Reinforcement Learning Based Physical-layer Authentication for Controller Area Networks

Liang Xiao*, Xiaozhen Lu*, Tangwei Xu*, Weihua Zhuang†, Huaiyu Dai‡
*Dept. of Information and Communication Engineering, Xiamen University, Xiamen, China. Email: lxiao@xmu.edu.cn
†Dept. of Electrical and Computer Engineering, University of Waterloo, Waterloo, Canada. Email: wzhuang@uwaterloo.ca
‡Dept. of Electrical and Computer Engineering, NC State Univesity, Raleigh, NC. Email: huaiyu_dai@ncsu.edu

***Abstract*—In controller area networks (CANs), electronic control units (ECUs) such as telematics ECUs and on-board diagnostic ports must protect the message exchange from spoofing attacks. In this paper, we propose a CAN bus authentication framework that exploits physical layer features of the messages, including message arrival intervals and signal voltages, and applies reinforcement learning to choose the authentication mode and parameter. By applying the Dyna architecture and using a double estimator, this scheme improves the utility in terms of authentication accuracy without changing the CAN bus protocol or the ECU components and requiring knowledge of the spoofing model. We also propose a deep learning version to further improve the authentication efficiency for the CAN bus. The learning scheme applies a hierarchical structure to reduce the exploration time, and uses two deep neural networks to compress the high-dimensional state space and to fully exploit the physical authentication experiences. We provide the computational complexity and the performance analysis. Experimental results verify the theoretical analysis and show that our proposed schemes significantly improve the authentication accuracy as compared with benchmark schemes.**

***Index Terms*—Controller area networks, authentication, spoofing attacks, reinforcement learning.**

I. Introduction

Electronic control units (ECUs) in controller area networks (CANs) exchange service information and control signals to support advanced driver assistance systems and comfort applications for vehicles [2]–[4]. However, CAN bus has to detect spoofing attacks that are launched via social engineering, worms at dealerships and the air update [5].

Spoofing attackers can compromise the in-vehicle onboard diagnostic (OBD-II) ports and telematics ECUs, and keep sending spoofing messages with high priority to the CAN bus [6]–[8], which further results in denial-of-service attacks and even life-threatening disasters to drivers and passengers. A smart attacker can sniff the CAN bus status, apply data mining to exploit the message ID database such as [9] and use an automotive diagnostic tool such as VAS5052 developed by Audi/Volkswagen in [10], to obtain the ID of each command message. More specifically, the attackers can store some important messages such as the braking command and send these messages later with the falsified ID from the compromised ECUs, which results in man-in-the-middle attacks and further controls the steering column of the vehicles to cause severe traffic accidents [5], [11]–[14].

Current CAN bus authentication mostly uses specific message authentication codes (MACs) to verify the ECU messages [10], [15]. For instance, the CAN bus as designed in [10] uses encryption session keys to build the MAC and to authenticate the data frames. However, the session keys required by these authentication schemes are not always applicable due to the limited communication bandwidth of the CAN bus, the restricted computational sources of the ECUs and the weak integrity checking [16]. On the other hand, the in-vehicle encryption in [17] has to choose a secure ECU to generate the long-term symmetric keys and may suffer from high computational and communication overheads for a CAN bus with restricted computational resources.

To the best of our knowledge, CAN buses usually apply unidirectional authentication such as the signal voltage based intrusion detection scheme in [18] instead of the bidirectional schemes due to the limited computational and bandwidth capacity. For example, the four-way authentication in [19] that uses 256 bits in the handshake process to verify each message is not applicable to the CAN bus message with less than 128 bits according to [18].

Therefore, physical (PHY) layer authentication schemes are further considered, which exploit PHY features such as the message arrival intervals [20]–[23] and signal voltages [8], [18], [24], [25] to detect spoofing messages. More specifically, the spoofing attackers that use the compromised ECU at a lower sampling rate than the monitor cannot accurately obtain the signal voltages of the other ECUs. Due to the distinctive hardware configurations of the CAN bus, especially the transistors with drain-to-source on-state resistance [18], [25], the inimitable signal voltages cannot be forged by the attackers. Each message sent by the ECUs has a unique arbitration ID that represents the transmission priority according to [26]. The message periodicity indicates whether the CAN bus receives a spoofing message with the assigned arbitration ID [27].

As a novel physical feature based automotive intrusion detection scheme, VoltageIDS in [25] uses the signal voltages to detect both masquerade attacks and bus-off attacks without any modulation of the current CAN bus systems. In addition, the intrusion detection scheme named Viden as proposed in [18] uses the voltage profiles as fingerprints to pinpoint the malicious ECUs. However, a CAN bus has to use faster authentication to support time sensitive applications and emergency control signals.

In this paper, we propose a CAN bus authentication scheme that exploits the arrival intervals of the periodic messages and signal voltages of messages to detect spoofing attacks and the resulting man-in-the-middle attacks, replay attacks and denial-of-service attacks. More specifically, this scheme

This paper was presented in part at IEEE ICC 2019 [1].

samples the waveform of the received message and measures the output voltages from the two dedicated wires, called CAN-High (CANH) and CAN-Low (CANL), which are used to transmit messages to avoid electromagnetic interference [28]. This scheme also measures the arrival interval from the previous message with the same arbitration ID if the claiming arbitration ID is periodic.

Instead of using the heuristic algorithms (such as the genetic algorithms [29]) that achieve the local optimal policy or the watermarking techniques (such as the semi-fragile watermark [30]) that focus on the image authentication, we apply reinforcement learning (RL) to optimize the authentication policy for the monitor without knowing the frequency of spoofing messages sent by the attacker. More specifically, our scheme applies reinforcement learning to choose the authentication mode and test threshold in a hypothesis test that compares the physical layer features with the records for improving the authentication accuracy.

The authentication policy is chosen based on the state that consists of the message priority, number of total accepted messages, and number of the falsely accepted messages in a given time duration. By applying the Dyna architecture, this scheme exploits the simulated authentication experiences for planning to reduce random exploration in the original authentication process. Based on a double estimator, this scheme avoids the over-estimation of the Q-values and thus the suboptimal authentication policies. Different from VoltageIDS [25] that relies on signal voltage features and Viden [18] that depends on signal voltage profiles, our scheme relies on both the arrival intervals of periodic messages and signal voltages to improve the utility consisting of the message priority and authentication accuracy, without depending on any labelled signal voltages.

We further propose a deep RL based CAN bus authentication scheme, which improves the authentication accuracy and reduces the optimization latency based on a hierarchical structure [31], and design two deep neural networks (DNNs), i.e., top-level DNN and bottom-level DNN. It uses the two DNNs to compress the high-dimensional state space and to fully exploit the PHY authentication experiences, each of which contains two fully connected (FC) layers rather than the convolutional layers to quickly extract the authentication features and thus reduce the sample complexity. The top-level DNN outputs the authentication mode distribution of the current state. With the chosen authentication mode and state as input, the bottom-level DNN outputs the test threshold distribution. By applying the experience replay technique to update the DNN weights, the proposed scheme improves the learning efficiency in the CAN bus authentication.

The computational complexity and performance are provided to evaluate the authentication efficacy. Experiments are performed on a CAN bus connected to 18 legitimate ECUs, a monitor and a compromised ECU. Experimental results show that our proposed schemes improve the authentication accuracy, as compared with existing schemes.

The rest of this paper is organized as follows. Section II gives an overview of the related work. Section III presents the CAN bus model, followed by the CAN authentication framework in Section IV. We propose an RL based CAN bus authentication scheme in Section V and a deep RL version in Section VI. The computational complexity and performance are provided in Section VII, followed by the experimental results in Section VIII. We draw the conclusion and discuss the future work in Section IX. For ease of reference, some important symbols are summarized in Table I.

TABLE I
LIST OF IMPORTANT SYMBOLS

Symbol	Description
L	Number of the ECUs
N_{L}	Number of the legitimate messages in T time slots
Γ	Number of the feasible test thresholds
N_{Y}	Maximum number of the spoofing messages in T time slots
f_{S}	Message voltage sampling rate
J	Maximum number of the IDs
$d^{(k)}$	Arbitration ID of the message received at time slot k
$l^{(k)}$	Periodicity of the message received at time slot k
$\rho^{(k)}$	Transmission priority of the message at time slot k
W	Number of the voltage samples for each message
$\{\nu_i^{(k)}/\mu_i^{(k)}\}_{1\leq i\leq W}$	Extracted CANH/CANL voltages of the message received at time slot k
$\tau^{(k)}$	Arrival interval from previous message with arbitration ID $d^{(k)}$
$N_{\mathrm{A}}^{(k)}$	Number of the received messages in T time slots
$N_{\mathrm{P}}^{(k)}$	Number of the accepted messages in the previous $N_{\mathrm{A}}^{(k)}$ messages
$N_{\mathrm{F}}^{(k)}$	Number of the falsely accepted messages in the previous $N_{\mathrm{A}}^{(k)}$ messages
c_{F}	Importance of the falsely accepted messages
c_{T}	Authentication latency coefficient

II. RELATED WORK

In this section, we review some relevant works to our study, besides those already mentioned above. A group of pioneering CAN bus intrusion detection schemes apply supervised learning to evaluate the measured signal voltages based on the labelled voltage dataset. For example, Choi *et al*. propose an automotive intrusion detection scheme in [25], which uses support vector machine (SVM) and bagged decision tree classification model to evaluate the signal voltages in the detection of both masquerade attacks and bus-off attacks accurately. The spoofing detection scheme in [28] applies

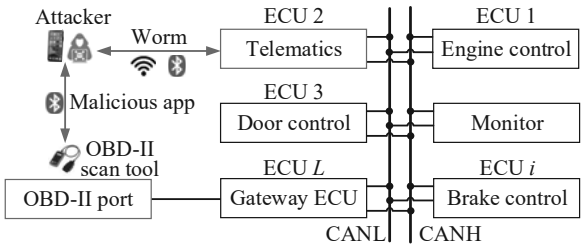


Fig. 1. Illustration of the CAN bus with a monitor and L ECUs, in which an outside attacker compromises the telematics ECU and/or the OBD-II port.

SVM, neural network and bagged decision tree to improve the spoofing detection accuracy at a high sampling rate. The spoofing detection system in [24] reduces the required sampling rate of the voltage signals for a low-speed CAN bus. The signal voltage based spoofing detection scheme in [8] uses the logistic regression to further reduce the sampling rate for a high-speed CAN bus. The CAN bus authentication scheme in [32] applies the maximum-likelihood estimation for the Poisson distribution based data arrival interval scenarios.

Message arrival intervals can be used to verify the periodic messages for CAN buses. For instance, the intrusion detection algorithm in [20] that builds a hypothesis test with a fixed threshold to evaluate the message arrival intervals for the detection of both injection attacks and denial-of-service attacks suffers from a high false alarm rate if the attacker sends spoofing messages with high priority. The CAN bus intrusion detection system in [21] that exploits the message arrival intervals to detect fabrication attacks, suspension attacks, and masquerade attacks suffers from a low detection accuracy for periodic messages with lower priority, due to the changed message arrival intervals. The CAN-FD authentication scheme as developed in [22] applies Bloom filtering to measure the message periodicity to detect replay attacks, which also suffers from detection performance degradation for the periodic messages with low priority due to the queuing delay. The anomaly detection approach proposed in [23] applies a distributed long-short-term-memory framework to evaluate the arrival intervals for intra-vehicle networks that support deep neural networks but sometimes falsely views the delayed low priority messages as spoofing.

Reinforcement learning has been used to optimize the test threshold for wireless authentication without the knowledge of attack strategies. For instance, the authentication scheme in [33] applies Q-learning and Dyna-Q to optimize the test threshold and thus improve the authentication accuracy. The authentication scheme in [34] uses neural episodic control to select the authentication data such as the received signal strength indicators and the arrival intervals of the ambient radio signals for wireless networks instead of CAN buses.

III. System Model

A. Network Model

In this work, we consider a CAN bus in a vehicle that can transport sensitive information such as braking command and monitor the vehicle's status, which consists of a monitor and L ECUs such as the engine control ECU, the telematics ECU, and the routing gateway ECU, as shown in Fig. 1. All the L ECUs are connected via two dedicated wires, i.e., CANH and CANL. According to [35] and [36], the OBD-II port is assumed to have unrestricted access to the high-speed CAN bus with baud rate up to 1 Mbps or the low-speed CAN bus with baud rate lower than 125 Kbps via a routing gateway ECU instead of an access control gateway ECU. For example, the CAN bus uses a diagnostic tool to read the received messages and uses a scan tool to send messages through the OBD-II port. The ECUs have distinctive hardware configurations [18]. Each ECU generates a waveform with $2.75 \sim 4.5$ V on the CANH and $0.5 \sim 2.25$ V on the CANL to send a dominant bit (0) [28]. The recessive bit (1) usually corresponds to the waveform at $2.25 \sim 2.75$ V voltages on both CANH and CANL.

Radio devices such as smartphones send busty messages with varying arrival interval ranging from 10 ms to 500 ms via the telematics ECU or the OBD-II port using Bluetooth and WiFi. On the other hand, the engine control ECU sends the engine revolution message with arbitration ID 0x2C4 every 24 ms, and the steering ECU sends the angle message with arbitration ID 0x025 every 20 ms [21]. Time is partitioned to slots of equal length. Let $l^{(k)} \in \{0, 1\}$ indicate the periodicity of the message received at time slot k. If $l^{(k)} = 1$, the message is periodic with the message arrival interval denoted by $\tau^{(k)}$.

Without loss of generality, we assume that at most one ECU sends a message in each time slot to the CAN bus. The ECU that is assigned with lower arbitration ID gains access to the CAN bus if there are more than one ECU sending messages simultaneously. The intra-vehicle network has J feasible arbitration IDs. Each message has a unique arbitration ID, which represents the message type and priority. At time slot k, the CAN bus receives a message with ID $d^{(k)} \in \{1, 2, \cdots, J\}$ corresponding to ECU j. According to [26], the priority of the message denoted by $\rho^{(k)}$ decreases with the arbitration ID $d^{(k)}$, with $\rho^{(k)} = 0$ when $d^{(k)} = J$, i.e., $\rho^{(k)} = 1 - d^{(k)}/J$. According to [21], a higher priority message sent by other ECUs can change the arrival intervals of the periodic messages.

Each message consists of at most 128 dominant bits (0) and recessive bits (1), including the start of the frame (SOF), the arbitration ID, the extended ID field, the data, the acknowledgement (ACK) and the end of the frame (EOF), as shown in Fig. 2. According to [25], the voltages of the dominant bits (0) in the extended ID field are more active than the other parts of the message. Upon receiving a message, the monitor uses an analog-to-digital converter at rate $f_{\rm S}$ to measure the waveform of the dominant bits (0) in the extended ID field of the message and obtain W CANH voltages $\{\nu_i^{(k)}\}_{1 \le i \le W}$ and W CANL voltages $\{\mu_i^{(k)}\}_{1 \le i \le W}$. A falsely accepted message may cause the abnormal behavior such as the failure fuel level in [2].

B. Intra-vehicle Attack Model

We consider an attacker, Eve, who compromises a telematics ECU or OBD-II port, and sends a number of the spoofing mes-

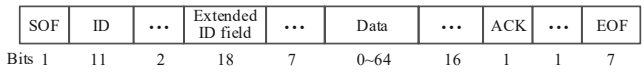


Fig. 2. CAN bus message format with $64 \sim 128$ bits.

sages with the arbitration IDs of the other ECUs [12], which compromises the arrival intervals of the periodic messages. The spoofing message can result in the falsification of the fuel level, the instruction failure alarm on the instrument panels, and the control lost of the engine and brake of the vehicle [11].

Eve installs a malware such as a malicious self-diagnostic application (app) from the Android application Market on her smartphone to compromise the OBD-II port via the OBD-II scan tool [10], as shown in Fig. 1. Eve can also inject worms such as à la Stuxnet to compromise the telematics ECU [5], and control it to send spoofing messages. By storing some important messages such as the braking command, Eve controls the compromised ECUs to send these messages later with the falsified ECU IDs, which results in man-in-the-middle attacks. Eve can also keep sending spoofing messages with high priority via the compromised ECUs to result in denial-of-service attacks [37].

According to [21] and [18], Eve changes the arrival intervals of the periodic messages once sending out a spoofing message to the CAN bus and cannot forge the signal voltage of another ECU due to the distinct drain-to-source on-state resistance of the transistor. Eve is assumed to send at most $N_{\rm Y}$ spoofing messages in T time slots, and choose the number of spoofing messages to send to the CAN bus in the future T time slots, denoted by $y^{(k)} \in \{0, 1, \cdots, N_{\rm Y}\}$ that is claimed to be sent by ECU i, which aims to prevent ECU i from sending these messages.

IV. CAN Bus Authentication Framework

We propose an unidirectional based CAN bus authentication framework that exploits the physical layer information, including message arrival intervals and signal voltages to detect spoofing attacks, which saves the computational and communication overhead compared with the bidirectional authentication. More specifically, a monitor connects to the CAN bus and measures the physical features to determine whether the message under test is indeed sent by the ECU that is assigned with the claimed arbitration ID. This framework provides lightweight authentication services without changing the CAN bus protocol or the ECU components in intra-vehicle systems.

Without loss of generality, the CAN bus is assumed to receive a message that claims arbitration ID $d^{(k)}$ at time slot k, which is either sent by the claimed ECU or an attacker Eve who fakes the message. This framework monitors the waveform of the dominant bits (0) in the extended ID field of the message as shown in Fig. 2 and measures the corresponding signal voltages at W uniformly distributed time. The resulting CANH voltages are $\{\nu_i^{(k)}\}_{1 \le i \le W}$ and the CANL voltages are $\{\mu_i^{(k)}\}_{1 \le i \le W}$. The monitor evaluates the message priority $\rho^{(k)}$ and measures the message arrival interval $\tau^{(k)}$ from the previous authenticated message with this ID if the message with the claimed arbitration ID $d^{(k)}$ is periodic.

The scheme chooses the authentication features or mode denoted by $x_1^{(k)} \in \{0, 1\}$. If $x_1^{(k)} = 0$, the authentication depends on the difference between the current signal voltages and the samples of the previous authenticated message with arbitration ID $d^{(k)}$. If $x_1^{(k)} = 1$, the authentication depends on both the signal voltage samples and the message arrival interval. In the framework, authentication mode 0 is more effective for non-periodic messages, while authentication mode 1 is more accurate for periodic messages. In particular, the framework switches to the signal voltages if the message arrival interval is known by the attacker or changed by the message transmission failure or higher priority messages.

This framework builds a z-score based hypothesis test to compare the current physical layer features with the feature records. The test statistic denoted by Δ is compared with the test threshold $x_2^{(k)}$, which is equally quantized into Γ levels, i.e., $x_2^{(k)} \in \{i/\Gamma | 1 \le i \le \Gamma\}$. If $\Delta \le x_2^{(k)}$, the monitor accepts the message and updates the voltage records of the message. Otherwise, the monitor rejects the message and sends an active error flag to the CAN bus rather than a specific ECU. After the authentication, a radio device such as a smartphone outside the intra-vehicle network sends a feedback signal to the monitor if there are some abnormal behaviors such as the falsification of the fuel level caused by the falsely accepted message. According to [18], the monitor must finish the authentication before the ACK of the message, as shown in Fig. 2.

V. RL Based CAN Bus Authentication

We propose an RL based CAN bus authentication scheme (RLA), which chooses both the authentication mode and test threshold $[x_1^{(k)}, x_2^{(k)}]$. This authentication scheme uses the Dyna architecture to simulate hypothetical authentication experiences to update the learning parameters such as the Q-values, and applies a double estimator to avoid over-estimation.

Upon receiving the message at time slot k, the authentication formulates the current state denoted by $\boldsymbol{s}^{(k)}$ that consists of message priority $\rho^{(k)}$, the number of the accepted messages in the authentication, $N_{\rm P}^{(k-1)}$, and the number of the falsely accepted messages in the previous T time slots, $N_{\rm F}^{(k-1)}$, that is obtained from the feedback from the radio device, i.e.,

$$\boldsymbol{s}^{(k)} = \left[\rho^{(k)}, N_{\rm P}^{(k-1)}, N_{\rm F}^{(k-1)}\right]. \tag{1}$$

Based on a double estimator, the authentication policy $\boldsymbol{x}^{(k)} = [x_1^{(k)}, x_2^{(k)}]$ is chosen based on ϵ-greedy using the sum of the first Q-value $\boldsymbol{Q}^{\rm A}(\boldsymbol{s}^{(k)}, \cdot)$ and second Q-value

Algorithm 1: RL based CAN bus authentication

Initialize λ, β, J, W, T, $N_{\mathrm{A}}^{(0)}$, $N_{\mathrm{P}}^{(0)}$, $N_{\mathrm{F}}^{(0)}$, $\{\bar{\tau}_i\}_{1\le i\le J}$, $\{\bar{\nu}_i\}_{1\le i\le J}$, $\{\tilde{\nu}_i\}_{1\le i\le J}$, $\{\bar{\mu}_i\}_{1\le i\le J}$, $\{\tilde{\mu}_i\}_{1\le i\le J}$, $\boldsymbol{Q}^{\mathrm{A}}=\mathbf{0}$, and $\boldsymbol{Q}^{\mathrm{B}}=\mathbf{0}$
for $k=1,2,3,...$ **do**
  Read the arbitration ID $d^{(k)}$
  $\rho^{(k)}=1-\frac{d^{(k)}}{J}$
  Extract W CANH voltages $\left\{\nu_i^{(k)}\right\}_{1\le i\le W}$
  Extract W CANL voltages $\left\{\mu_i^{(k)}\right\}_{1\le i\le W}$
  if Message with arbitration ID $d^{(k)}$ is periodic **then**
    Measure $\tau^{(k)}$
  end if
  Form $\boldsymbol{s}^{(k)}$ via (1)
  Select $\boldsymbol{x}^{(k)}=\left[x_i^{(k)}\right]_{1\le i\le 2}$ via (2)
  Calculate Δ via (3)
  if $\Delta\le x_2^{(k)}$ **then**
    Accept the message
    Update the voltage records
  else
    Send an active error flag
  end if
  Count $N_{\mathrm{A}}^{(k)}$ and $N_{\mathrm{P}}^{(k)}$
  Measure $N_{\mathrm{F}}^{(k)}$
  Compute $u^{(k)}$ via (4)
  Randomly choose to update $\boldsymbol{Q}^{\mathrm{A}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)}\right)$ via (5) or $\boldsymbol{Q}^{\mathrm{B}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)}\right)$ via (6)
  $\mathbf{C}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)},\boldsymbol{s}^{(k+1)}\right)=\mathbf{C}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)},\boldsymbol{s}^{(k+1)}\right)+1$
  Calculate M via (7)
  $\mathbf{R}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)},M\right)=u^{(k)}$
  for $i=1,2,\cdots,D$ **do**
    Randomly choose a simulated state-action pair $(\tilde{\boldsymbol{s}},\tilde{\boldsymbol{x}})$
    Choose $\tilde{\boldsymbol{s}}'$ with probability $\frac{1}{M}\mathbf{C}\left(\tilde{\boldsymbol{s}},\tilde{\boldsymbol{x}},\tilde{\boldsymbol{s}}'\right)$
    Compute $\tilde{r}$ via (8)
    Randomly choose to update $\boldsymbol{Q}^{\mathrm{A}}\left(\tilde{\boldsymbol{s}},\tilde{\boldsymbol{x}}\right)$ via (5) or $\boldsymbol{Q}^{\mathrm{B}}\left(\tilde{\boldsymbol{s}},\tilde{\boldsymbol{x}}\right)$ via (6)
  end for
end for

$\boldsymbol{Q}^{\mathrm{B}}(\boldsymbol{s}^{(k)},\cdot)$. The policy distribution is given by

$$\Pr\left(\boldsymbol{x}^{(k)}=\hat{\boldsymbol{x}}\right)=\begin{cases}1-\epsilon, & \hat{\boldsymbol{x}}=\arg\max\limits_{\boldsymbol{x}'\in\mathbf{X}}\left\{\boldsymbol{Q}^{\mathrm{A}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}'\right)+\boldsymbol{Q}^{\mathrm{B}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}'\right)\right\}\\ \epsilon, & \text{o.w}\end{cases}\tag{2}$$

where $\mathbf{X}$ is the action set that consists of the 2Γ feasible authentication policies.

The message arrival interval records of the J different types of messages $\{\bar{\tau}_i\}_{1\le i\le J}$ are initialized by the vehicle manufacturers. In particular, the arrival interval records of the non-periodic messages are set as 0. The voltage records $\{\bar{\nu}_i\}_{1\le i\le J}$, $\{\tilde{\nu}_i\}_{1\le i\le J}$, $\{\bar{\mu}_i\}_{1\le i\le J}$, and $\{\tilde{\mu}_i\}_{1\le i\le J}$ are initialized with the Viden scheme in [18]. Let $\mathbf{I}(\cdot)$ be the indictor function, with value 1 if true and 0 otherwise. This scheme builds the test statistic Δ as follows

$$\Delta=\frac{1}{\tilde{\nu}_d}\left(\frac{\sum_{i=1}^{W}\nu_i^{(k)}}{W}-\bar{\nu}_d\right)^2+\frac{1}{\tilde{\mu}_d}\left(\frac{\sum_{i=1}^{W}\mu_i^{(k)}}{W}-\bar{\mu}_d\right)^2+\mathbf{I}\left(x_1^{(k)}\right)\mathbf{I}\left(l^{(k)}\right)\left(\frac{\tau^{(k)}}{\bar{\tau}_d}-1\right)^2.\tag{3}$$

If $\Delta\le x_2^{(k)}$, the CANH voltage records $\bar{\nu}_d$ and $\tilde{\nu}_d$ are updated with the average and the variance of $\{\nu_i^{(k)}\}_{1\le i\le W}$, respectively. In this case, the CANL voltage records $\bar{\mu}_d$ and $\tilde{\mu}_d$ are updated with $\{\mu_i^{(k)}\}_{1\le i\le W}$.

The monitor counts the number of the received messages denoted by $N_{\mathrm{A}}^{(k)}$ and the number of the accepted messages $N_{\mathrm{P}}^{(k)}$ in the previous $T-1$ time slots and current time slot. The scheme measures the number of the time slots that received the feedback from the radio device outside the intra-vehicle network in the previous $T-1$ time slots and current time slot, which is used as the number of spoofing messages that are falsely accepted in the $N_{\mathrm{A}}^{(k)}$ messages, i.e., $N_{\mathrm{F}}^{(k)}$. The authentication weight denoted by c_{F} represents the importance of the falsely accepted messages, and the latency coefficient denoted by c_{T} represents the importance of the authentication response time in the utility evaluation. This scheme computes the utility that represents the optimization objective via

$$u^{(k)}=\rho^{(k)}N_{\mathrm{P}}^{(k)}-c_{\mathrm{F}}\rho^{(k)}N_{\mathrm{F}}^{(k)}-c_{\mathrm{T}}x_1^{(k)}\tag{4}$$

where three terms are the estimated importance of the accepted messages, the penalty of the falsely accepted messages, and the estimated authentication cost, respectively.

The learning rate denoted by $\lambda\in(0,1]$ is chosen to control the weight of the current experience, and the discount factor $\beta\in(0,1]$ indicates the weight on the future utility. The first Q-value is updated based on the second Q-value $\boldsymbol{Q}^{\mathrm{B}}(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)})$ according to the iterative Bellman equation via

$$\boldsymbol{Q}^{\mathrm{A}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)}\right)\leftarrow(1-\lambda)\boldsymbol{Q}^{\mathrm{A}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)}\right)+\lambda\left(u^{(k)}+\beta\boldsymbol{Q}^{\mathrm{B}}\left(\boldsymbol{s}^{(k+1)},\arg\max_{\boldsymbol{x}'\in\mathbf{X}}\boldsymbol{Q}^{\mathrm{A}}\left(\boldsymbol{s}^{(k+1)},\boldsymbol{x}'\right)\right)\right).\tag{5}$$

Similarly, the second Q-value $\boldsymbol{Q}^{\mathrm{B}}(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)})$ is updated as follows,

$$\boldsymbol{Q}^{\mathrm{B}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)}\right)\leftarrow(1-\lambda)\boldsymbol{Q}^{\mathrm{B}}\left(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)}\right)+\lambda\left(u^{(k)}+\beta\boldsymbol{Q}^{\mathrm{A}}\left(\boldsymbol{s}^{(k+1)},\arg\max_{\boldsymbol{x}'\in\mathbf{X}}\boldsymbol{Q}^{\mathrm{B}}\left(\boldsymbol{s}^{(k+1)},\boldsymbol{x}'\right)\right)\right).\tag{6}$$

According to [38], this scheme randomly chooses to update the first Q-value $\boldsymbol{Q}^{\mathrm{A}}(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)})$ or the second Q-value $\boldsymbol{Q}^{\mathrm{B}}(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)})$ each time slot.

The number of the occurrences from $(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)})$ to $\boldsymbol{s}^{(k+1)}$ is denoted by $\mathbf{C}(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)},\boldsymbol{s}^{(k+1)})$. The number of occurrences of the state-action pair $(\boldsymbol{s}^{(k)},\boldsymbol{x}^{(k)})$ denoted by M is

given by

$$M = \sum_{\hat{\boldsymbol{s}}} \mathbf{C}\left(\boldsymbol{s}^{(k)}, \boldsymbol{x}^{(k)}, \hat{\boldsymbol{s}}\right). \tag{7}$$

The reward record $\mathbf{R}(\boldsymbol{s}^{(k)}, \boldsymbol{x}^{(k)}, M) = u^{(k)}$. By applying the Dyna architecture as in [34], the scheme simulates D hypothetical authentication experiences to update the two Q-values, as summarized in Algorithm 1. More specifically, in the i-th simulation, the monitor randomly chooses a simulated state-action pair $(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{x}})$ and selects the next simulated state $\tilde{\boldsymbol{s}}'$ based on $\mathbf{C}(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{s}}')$ and M. The simulated reward denoted by $\tilde{r}$ is calculated by

$$\tilde{r} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{R}\left(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{x}}, m\right) \tag{8}$$

which replaces $u^{(k)}$ in (5) and (6) to update $\boldsymbol{Q}^{\mathrm{A}}\left(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{x}}\right)$ and $\boldsymbol{Q}^{\mathrm{B}}\left(\tilde{\boldsymbol{s}}, \tilde{\boldsymbol{x}}\right)$.

VI. Deep RL Based CAN Bus Authentication

We propose a deep RL based CAN bus authentication scheme named DRLA in Algorithm 2 to further improve the authentication performance for the CAN buses that support deep learning. The scheme uses a hierarchical structure and two lightweight deep neural networks with the fully connected layers to quickly extract the authentication features and thus reduce the authentication latency. Based on two-layer DNNs, this scheme uses the top-level DNN to choose the authentication mode $x_1^{(k)}$ and the bottom-level DNN to select the test threshold $x_2^{(k)}$ with the chosen authentication mode.

Similar to Algorithm 1, the scheme formulates the state $\boldsymbol{s}^{(k)}$ based on message priority $\rho^{(k)}$, the number of the accepted messages by the authentication, $N_{\mathrm{P}}^{(k-1)}$, and the number of the falsely accepted messages, $N_{\mathrm{F}}^{(k-1)}$, in the previous T time slots, which is the input to the top-level DNN. As shown in Fig. 3, the top-level DNN with weights $\boldsymbol{\theta}$ includes two FC layers and outputs the top Q-values denoted by $\boldsymbol{Q}^*\left(\boldsymbol{s}^{(k)}, \cdot; \boldsymbol{\theta}\right)$. The two FC layers consist of n_1 and n_2 rectified linear units as the activation function. The authentication mode, $x_1^{(k)}$, is chosen based on the top Q-values according to ϵ-greedy. With $x_1^{(k)}$ and $\boldsymbol{s}^{(k)}$ as the input, the bottom-level DNN with weights $\boldsymbol{\varphi}$ has two FC layers that contain m_1 and m_2 rectified linear units, respectively, and outputs the bottom Q-values denoted by $\boldsymbol{Q}'\left(\boldsymbol{s}^{(k)}, x_1^{(k)}, \cdot; \boldsymbol{\varphi}\right)$. The test threshold $x_2^{(k)}$ is chosen based on $\boldsymbol{Q}'\left(\boldsymbol{s}^{(k)}, x_1^{(k)}, \cdot; \boldsymbol{\varphi}\right)$ with ϵ-greedy.

The test statistic Δ given by (3) is compared with the chosen test threshold $x_2^{(k)}$. If $\Delta \leq x_2^{(k)}$, the monitor updates the voltage records $\bar{\nu}_d$, $\tilde{\nu}_d$, $\bar{\mu}_d$ and $\tilde{\mu}_d$ similar to Algorithm 1, and sends an active error flag otherwise. The monitor measures $N_{\mathrm{F}}^{(k)}$ and uses (4) to calculate the utility $u^{(k)}$.

The weights of the DNNs are updated based on the authentication experience $\{\boldsymbol{s}^{(k)}, \boldsymbol{x}^{(k)}, u^{(k)}\}$, which is then stored in the memory pool $\mathcal{D}$, i.e.,

$$\mathcal{D} \leftarrow \mathcal{D} \cup \left\{\boldsymbol{s}^{(k)}, \boldsymbol{x}^{(k)}, u^{(k)}\right\}. \tag{9}$$

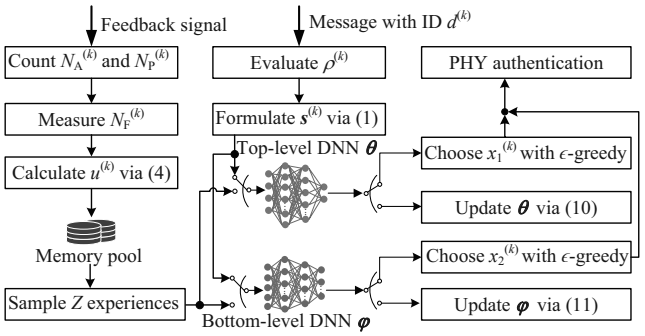


Fig. 3. Illustration of the Deep RL based CAN bus authentication.

By applying the experience replay technique, the scheme randomly samples Z experiences from $\mathcal{D}$ to update the top-level DNN weights $\boldsymbol{\theta}$. According to the Adam method [39], the top-level DNN weights are updated by

$$\boldsymbol{\theta} = \arg\min_{\boldsymbol{\theta}'} \mathbb{E}\left[\left(u^{(k-1)} + \beta \max_{x_1' \in \{0,1\}} \boldsymbol{Q}^*\left(\boldsymbol{s}^{(k)}, x_1'; \boldsymbol{\theta}'\right) - \boldsymbol{Q}^*\left(\boldsymbol{s}^{(k-1)}, x_1^{(k-1)}; \boldsymbol{\theta}'\right)\right)^2\right]. \tag{10}$$

The weights of the bottom-level DNN are updated with the Z authentication experiences using Adam,

$$\boldsymbol{\varphi} = \arg\min_{\boldsymbol{\varphi}'} \mathbb{E}\left[\left(u^{(k-1)} - \boldsymbol{Q}'\left(\boldsymbol{s}^{(k-1)}, \boldsymbol{x}^{(k-1)}; \boldsymbol{\varphi}'\right) + \beta \max_{x_2' \in \left\{\frac{i}{\Gamma}\right\}_{i \in \{1,2,\cdots,\Gamma\}}} \boldsymbol{Q}'\left(\boldsymbol{s}^{(k)}, x_1^{(k)}, x_2'; \boldsymbol{\varphi}'\right)\right)^2\right]. \tag{11}$$

VII. Performance Analyses

In this section, we investigate the computational complexity of our proposed authentication schemes and the performance bound including the authentication accuracy. More specifically, this scheme uses h_{T} multiplications to update the weights of the top-level DNN, which depends on the n_1 rectified linear units in the first FC layer, the n_2 rectified linear units in the second FC layer, the Z sampled authentication experiences and the two authentication modes. By [40], we have

$$h_{\mathrm{T}} = 12Zn_1 + 3Zn_1n_2 + 11Zn_2 + 8Z. \tag{12}$$

Similarly, the bottom-level DNN needs h_{B} multiplications to update its weights, which relies on the m_1 rectified linear units in the first FC layer, the m_2 rectified linear units in the second FC layer, the size of the sample authentication experiences Z and the Γ feasible test thresholds given by

$$h_{\mathrm{B}} = 15Zm_1 + 3Zm_1m_2 + (4\Gamma + 3)Zm_2 + 4Z\Gamma. \tag{13}$$

According to [41], the number of the rectified linear units in the first FC layer of the top-level DNN in Algorithm 2

Algorithm 2: Deep RL based CAN bus authentication

Initialize β, J, W, T, Z, $N_{\mathrm{A}}^{(0)}$, $N_{\mathrm{P}}^{(0)}$, $N_{\mathrm{F}}^{(0)}$, $\{\bar{\tau}_i\}_{1\leq i\leq J}$, $\{\bar{\nu}_i\}_{1\leq i\leq J}$, $\{\tilde{\nu}_i\}_{1\leq i\leq J}$, $\{\bar{\mu}_i\}_{1\leq i\leq J}$, $\{\tilde{\mu}_i\}_{1\leq i\leq J}$, $\boldsymbol{\theta}$, $\boldsymbol{\varphi}$, and $\mathcal{D}=\varnothing$
for $k=1,2,3,...$ **do**
  Read the arbitration ID $d^{(k)}$
  $\rho^{(k)}=1-\frac{d^{(k)}}{J}$
  Extract W CANH voltages $\left\{\nu_i^{(k)}\right\}_{1\leq i\leq W}$
  Extract W CANL voltages $\left\{\mu_i^{(k)}\right\}_{1\leq i\leq W}$
  if Message with arbitration ID $d^{(k)}$ is periodic **then**
    Measure $\tau^{(k)}$
  end if
  Form $\boldsymbol{s}^{(k)}$ via (1)
  Input $\boldsymbol{s}^{(k)}$ to the top-level DNN
  Choose $x_1^{(k)}$ based on the top-level DNN outputs, $\boldsymbol{Q}^*\left(\boldsymbol{s}^{(k)},\cdot;\boldsymbol{\theta}\right)$, with ϵ-greedy
  Input $\boldsymbol{s}^{(k)}$ and $x_1^{(k)}$ to the bottom-level DNN
  Choose $x_2^{(k)}$ based on the bottom-level DNN outputs, $\boldsymbol{Q}'\left(\boldsymbol{s}^{(k)},x_1^{(k)},\cdot;\boldsymbol{\varphi}\right)$, with ϵ-greedy
  Calculate Δ via (3)
  if $\Delta\leq x_2^{(k)}$ **then**
    Accept the message
    Update the voltage records
  else
    Send an active error flag
  end if
  Count $N_{\mathrm{A}}^{(k)}$ and $N_{\mathrm{P}}^{(k)}$
  Measure $N_{\mathrm{F}}^{(k)}$
  Compute $u^{(k)}$ via (4)
  Save the authentication experience via (9)
  Sample Z experiences from $\mathcal{D}$
  Update $\boldsymbol{\theta}$ via (10)
  Update $\boldsymbol{\varphi}$ via (11)
end for

depends on the number of the learning samples k and the two authentication modes, i.e.,

$$n_1=2\sqrt{2k}. \tag{14}$$

The number of the rectified linear units in the second FC layer of the top-level DNN relies on the k learning samples and two authentication modes given by

$$n_2=\sqrt{2k}. \tag{15}$$

The number of the rectified linear units in the first FC layer of the bottom-level DNN in Algorithm 2 relies on the number of the learning samples k and the number of the feasible test thresholds Γ given by

$$m_1=\sqrt{k\Gamma}+2\sqrt{\frac{k}{\Gamma}}. \tag{16}$$

The number of the rectified linear units m_1 in the second FC layer of the bottom-level DNN is given by

$$m_2=\sqrt{k\Gamma}. \tag{17}$$

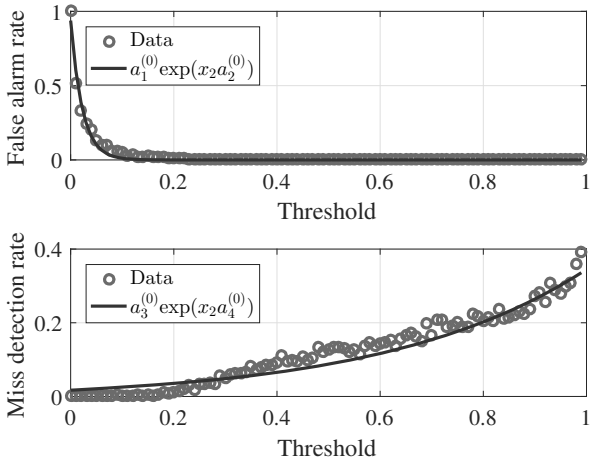


Fig. 4. Nonlinear regression of the false alarm rate and the miss detection rate based on the CAN bus that consists of a monitor, 18 legitimate ECUs and a compromised ECU as shown in Fig. 5.

Theorem 1. *The computational complexity of QLA, Viden, RLA and DRLA is given by* $O\left(k\Gamma\right)$, $O\left(k\right)$, $O\left(kD\Gamma\right)$, *and* $O\left(kZ\Gamma\right)$, *respectively.*

Proof. See Appendix A. □

Remark: The complexity of the four authentication schemes increases with the number of the learning samples k. Our proposed RLA also relies on the D simulated authentication experiences and the Γ feasible test thresholds. The computational complexity of DRLA increases with the number of the feasible test thresholds, Γ, and the size of the sample authentication experiences, Z, and uses the deep neural networks instead of the convolutional neural networks in [34], to reduce the sample complexity.

The CAN bus authentication against a spoofing attacker Eve can be formulated as an authentication game. In this game, the monitor chooses the authentication mode $x_1\in\{0,1\}$ and test threshold $x_2\in(0,1]$, while Eve selects the number of the spoofing messages in T time slots $y\in[0,N_{\mathrm{Y}}]$. For simplicity, the CAN bus is assumed to receive N_{L} legal messages with transmission priority ρ from the L ECUs in T time slots. Eve is assumed to use the arbitration IDs of a compromised ECU, observe the messages transmitted on the CAN bus, send y spoofing messages in T time slots, and aim to maximize its utility denoted by u_{S}. Decreasing with the utility of the monitor u and the spoofing cost c_{S}, the utility of Eve is modelled by

$$u_{\mathrm{S}}=-u-c_{\mathrm{S}}y. \tag{18}$$

We perform 100 experiments to obtain data of the false alarm rate and miss detection rate in Fig. 4. In each experiment, 18 legitimate ECUs and a compromised ECU exchange 1800 messages over a CAN bus at 500 Kbps baud rate as

shown in Fig. 5. The monitor, consisting of an ECU, an oscilloscope and a laptop, samples the voltage signals at a rate of 100 MS/s to obtain 360 CANH voltages and 360 CANL voltages for each message. The test statistic is calculated via (3) and compared with the given test threshold to obtain the false alarm rate and miss detection rate every 1800 messages.

The false alarm rate of the proposed authentication scheme $p_{\rm F}$ is modelled with two regression parameters, denoted by $0 < a_1^{(i)} \leq 1$ and $a_2^{(i)} < 0$, for two authentication modes. As shown in Fig. 4, $p_{\rm F}$ of authentication mode i is assumed to be an exponential function given by

$$p_{\rm F} = a_1^{(i)} \exp\left(x_2 a_2^{(i)}\right). \tag{19}$$

Similarly, the distribution of the miss detection rate $p_{\rm M}$ for authentication mode i is given by

$$p_{\rm M} = a_3^{(i)} \exp\left(x_2 a_4^{(i)}\right) \tag{20}$$

where $0 < a_3^{(i)} \leq \exp(-a_4^{(i)})$, and $a_4^{(i)} > 0$. For $\forall i \in \{0, 1\}$, let

$$\alpha_i = \frac{a_3^{(i)} a_4^{(i)} N_{\rm Y}}{a_1^{(i)} a_2^{(i)} N_{\rm L}} \tag{21}$$

$$\xi_i = \frac{a_2^{(i)}}{a_2^{(i)} - a_4^{(i)}}. \tag{22}$$

Theorem 2. *The performance bound of the RL based CAN bus authentication is given by*

$$p_{\rm F} \geq a_1^{(0)} \left(\alpha_0 - c_{\rm F}\alpha_0\right)^{\xi_0} \tag{23}$$

$$p_{\rm M} \geq a_3^{(0)} \left(\alpha_0 - c_{\rm F}\alpha_0\right)^{\xi_0 - 1} \tag{24}$$

$$u \leq \rho N_{\rm L} - \rho \left(N_{\rm L} a_1^{(0)} - \frac{N_{\rm Y} a_3^{(0)}}{\alpha_0}\right) \left(\alpha_0 - c_{\rm F}\alpha_0\right)^{\xi_0} \tag{25}$$

if

$$c_{\rm S} \leq -\frac{\rho a_3^{(0)}}{\alpha_0} \left(\alpha_0 - c_{\rm F}\alpha_0\right)^{\xi_0} \tag{26}$$

$$\alpha_0 < 1 - c_{\rm F} \leq \alpha_0 \exp\left(a_2^{(0)} - a_4^{(0)}\right) \tag{27}$$

$$\rho \left(N_{\rm L} a_1^{(0)} - \frac{N_{\rm Y} a_3^{(0)}}{\alpha_0}\right) \left(\alpha_0 - c_{\rm F}\alpha_0\right)^{\xi_0} \geq \rho \left(N_{\rm L} a_1^{(1)} - \frac{N_{\rm Y} a_3^{(1)}}{\alpha_1}\right) \left(\alpha_1 - c_{\rm F}\alpha_1\right)^{\xi_1} + c_{\rm T} \tag{28}$$

Proof. See Appendix B. □

Remark: In the case given by (26)-(28), the monitor chooses to use the signal voltages in the authentication (i.e., $x_1 = 0$), if the transmission priority of the message ρ is higher than a bound given by the $N_{\rm L}$ legal messages sent by the L ECUs and the maximum number of the received spoofing messages $N_{\rm Y}$ in T time slots. In this case, the authentication accuracy given by (23) and (24) relies on the number of the received legal messages $N_{\rm L}$ and the maximum number of the received spoofing messages $N_{\rm Y}$ in T time slots. The utility of the monitor given by (25) relies on the message priority ρ. According to [38], if (26)-(28) hold, the proposed RLA in Algorithm 1 can eventually converge to the optimal authentication policy given by

$$\boldsymbol{x}^* = \left[0, \frac{1}{a_2^{(0)} - a_4^{(0)}} \ln\left(\alpha_0 - c_{\rm F}\alpha_0\right)\right] \tag{29}$$

after sufficient interactions, and thus achieve the performance bound given by (23)-(25).

VIII. Experimental Results

Experiments were performed on a CAN bus at 500 Kbps baud rate that allows about 3906 messages per second. Eighteen legitimate ECUs, a monitor, and a compromised ECU are connected on the CAN bus, as shown in Fig. 5. The monitor is an ECU connected to a Dell E6430 laptop that has a 2.4 GHz dual-core processor and 16 GB of RAM to store the voltage data measured by an oscilloscope based on the extended ID field of the received messages at 100 MS/s sampling rate. Note that the monitor in a practical vehicle can be implemented with a sampling chip such as ADC08100CIMTC/NOPB and a Raspberry Pi 4 for lower costs.

The attacker Eve compromises the telematics ECU via WiFi and fabricates the messages with the legal arbitration IDs of the compromised ECU. Eve measures the CAN bus signal features, applies the greedy algorithm to choose a number $N_{\rm Y}$ from 0 to 800 and sends $N_{\rm Y}$ spoofing messages in 200 ms.

Typical message voltages of the ECUs on the CANH and the CANL are shown in Fig. 6. The CAN bus has 18 message types that consist of 14 periodic messages with arbitration IDs 0x000 $\sim$ 0x00D and 4 non-periodic messages with arbitration IDs 0x00E $\sim$ 0x011. The arrival intervals of the periodic messages are summarized in Table II. The legitimate ECUs and the compromised ECU send at most 1458 messages every second, i.e., 38% CAN bus load. The monitor estimates the false alarm rate and miss detection rate every 50 messages within 200 ms.

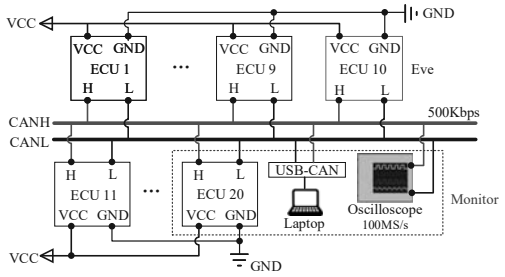
The learning parameters λ, β and ϵ are set as 0.1, 0.2 and 0.1, respectively, based on the experiments not shown here for accurate authentication. The first FC layer in the top-level DNN has 32 filters and the second FC layer has 16 filters. The two FC layers in the bottom-level DNN have 256 and 128 filters, respectively. The experience pool can store 3000 experiences, and the minibatch has 64 experiences.

TABLE II
Arrival Intervals of the Periodic Messages

Message ID	**Time interval** (ms)
0x000 $\sim$ 0x003	10
0x004 $\sim$ 0x006	20
0x007 $\sim$ 0x009	40
0x00A $\sim$ 0x00B	100
0x00C $\sim$ 0x00D	500

The authentication performance of our proposed RLA improves over time, as shown in Fig. 7. For example, RLA decreases the false alarm rate by 60.4% to 4.2%, and decreases the miss detection rate by 65.3% to 4.9% after 2000 time slots. Compared with the PHY authentication scheme named QLA in

(a) CAN bus snapshot



(b) Network topology

Fig. 5. Experimental setting of the CAN bus authentication with 20 ECUs, in which one compromised ECU sends spoofing messages.

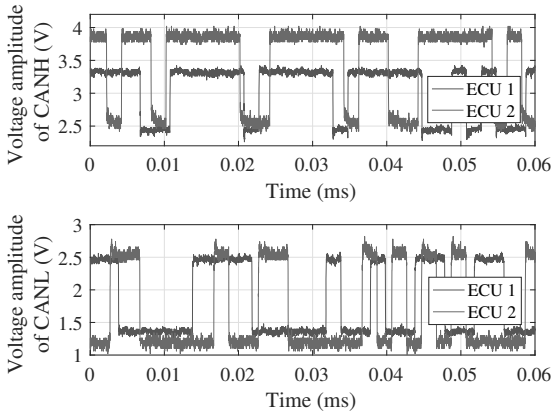
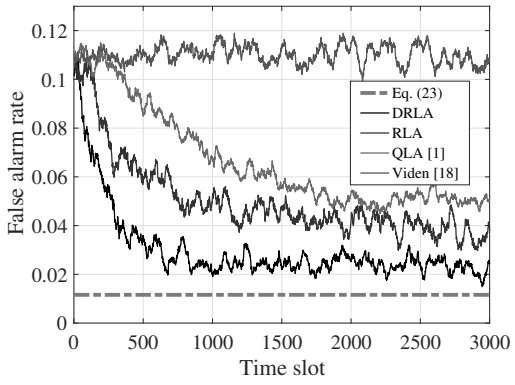
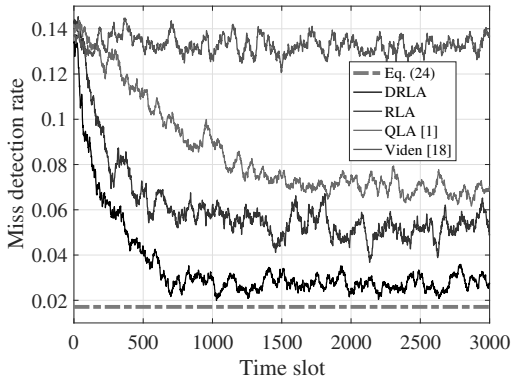


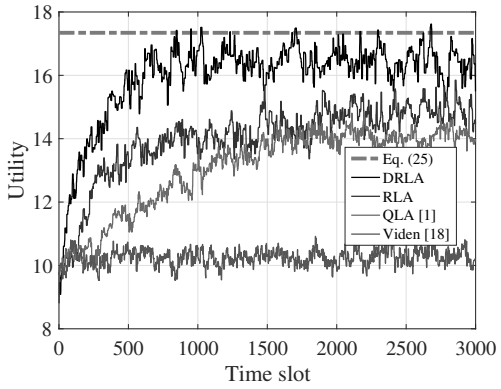
Fig. 6. Examples of the signal voltages of two ECUs on the CAN bus as shown in Fig. 5.



(a) False alarm rate



(b) Miss detection rate



(c) Utility

Fig. 7. Performance of the CAN bus authentication schemes with 18 legitimate ECUs in the CAN bus as shown in Fig. 5 against the attacker that sends $0 \sim 800$ spoofing messages in 200 ms.

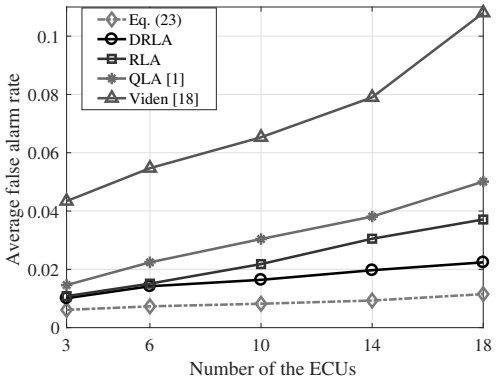
[1], RLA reduces the false alarm rate by 22.2%, decreases the miss detection rate by 34.3%, and increases the utility by 8.8% at the 2000-th time slot. The proposed RLA also improves the authentication performance as compared with Viden in [18]. For instance, RLA reduces the false alarm rate by 63.8% and the miss detection rate by 63.4%, and increases the utility by 49.0% after 2000 time slots compared with Viden. The reason is that RLA exploits both the arrival intervals and signal voltages to improve the authentication performance. In addition, RLA takes 96.5 μs to verify a message, including the physical feature measurement and authentication evaluation, which is shorter than the minimum message arrival time of 128 μs of the CAN bus in Fig. 5.

DRLA further improves the authentication performance of RLA, e.g., DRLA further decreases the false alarm rate by 40.0% to 2.5%, reduces the miss detection rate by 51.1% to 2.4%, increases the utility by 12.5% at time slot 2000, and decreases the convergence time by 50.0%. DRLA reaches the performance bound given by (23)-(25) after 1000 time slots, verifying the analysis results in Theorem 2. With a slightly higher computational complexity than QLA and Viden, DRLA can be implemented in a monitor equipped with a Dell E6430 laptop with a 2.4 GHz dual-core processor and 16 GB of RAM.

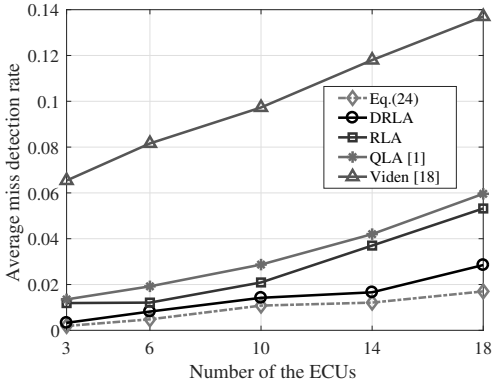
The authentication performance for the $3 \sim 18$ ECUs averaged over 500 time slots is presented in Fig. 8, showing that the proposed schemes work well in typical ECU settings. For example, RLA has the false alarm rate less than 3.7%

TABLE III
PERFORMANCE COMPARISON OF CAN BUS AUTHENTICATION SCHEMES

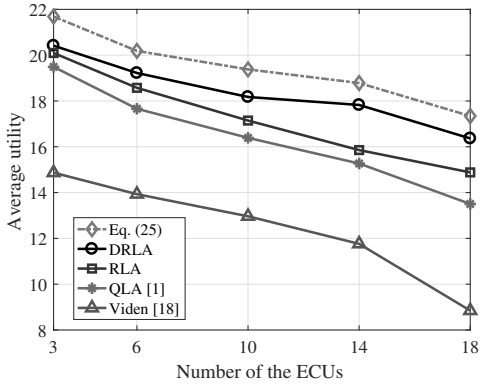
Performance / Method	Periodic messages		Non-periodic messages	
	False alarm rate (%)	**Miss detection rate** (%)	**False alarm rate** (%)	**Miss detection rate** (%)
DRLA	**0.8**	**1.0**	2.1	4.4
RLA	3.9	3.0	**4.1**	**5.2**
QLA [1]	5.1	6.8	4.9	6.7
Viden [18]	11.4	13.2	10.8	13.1
VoltageIDS [25]	**10.0**	**4.3**	9.1	4.5
IDA [20]	5.5	6.9	**44.2**	**55.6**



(a) Average false alarm rate

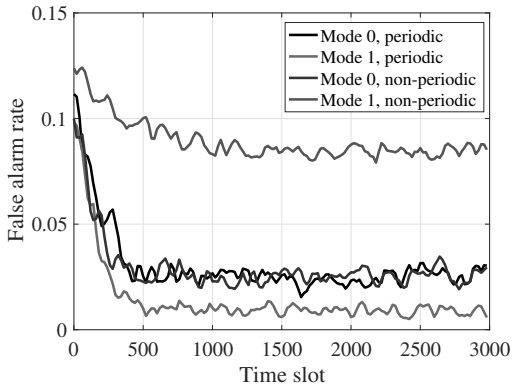


(b) Average miss detection rate

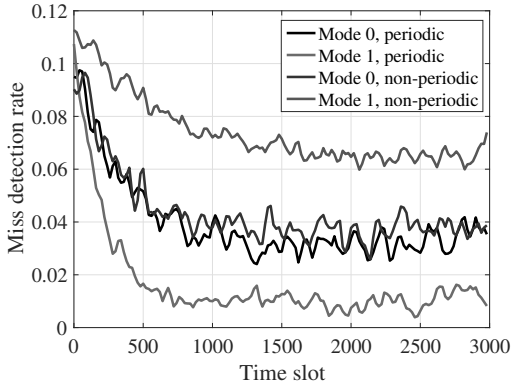


(c) Average utility

Fig. 8. Performance of the RL based CAN bus authentication schemes in the CAN bus as shown in Fig. 5 for the {3, 6, 10, 14, 18} legitimate ECUs averaged over 500 time slots against the attacker that sends $0 \sim 800$ spoofing messages in 200 ms.



(a) False alarm rate



(b) Miss detection rate

Fig. 9. Authentication accuracy for both periodic and non-periodic messages on the CAN bus equipped with 18 legitimate ECUs under two authentication modes.

and miss detection rate less than 4.59%, if the number of ECUs is less than 18. Compared with QLA and Viden, RLA decreases the false alarm rate by 35.9% and 66.1%, decreases the miss detection rate by 27.2% and 78.5%, and increases the utility by 4.6% and 32.2%, respectively, with 10 ECUs. The performance of DRLA further improves, with 24.8% reduction of the false alarm rate, 32.1% reduction of the miss detection rate, and 5.7% increase of the utility.

As shown in Table III, the accuracy averaged over 100 runs and 500 time slots of our proposed schemes exceeds existing CAN bus authentication schemes, including QLA [1], Viden [18], VoltageIDS [25] and the intrusion detection algorithm named IDA [20]. For example, the proposed DRLA has 8.0‰

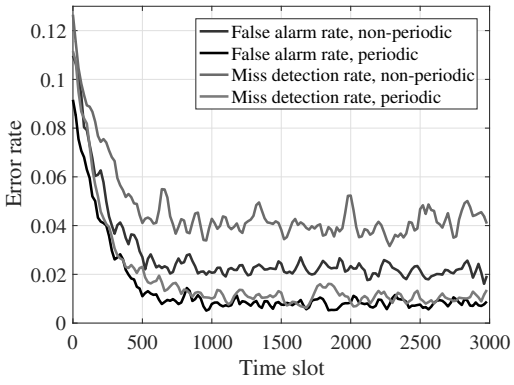


Fig. 10. Authentication accuracy of DRLA for both periodic and non-periodic messages on the CAN bus consisting of 18 legitimate ECUs.

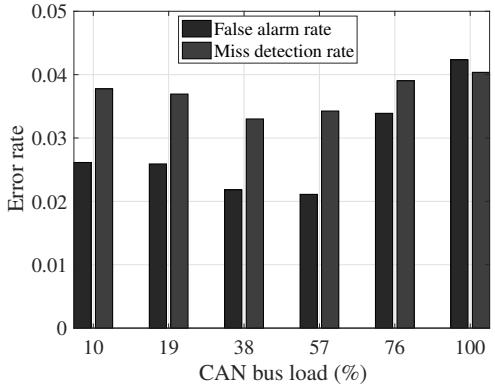


Fig. 11. Authentication accuracy of DRLA for given CAN bus load averaged over 100 runs each consisting of 500 time slots.

false alarm rate and 1.0% miss detection rate when dealing with periodic messages, which is 92.0% and 76.7% lower than VoltageIDS, respectively. The performance gain results from the authentication basis that consists of both the arrival intervals and signal voltages. Our proposed RLA reduces the false alarm rate by 90.7% and the miss detection rate by 90.6% compared with IDA. That is because IDA exploits the arrival intervals and thus fails to authenticate the non-periodic messages.

The authentication accuracy of the two modes for both periodic and non-periodic messages on the CAN bus with 18 legitimate ECUs in Fig. 9 shows that mode 0 is more effective for non-periodic messages, while mode 1 authenticates the periodic messages more accurately. For example, mode 1 reduces the false alarm rate by 74.3% to 6.0‰ and decreases the miss detection rate by 78.0% to 7.5‰ after 2000 time slots for periodic messages, as compared with mode 0. On the other hand, mode 0 has 2.7% false alarm rate and 3.8% miss detection rate at the 2000-th time slot when dealing with non-periodic messages, which is 68.9% and 41.6% higher than mode 1, respectively.

Our proposed DRLA is more accurate for periodic messages, as shown in Fig. 10. For instance, DRLA with periodic messages reduces the false alarm rate by 62.0% to 8.0‰ and decreases the miss detection rate by 75.2% to 1.1% after 1000 time slots, as compared with non-periodic messages.

As shown in Fig. 11, the authentication accuracy of our proposed DRLA averaged over 100 runs each consisting of 500 time slots is provided for the CAN bus load ranging from 10.0% to 100.0%. Both the false alarm rate and the miss detection rate first increase with the load and decrease afterwards. For example, the false alarm rate slightly decreases from 2.6% to 2.1% if the load changes from 10.0% to 57.0%, and increases to 4.2% if the load improves up to 100.0%. Besides, DRLA decreases the miss detection rate from 3.8% to 3.3% if the load increases from 10.0% to 38.0%, and increases the miss detection rate by 21.2% to 4.0% if the load changes from 38.0% to 100.0%.

IX. Conclusion and Future Work

In this paper, we have proposed an RL based CAN bus authentication scheme that exploits both the signal voltages and arrival intervals of messages to detect spoofing attacks. This scheme simulates hypothetical authentication experiences with the Dyna architecture to reduce the initial random exploration and uses a double estimator to avoid the over-estimation of the Q-values in the selection of the authentication mode and test threshold without the knowledge of the spoofing model. We also have designed a deep RL version to further improve the authentication performance and discussed its computational complexity and performance. Experiments on a CAN bus with 18 legitimate ECUs have been performed, with results showing that its authentication performance gains over QLA [1], Viden [18], VoltageIDS [25] and IDA [20]. For example, RLA reduces the false alarm rate by 60.4% to 4.2%, reduces the miss detection rate by 65.3% to 4.9%, and increases the utility by 52.7% compared with Viden. DRLA further reduces the false alarm rate by 34.5% and miss detection rate by 43.9%, and increases the utility by 15.2%.

Our scheme depends on a trusted monitor and has a lower authentication accuracy once the legitimate messages are falsely rejected. In the future, we plan to improve the authentication accuracy and robustness by incorporating multiple handshakes for the CAN-FD with a high baud rate and a large frame size. The monitor and the ECUs have to be protected with handshakes against the attackers that can compromise the in-vehicle nodes and provide better feedback signals to further reduce the false alarm rate. In addition, our scheme can combine the positive-slope with the negative-slope signal voltages in both the time and frequency domains to detect spoofing attacks more accurately and use transfer learning to exploit the similar authentication experiences to further accelerate the optimization speed.

Appendix A
Proof of Theorem 1

Proof. According to [42], the complexity of QLA in [1] and Viden in [18] is given by $O(k\Gamma)$ and $O(k)$. Compared with QLA, our proposed RLA updates the Q-values with D more times every time slot, and thus has complexity given by $O(kD\Gamma)$.

According to [43], by (12)-(17), the computational complexity of DRLA is given by

$$O\Big(Z\big(12n_1+3n_1n_2+11n_2+8+15m_1 + 3m_1m_2+(4\Gamma+3)\,m_2+4\Gamma\big)\Big) \tag{30}$$

$$= O\left(35Z\sqrt{2k}+18Zk+(4\Gamma+18)\,Z\sqrt{k\Gamma} + 30Z\sqrt{\frac{k}{\Gamma}}+3Zk\Gamma+4Z\Gamma+8Z\right) \tag{31}$$

$$= O\left(Z\sqrt{2k}+Zk+Z\sqrt{k\Gamma}+Z\sqrt{\frac{k}{\Gamma}}+Zk\Gamma\right) \tag{32}$$

$$= O\Big(Zk\Gamma\Big). \tag{33}$$

□

Appendix B
Proof of Theorem 2

Proof. By (4), (19) and (20), we have

$$u\left(\boldsymbol{x},y\right)=\rho N_{\mathrm{L}}-\rho N_{\mathrm{L}}a_1^{(x_1)}\exp\left(x_2a_2^{(x_1)}\right) + \rho y\left(1-c_{\mathrm{F}}\right)a_3^{(x_1)}\exp\left(x_2a_4^{(x_1)}\right)-c_{\mathrm{T}}x_1. \tag{34}$$

For given $\forall x_1\in\{0,1\}$, let

$$\tilde{x}_2=\frac{1}{a_2^{(x_1)}-a_4^{(x_1)}}\ln\left(\alpha_{x_1}-c_{\mathrm{F}}\alpha_{x_1}\right). \tag{35}$$

By (18) and (34), if (26) holds, we have

$$\frac{\partial u_{\mathrm{S}}\left([0,\tilde{x}_2],y\right)}{\partial y}=-\frac{\rho a_3^{(0)}}{\alpha_0}\left(\alpha_0-c_{\mathrm{F}}\alpha_0\right)^{\xi_0}-c_{\mathrm{S}}\geq 0,\forall y\in[0,N_{\mathrm{Y}}]. \tag{36}$$

Thus, we have

$$u_{\mathrm{S}}\left(\left[0,\frac{1}{a_2^{(0)}-a_4^{(0)}}\ln\left(\alpha_0-c_{\mathrm{F}}\alpha_0\right)\right],N_{\mathrm{Y}}\right) \geq u_{\mathrm{S}}\left(\left[0,\frac{1}{a_2^{(0)}-a_4^{(0)}}\ln\left(\alpha_0-c_{\mathrm{F}}\alpha_0\right)\right],y\right),\forall y\in[0,N_{\mathrm{Y}}]. \tag{37}$$

By (34), we have

$$\frac{\partial^2u\big(\boldsymbol{x},N_{\mathrm{Y}}\big)}{\partial x_2^2}=-\rho N_{\mathrm{L}}a_1^{(x_1)}\left(a_2^{(x_1)}\right)^2\exp\left(x_2a_2^{(x_1)}\right) + \rho N_{\mathrm{Y}}\left(1-c_{\mathrm{F}}\right)a_3^{(x_1)}\left(a_4^{(x_1)}\right)^2\exp\left(x_2a_4^{(x_1)}\right) \leq 0,\forall x_1\in\{0,1\},x_2\in(0,1]. \tag{38}$$

Thus, we have

$$\left.\frac{\partial u\big(\boldsymbol{x},N_{\mathrm{Y}}\big)}{\partial x_2}\right|_{x_2=\tilde{x}_2} = -\rho N_{\mathrm{L}}a_1^{(x_1)}a_2^{(x_1)}\left(\alpha_{x_1}-c_{\mathrm{F}}\alpha_{x_1}\right)^{\xi_{x_1}} + \frac{\rho N_{\mathrm{Y}}a_3^{(x_1)}a_4^{(x_1)}}{\alpha_{x_1}}\left(\alpha_{x_1}-c_{\mathrm{F}}\alpha_{x_1}\right)^{\xi_{x_1}} = 0,\forall x_1\in\{0,1\}. \tag{39}$$

Thus, if (26)-(28) hold, we have

$$u\left(\left[0,\frac{1}{a_2^{(0)}-a_4^{(0)}}\ln\left(\alpha_0-c_{\mathrm{F}}\alpha_0\right)\right],N_{\mathrm{Y}}\right) \geq u\big(\boldsymbol{x},N_{\mathrm{Y}}\big),\forall x_1\in\{0,1\},x_2\in(0,1]. \tag{40}$$

Thus, by (37) and (40), we have a Nash equilibrium of the authentication game given by

$$\left(\boldsymbol{x}^*,y^*\right)=\left(\left[0,\frac{1}{a_2^{(0)}-a_4^{(0)}}\ln\left(\alpha_0-c_{\mathrm{F}}\alpha_0\right)\right],N_{\mathrm{Y}}\right). \tag{41}$$

According to [44], by (19), (20) and (34), we have the performance bound given by (23)-(25). □

References

[1] T. Xu, X. Lu, L. Xiao, Y. Tang, and H. Dai, "Voltage based authentication for controller area networks with reinforcement learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019.

[2] W. Zeng, M. Khalid, and S. Chowdhury, "In-vehicle networks outlook: Achievements and challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1552–1571, Apr. 2016.

[3] Y. Zhang, M. Chen, N. Guizani, D. Wu, *et al.*, "SOVCAN: Safety-oriented vehicular controller area network," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 94–99, Aug. 2017.

[4] H. Mun, K. Han, and D. H. Lee, "Ensuring safety and security in CAN-based automotive embedded systems: A combination of design optimization and secure communication," *IEEE Trans. Veh. Technol.*, vol. 69, no. 7, pp. 7078–7091, Jul. 2020.

[5] S. Checkoway, D. Mccoy, B. Kantor, *et al.*, "Comprehensive experimental analyses of automotive attack surfaces," in *Proc. USENIX Security Symp.*, pp. 77–92, San Francisco, CA, Aug. 2011.

[6] B. Groza and P. Murvay, "Security solutions for the controller area network: Bringing authentication to in-vehicle networks," *IEEE Veh. Technol. Mag.*, vol. 13, no. 1, pp. 40–47, Mar. 2018.

[7] X. Ying, S. Sagong, A. Clark, L. Bushnell, *et al.*, "Shape of the cloak: Formal analysis of clock skew-based intrusion detection system in controller area networks," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 9, pp. 2300–2314, Sept. 2019.

[8] M. Kneib and C. Huth, "Scission: Signal characteristic-based sender identification and intrusion detection in automotive networks," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*, pp. 787–800, Toronto, Canada, Oct. 2018.

[9] "Vehicle CAN Database." https://www.vboxautomotive.co.uk/index.php/en/customer-area/vehicle-can-database.

[10] S. Woo, J. Jo, and L. Dong, "A practical wireless attack on the connected car and security protocol for in-vehicle CAN," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 993–1006, Apr. 2015.

[11] C. Miller and C. Valasek, "Remote exploitation of an unaltered passenger vehicle," *Black Hat USA*, Las Vegas, NV, Aug. 2015.

[12] J. Liu, S. Zhang, W. Sun, and Y. Shi, "In-vehicle network attacks and countermeasures: Challenges and future directions," *IEEE Network*, vol. 31, no. 5, pp. 50–58, Sept. 2017.

[13] W. Si, D. Starobinski, and M. Laifenfeld, "A robust load balancing and routing protocol for intra-car hybrid wired/wireless networks," *IEEE Trans. Mobile Comput.*, vol. 18, no. 2, pp. 250–263, Feb. 2019.

[14] A. Ferdowsi, S. Ali, W. Saad, and N. B. Mandayam, "Cyber-physical security and safety of autonomous connected vehicles: Optimal control meets multi-armed bandit learning," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7228–7244, Oct. 2019.

[15] B. Groza, S. Murvay, A. Herrewege, and I. Verbauwhede, "LiBrA-CAN: Lightweight broadcast authentication for controller area networks," *ACM Trans. Embedded Comput. Syst.*, vol. 16, no. 3, pp. 1–28, Apr. 2017.

[16] C. Lin, Q. Zhu, C. Phung, and A. Vincentelli, "Security-aware mapping for CAN-based real-time distributed automotive systems," in *Proc. IEEE/ACM Int. Conf. Comput. Aided Design (ICCAD)*, pp. 115–121, San Jose, CA, Nov. 2013.

[17] Z. Lu, Q. Wang, X. Chen, *et al.*, "LEAP: A lightweight encryption and authentication protocol for in-vehicle communications," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, pp. 1158–1164, Auckland, New Zealand, Oct. 2019.

[18] K. Cho and K. Shin, "Viden: Attacker identification on in-vehicle networks," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*, pp. 1109–1123, Dallas, TX, Oct. 2017.

[19] M. A. Jan, F. Khan, M. Alam, and M. Usman, "A payload-based mutual authentication scheme for Internet of Things," *Future Generation Computer Systems*, vol. 92, pp. 1028–1039, Sept. 2019.

[20] H. M. Song, H. R. Kim, and H. K. Kim, "Intrusion detection system based on the analysis of time intervals of CAN messages for in-vehicle network," in *Proc. IEEE Int. conf. information netw. (ICOIN)*, pp. 63–68, Kota Kinabalu, Malaysia, Jan. 2016.

[21] K. Cho and K. Shin, "Fingerprinting electronic control units for vehicle intrusion detection," in *Proc. USENIX Security Symp.*, pp. 911–927, Austin, TX, May 2016.

[22] B. Groza and P.-S. Murvay, "Efficient intrusion detection with Bloom filtering in controller area networks," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 4, pp. 1037–1051, Apr. 2019.

[23] K. Zhu, Z. Chen, Y. Peng, and L. Zhang, "Mobile edge assisted literal multi-dimensional anomaly detection of in-vehicle network using LSTM," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4275–4284, May 2019.

[24] P. Murvay and B. Groza, "Source identification using signal characteristics in controller area networks," *IEEE Signal Process. Lett.*, vol. 21, no. 4, pp. 395–399, Apr. 2014.

[25] W. Choi, K. Joo, H. Jo, M. Park, *et al.*, "VoltageIDS: Low-level communication characteristics for automotive intrusion detection system," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 8, pp. 2114–2129, Mar. 2018.

[26] K. Cho and K. Shin, "Error handling of in-vehicle networks makes them vulnerable," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*, pp. 1044–1055, Hofburg Palace, Austria, Oct. 2016.

[27] H. Olufowobi, C. Young, J. Zambreno, and G. Bloom, "SAIDuCANT: Specification-based automotive intrusion detection using controller area network (CAN) timing," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1484–1494, Feb. 2020.

[28] W. Choi, H. Jo, S. Woo, *et al.*, "Identifying ECUs using inimitable characteristics of signals in controller area networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 4757–4770, Jun. 2018.

[29] K. Lee and H. Byun, "A new face authentication system for memory-constrained devices," *IEEE Trans. Consum. Electron.*, vol. 49, no. 4, pp. 1214–1222, Nov. 2003.

[30] K. Maeno, Q. Sun, S.-F. Chang, and M. Suto, "New semi-fragile image authentication watermarking techniques using random bias and nonuniform quantization," *IEEE Trans. Multimedia*, vol. 8, no. 1, pp. 32–45, Feb. 2006.

[31] T. Kulkarni, K. Narasimhan, A. Saeedi, and J. Tenenbaum, "Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation," in *Proc. Conf. Advances Neural Inf. Processing Systems (NIPS)*, pp. 3675–3683, Barcelona, Spain, Dec. 2016.

[32] Y. Lei, Y. Yuan, and J. Zhao, "Model-based detection and monitoring of the intermittent connections for CAN networks," *IEEE Trans. Ind. Electron.*, vol. 61, no. 6, pp. 2912–2921, Jun. 2014.

[33] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10037–10047, Dec. 2016.

[34] X. Lu, L. Xiao, T. Xu, Y. Zhao, *et al.*, "Reinforcement learning based PHY authentication for VANETs," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3068–3079, Mar. 2020.

[35] S. Woo, H. J. Jo, I. S. Kim, and D. H. Lee, "A practical security architecture for in-vehicle CAN-FD," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 8, pp. 2248–2261, Aug. 2016.

[36] H. J. Jo, J. H. Kim, H.-Y. Choi, *et al.*, "MAuth-CAN: masquerade-attack-proof authentication for in-vehicle networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2204–2218, Feb. 2020.

[37] A. Humayed and B. Luo, "Using ID-hopping to defend against targeted DoS on CAN," in *Proc. ACM Int. Safe Control of Connected and Autonomous Vehicles*, pp. 19–26, Pittsburgh, PA, Apr. 2017.

[38] R. Sutton and A. Barto, *Reinforcement learning: An introduction*. MIT press, Cambridge, MA, 2018.

[39] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learning Representations (ICLR)*, San Diego, CA, May 2015.

[40] G. Ou and Y. Murphey, "Multi-class pattern classification using neural networks," *Pattern Recognition*, vol. 40, no. 1, pp. 4–18, Jan. 2007.

[41] G. Huang, "Learning capability and storage capacity of two-hidden-layer feedforward networks," *IEEE Trans. Neural Netw.*, vol. 14, no. 2, pp. 274–281, Mar. 2003.

[42] C. Jin, Z. Allen-Zhu, S. Bubeck, and M. I. Jordan, "Is Q-learning provably efficient?," in *Proc. Conf. Advances Neural Information Process. Syst. (NIPS)*, pp. 4868–4878, Montreal, Canada, Dec. 2018.

[43] H. Zeng, R. Chen, C. Zhang, and V. Prasanna, "A framework for generating high throughput CNN implementations on FPGAs," in *Proc. ACM/SIGDA Int. Symp. Field-Programmable Gate Arrays (FPGA)*, pp. 117–126, Monterey, CA, Feb. 2018.

[44] E. Even-Dar, A. Kesselman, and Y. Mansour, "Convergence time to Nash equilibrium in load balancing," *ACM Trans. Algorithms*, vol. 3, no. 3, pp. 1–21, Aug. 2007.