B. Evaluation Metrics

We use *false alarm rate* (FA) and *miss detection rate* (MD) as metrics to evaluate the performance of the proposed authentication approaches. FA is defined as the rate at which messages from a legitimate vehicle are erroneously rejected as attack attempts, while MD refers to the rate at which the messages from the attacker are incorrectly accepted as legitimate ones. FA and MD are respectively expressed as

$$FA = \frac{\text{number of rejected legitimate messages}}{\text{total number of legitimate messages}}, \tag{20}$$

$$MD = \frac{\text{number of accepted attack messages}}{\text{total number of attack messages}}. \tag{21}$$

C. Performance Results and Discussions

In this subsection, we evaluate the performance of our approach under different values of the decision range parameter z and varying trajectories of Eve.

1) Overall FA and MD of the Proposed Scheme

We first study the FA and MD rates of our scheme with varying values of z, 64 subcarriers, and SNR = 5 dB. In this experiment, Eve and Alice move on a linear trajectory of 600 m. The suitable decision region can be chosen according to the different requirements of the FA. In this article, we are interested in the decision region between 3.6 to 5, i.e., $z \in [3.6, 5]$. As shown in Fig. 4, our scheme achieves the average FA and MD rates of 0.14 and 0.17, respectively. We observe that in the interested decision region, the FA decreases with the increase in z while the MD increases as z increases. Based on this result, we can conclude that the value of the parameter z is significant in determining the FA rate of the scheme and its value should be adjusted based on the accuracy requirement.

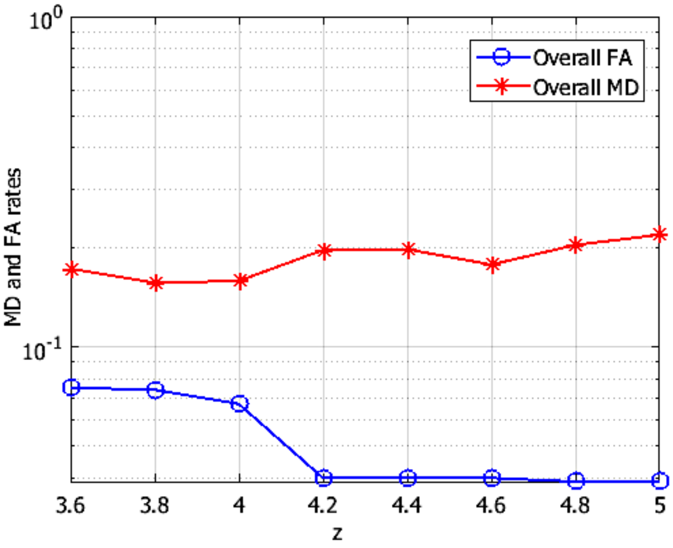


Fig. 4. Performance with varying decision region parameter z.

2) FA and MD of the Proposed Scheme Under Different Eve's Trajectories

Finally, we compare the performance of our scheme under two trajectories of Eve, one is a linear track of 600 m with Eve starting position set to 20 m away from Bob (”Near” case) and the other is a linear track of 600 m with Eve's starting position set to 30 m away from Bob (”Far” case) as depicted in Fig. 5. This test in important because distance significantly affects the quality of the received signal, which in turn affects the performance of the proposed approach. Similar to the first evaluation, our interested decision region is also $z \in [3.6, 5]$, with 64 subcarriers and SNR = 5 dB. Fig. 6 shows the performance of our scheme from which we can make two observations. First, like the first evaluation above, the FA decreases with the increase in z while the MD increases as z increases. Second, the ”Far” case shows a decrease in MD and an increase in FA than the ”Near” case. This is because the further Eve is from Bob, the more unstable and decorrelated her channel becomes compared to that of Alice and thus, the less likely it becomes for her signal to be accepted by Bob as though it is from Alice. Moreover, the more Alice moves further away from Bob, the more unstable her channel becomes, which leads to a slight increase in the FA rate of the ”Far” scenario. Overall, the results have demonstrated the effectiveness of our scheme both under varying ranges of the decision region parameter z and the trajectory of Eve. Moreover, the results have shown that the integration of PL as a security parameter has effectively addressed the limitations of the existing schemes and improved the performance of PLA in the IoV environment.

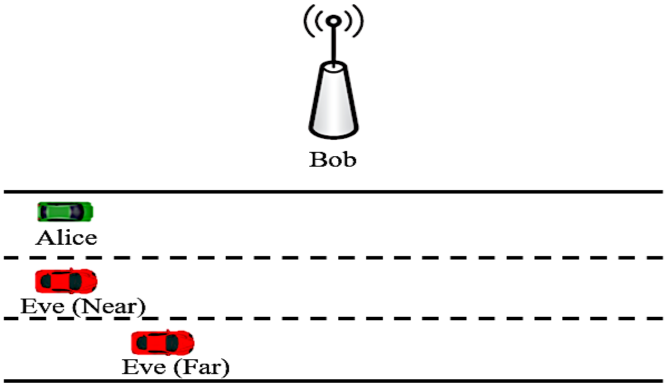


Fig. 5. Scenario design.

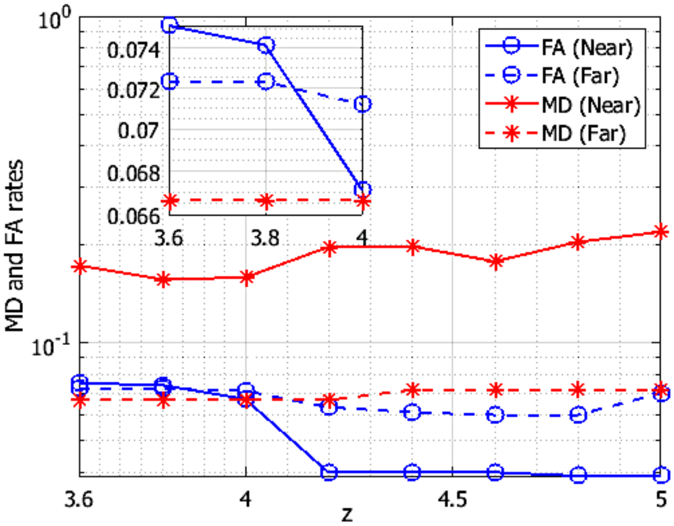


Fig. 6. Performance comparison under two trajectories of Eve.

V. Conclusion

This article proposed a novel PLA approach to address the challenges of the existing PLA schemes in the IoV environment. We adopted GP to predict the PL of a signal based on the relationship between the historical PL and CSI, and then compare the predicted PL with the expected one for authentication. We conducted simulations to demonstrate the authentication performance of our approach through channel attributes generated using the QuaDRiGa simulation platform. The results of the simulations have shown that our scheme can detect 83% of attack attempts with a false alarm rate of just 14%. Due to the scheme's design simplicity, we believe it is an appropriate choice to secure IoV application scenarios.

Acknowledgment

This work is supported by Jinan 'New Universities 20 Items' Introduced Innovation Team Project (No. 2021GXRC064) and NSFShandong (No. ZR2021LZH006).

References

[1] H. Qiu, M. Qiu, and R. Lu, "Secure v2x communication network based on intelligent pki and edge computing," *IEEE Network*, vol. 34, no. 2, pp. 172–178, 2020.

[2] D. Xu, P. Ren, and J. A. Ritcey, "Phy-layer cover-free coding for wireless pilot authentication in iov communications: Protocol design and ultra-security proof," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 171–187, 2019.

[3] N. V. Abhishek, M. N. Aman, T. J. Lim, and B. Sikdar, "Drive: Detecting malicious roadside units in the internet of vehicles with low latency data integrity," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 3270–3281, 2022.

[4] J. Contreras-Castillo, S. Zeadally, and J. A. Guerrero-Ibañez, "Internet of vehicles: Architecture, protocols, and security," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3701–3709, 2018.

[5] L. Wang and X. Liu, "Notsa: Novel obu with three-level security architecture for internet of vehicles," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3548–3558, 2018.

[6] M. Wazid, P. Bagga, A. K. Das, S. Shetty, J. J. P. C. Rodrigues, and Y. Park, "Akm-iov: Authenticated key management protocol in fog computing-based internet of vehicles deployment," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8804–8817, 2019.

[7] J. McEntyre and B. Kihei, "Zero-knowledge proof for enabling privacy-preserving electronic toll collection with vehicle-to-everything communications," in *2022 IEEE International Conference on Consumer Electronics (ICCE)*, 2022, pp. 1–6.

[8] R. Ma, J. Cao, Y. Zhang, C. Shang, L. Xiong, and H. Li, "A group-based multicast service authentication and data transmission scheme for 5g-v2x," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, pp. 23976–23992, 2022.

[9] H. Tan, Z. Gui, and I. Chung, "A secure and efficient certificateless authentication scheme with unsupervised anomaly detection in vanets," *IEEE Access*, vol. 6, pp. 74260–74276, 2018.

[10] M. Yang, S. Wei, R. Jiang, F. Ali, and B. Yang, "Single-message-based cooperative authentication scheme for intelligent transportation systems," *Computers & Electrical Engineering*, vol. 96, p. 107390, 2021.

[11] H. Qiu, M. Qiu, and R. Lu, "Secure v2x communication network based on intelligent pki and edge computing," *IEEE Network*, vol. 34, no. 2, pp. 172–178, 2020.

[12] S. Taha, M. Alhassany, and X. Shen, "Lightweight handover authentication scheme for 5g-based v2x communications," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[13] J. Zhang, H. Zhong, J. Cui, Y. Xu, and L. Liu, "Smaka: Secure many-to-many authentication and key agreement scheme for vehicular networks," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1810–1824, 2021.

[14] H.-M. Wang and Q.-Y. Fu, "Channel-prediction-based one-class mobile iot device authentication," *IEEE Internet of Things Journal*, vol. 9, no. 10, pp. 7731–7745, 2022.

[15] J. Wang, Y. Shao, Y. Ge, and R. Yu, "Physical-layer authentication based on adaptive kalman filter for v2x communication," *Vehicular Communications*, vol. 26, p. 100281, 2020.

[16] S. Wang, K. Huang, X. Xu, Z. Zhong, and Y. Zhou, "Csi-based physical layer authentication via deep learning," *IEEE Wireless Communications Letters*, vol. 11, no. 8, pp. 1748–1752, 2022.

[17] A. Abdelaziz, R. Burton, F. Barickman, J. Martin, J. Weston, and C. E. Koksal, "Enhanced authentication based on angle of signal arrivals," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4602–4614, 2019.

[18] J. Wang, Y. Shao, Y. Wang, Y. Ge, and R. Yu, "Physical layer authentication based on nonlinear kalman filter for v2x communication," *IEEE Access*, vol. 8, pp. 163746–163757, 2020.

[19] X. Yin, X. Fang, N. Zhang, P. Yang, X. Sha, and J. Qiu, "Online learning aided adaptive multiple attribute-based physical layer authentication in dynamic environments," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1106–1116, 2021.

[20] H. Fang, X. Wang, and L. Hanzo, "Learning-aided physical layer authentication as an intelligent process," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2260–2273, 2019.

[21] Y. Chen, P.-H. Ho, H. Wen, S. Y. Chang, and S. Real, "On physical-layer authentication via online transfer learning," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 1374–1385, 2022.

[22] F. Pan, Z. Pang, H. Wen, M. Luvisotto, M. Xiao, R.-F. Liao, and J. Chen, "Threshold-free physical layer authentication based on machine learning for industrial wireless cps," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6481–6491, 2019.

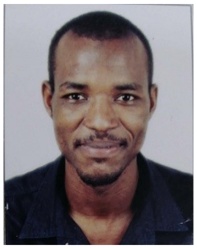
[23] V.-L. Nguyen, P.-C. Lin, and R.-H. Hwang, "Enhancing misbehavior detection in 5g vehicle-to-vehicle commu-

nications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9417–9430, 2020.

[24] S. So, J. Petit, and D. Starobinski, "Physical layer plausibility checks for misbehavior detection in v2x networks," in *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*, ser. WiSec '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 84–93.

[25] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[26] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "Quadriga: A 3-d multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 6, pp. 3242–3256, 2014.



Mubarak Umar received his B.Sc. and M.Sc. degrees in Computer Science from Bayero University, Kano, Nigeria, in 2011 and 2015, respectively, and the Ph.D. degree in Computer Software and Theory from Shaanxi Normal University, China. He is also a Lecturer with the Department of Information Technology, Bayero University, Kano, Nigeria. He is currently a postdoctoral researcher at the School of Computer Science and Technology, Xidian University, China. His research interests include physical layer authentication in the Internet of Vehicles (IoV) and the Internet of Medical Things (IoMT), and information security of wireless communication.

Jiandong Wang received his M.S. degree from Xidian University, Xi'an, China, in 2015. He is studying for a doctorate in the School of Computer Science and Technology at Xidian University. His research interests include smart cities and IoT.

Dr. Lei Liu is a full professor at the school of software at Shandong University. He obtained his M.S. and Ph.D. degrees in 2005 and 2010 from Bradford University, UK, respectively. Dr. Liu has published over 70 research papers in international conferences and journals. His research interest includes network performance engineering, 5g technology, quality of service, IoT, and UAVs.

Zewei Guo received his B.S. degree in Software Engineering from Shanxi University, Taiyuan, China, in 2015 and his M.S. degree in Computer Technology from Xidian University, Xi'an, China, in 2019. His research interest focuses on covert communication.

Shuguang Wang is a senior engineer at Shandong Institute of Standardization, No.146-6, Lishan Road, Jinan, China. He received his B.S. and M.S. degrees from Shandong University, Jinan, China, in 1998 and 2010, respectively. He is studying for a doctorate in the School of Computer Science and Technology at Xidian University. His research interests include smart cities, cyber security, and data security.