

Reinforcement Learning Enabled Vehicular Optimal Platoon Control under Uncertainties and Cyberattacks

Ziming Wang^{1,✉}, Karl H. Johansson², Apostolos I. Rikos¹

Cite this article: <https://doi.org/10.26599/JICV.2026.9210101>

ABSTRACT: This paper proposes a reinforcement learning (RL) based optimal platoon control framework for autonomous vehicles (AVs) subject to uncertainties and cyberattacks. Within the optimized backstepping design, an actor-critic-identifier architecture is employed to improve platoon control performance. A cyberattack compensation mechanism and a fuzzy logic based composite observer are integrated to reconstruct reliable vehicle information and estimate system uncertainties under compromised measurements, including the internal vehicular system states, uncertainties, and cyberattacks. In addition, a second-order tracking differentiator is introduced to smooth transient responses and substantially reduce the computational burden associated with differentiating the virtual controllers generated during the backstepping procedure. Based on the Lyapunov function based theory, the stability of the entire vehicle platoon is established, and all relevant errors in the multi-vehicle system are proven to be semiglobally uniformly ultimately bounded (SGUUB). Finally, numerical simulations of a multi-vehicle platooning system validate the effectiveness of the proposed framework and demonstrate its robustness against the simultaneous presence of uncertainties and cyberattacks. Moreover, a comparative study with a non-optimal control algorithm is conducted, highlighting the superior tracking performance and transient response of the proposed RL-enabled optimal control framework.

KEYWORDS: Optimal platoon control; adaptive control; cyberattacks; reinforcement learning (RL); autonomous vehicles (AVs).

1 Introduction

With the rapid growth of road traffic and increasing transportation demands, challenges such as traffic congestion, traffic safety, energy consumption, and environmental pollution have become increasingly prominent (Besselink and Johansson, 2017; Fei et al., 2024; Ioannou and Chien, 1993). Intelligent transportation systems (ITSs) have therefore attracted extensive attention because of their potential to improve traffic efficiency (Saccone, 2024), enhance road safety (Narri et al., 2025; Wang et al., 2025), and reduce adverse environmental impacts (Yang et al., 2025). As an important application of autonomous vehicles (AVs), vehicular platoon control coordinates a group of vehicles to travel at a prescribed velocity while maintaining desired inter-vehicle distances. The earliest large-scale implementation of this concept was conducted in the California PATH program in the mid-1980s (Shladover et al., 1991). Compared with conventional human-driven traffic, automated platooning enables reduced and more stable inter-vehicle spacing between consecutive vehicles, thereby increasing road capacity and improving traffic flow (Wang et al., 2026b; Wang et al., 2026c). In addition, close formation can reduce aerodynamic drag and fuel consumption, especially for highway transportation. Reliable platoon control also contributes to operational safety by enabling coordinated acceleration, braking, and disturbance rejection among vehicles. Consequently, the development of accurate and robust platoon control strategies has become a fundamental research topic in ITSs.

1.1. Related Works

Vehicular platoon control for nonlinear vehicle dynamics under uncertain operating conditions has been extensively investigated through adaptive control, neural network (NN)-based approximation,

and observer-based estimation techniques. (Besselink and Johansson, 2017) analyzed string stability for vehicle platoons subject to external disturbances, while (Liu et al., 2021a) developed cooperative extended-state observers to estimate unavailable inter-vehicle information under small time-headway policies. Adaptive NN based control was further employed by (Liu et al., 2022) to address unknown nonlinear dynamics and achieve distributed compound tracking. These studies have established important foundations for robust vehicle platoon coordination. However, existing approaches mainly focus on stability analysis and tracking control under uncertainties or disturbances, whereas resilient optimal control of high-order nonlinear platoons subject to corrupted measurements remains insufficiently explored.

However, the increasing dependence on inter-vehicle communication and onboard sensing makes platoons more susceptible to malicious cyberattacks. (Ren et al., 2023) investigated distributed consensus estimation for nonlinear systems under deception attacks, and (Chen et al., 2023) developed observer-based adaptive fuzzy consensus control against compromised information. For AVs, (Xie et al., 2022) introduced proportional-integral observers to compensate for replay attacks in distributed platooning. Although these approaches enhance cyber resilience, cyberattacks, unknown nonlinear dynamics, external disturbances, and unavailable vehicle states are frequently handled through separate mechanisms. Simultaneously estimating and compensating for these factors within a unified high-order platoon control framework remains challenging.

Reinforcement learning (RL) and adaptive dynamic programming have been increasingly applied to nonlinear optimal control because they can approximate solutions to the Hamilton-Jacobi-Bellman (HJB)

¹ Thrust of Artificial Intelligence, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511453, China. ² Division of Decision and Control Systems, KTH Royal Institute of Technology, Stockholm SE-100 44, Sweden.

✉ Corresponding author. E-mail: zwang216@connect.hkust-gz.edu.cn

Received: August 3, 2026; Revised: August 27, 2026; Accepted: September 14, 2026

© The Author(s) 2026. This is an open access article under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0, <http://creativecommons.org/licenses/by/4.0/>).

equation online. (Wen et al., 2020) incorporated actor-critic learning into optimized backstepping control for nonlinear strict-feedback systems, while (Song et al., 2024) studied RL-based optimal control under state constraints. Tracking differentiators have also been adopted to avoid repeated analytical differentiation of virtual controllers and reduce the complexity of recursive control design in (Zhao et al., 2022). Nevertheless, limited research has integrated attack compensation, composite state and disturbance observation, tracking differentiation, and actor-critic-identifier learning into a unified Lyapunov based framework for high-order vehicle platoons. This limitation motivates further investigation of resilient and RL-enabled optimal platoon control under uncertainties and cyberattacks.

1.2. Motivation

Despite significant advances in resilient vehicle platooning, cyberattack compensation, and learning-based optimal control, several fundamental challenges remain. Existing secure platoon control methods mainly focus on attack detection, state reconstruction, and closed-loop stabilization under compromised measurements, with greater emphasis on guaranteeing bounded tracking and platoon safety. For example, (He et al., 2021) developed a resilient estimation and attack-detection architecture for vehicle platoons under corrupted measurements, while (Mousavinejad and Vlacic, 2022) investigated secure platooning under multiple cyberattack scenarios using set-membership state estimation. However, less attention has been paid to the simultaneous treatment of time-varying cyberattacks, unknown nonlinear vehicle dynamics, uncertainties, and unavailable states within a unified high-order vehicular platoon model. This problem becomes more challenging because attack-induced estimation errors may enter the distributed tracking variables and propagate through the recursive backstepping procedure, thereby affecting both the virtual controllers and the actual control input.

Therefore, it is necessary to develop an RL-based optimal distributed platoon control framework that can simultaneously compensate for corrupted measurements, reconstruct unavailable vehicle states, and estimate unknown nonlinear dynamics and external disturbances. Moreover, although optimized backstepping methods based on actor-critic-identifier learning have been studied for nonlinear systems (Song et al., 2024; Wen et al., 2020), limited attention has been devoted to optimizing each virtual controller and the actual control input when cyberattacks and multiple uncertainties coexist in vehicle platoons. Since the virtual controllers depend on observer estimates, neighboring information, adaptive fuzzy parameters, and learned policies, their repeated analytical differentiation also results in substantial computational complexity. Hence, a unified design integrating projection-based cyberattack compensation, a fuzzy logic based composite observer, second-order tracking differentiators, and an actor-critic-identifier structure is required to achieve resilient state reconstruction, reduce recursive differentiation, optimize the complete control process, and guarantee bounded platoon tracking under uncertainties and cyberattacks. The technical modules described here are consistent with the controller architecture and semiglobally uniformly ultimately bounded (SGUUB) objective formulated in this paper.

1.3. Main Contributions

Motivated by the aforementioned limitations, this paper develops an RL-enabled optimal platoon control framework for high-order AVs subject

to uncertainties and cyberattacks. The main contributions are summarized as follows

1) A unified resilient estimation and compensation architecture is developed by combining a projection-based cyberattack compensation mechanism with a fuzzy logic based composite observer. The proposed architecture reconstructs unavailable vehicle states while simultaneously estimating uncertainties under corrupted position measurements.

2) An efficient optimized backstepping architecture is established by embedding second-order tracking differentiators and actor-critic-identifier RL structure into the recursive platoon control design. The proposed structure avoids repeated analytical differentiation of complex virtual controllers and enables online optimization of both virtual control policies and actual control inputs.

3) We develop a composite Lyapunov framework to analyze the complete closed-loop system, including tracking errors, observer errors, attack compensation errors, tracking differentiator errors, and neural-network approximation errors, proving the SGUUB stability of the proposed platoon system. Numerical simulations of a multi-vehicle platoon further demonstrate accurate position, velocity, and acceleration tracking, bounded learning parameters, and safe inter-vehicle spacing under uncertainties and cyberattacks.

Paper Organization. The remainder of this paper is organized as follows. Section 2 formulates the vehicle platoon control problem by introducing the third-order vehicle dynamics, the cyberattack and uncertainty models, the inter-vehicle communication topology. Section 3 presents the proposed optimal control framework and the closed loop stability established through a composite Lyapunov analysis. Section 4 provides numerical simulations to evaluate the tracking performance, learning behavior, robustness, and inter-vehicle safety of the proposed method, together with comparisons against existing control approaches. Finally, Section 5 concludes the paper and discusses potential directions for future research.

2 Problem Formulation

In this section, we introduce the third-order vehicular dynamical system model, along with the modeling of uncertainties and cyberattacks. Then we present the graph theory to form the vehicular communication network. Finally, we describe the control architecture and provide a detailed problem statement.

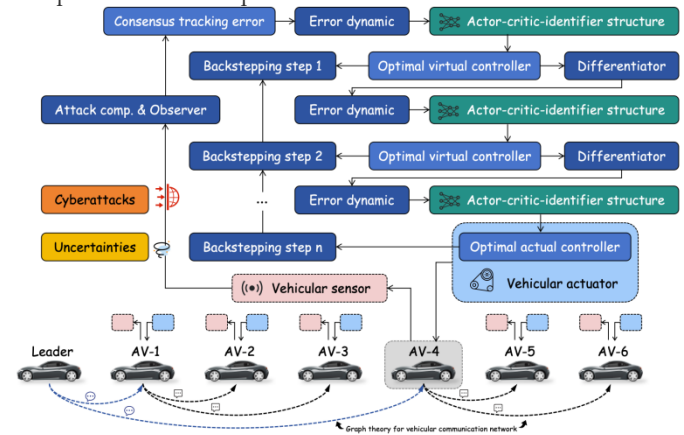


Fig. 1. The proposed RL based optimal platoon control architecture block diagram.

2.1. Vehicular Dynamical Model with Uncertainties and Cyberattacks

In this paper, we define the index of each vehicle as AV- i , $i = 1, \dots, n$. To address malicious attack of vehicle sensor outputs, a time-varying cyberattack model is considered. We define $\psi_i(t) = \varphi_i(t)x_i(t)$ as the cyberattack injection governed by an unknown time-varying coefficient $\varphi_i(t)$. Then the third-order vehicular dynamical system model of AVs is designed as

$$\dot{x}_i(t) = v_i(t), \quad (1)$$

$$\dot{v}_i(t) = a_i(t), \quad (2)$$

$$\dot{a}_i(t) = \frac{1}{\lambda_i} u_i(t) + f_i^v(v_i) + f_i^a(v_i, a_i) + d_i(t), \quad (3)$$

where x_i , v_i and a_i denote the position, velocity and acceleration of AVs. The traction force of vehicular engine (control input) for each vehicle is represented as u_i with the inertial lag parameter λ_i . And f_i^v , f_i^a denote the nonlinear functions. The uncertainties in this system are defined as d_i . Notably, the received attacked state is $(1 + \varphi_i)x_i$. In this paper, the time-varying cyberattack is considered in the inter-vehicle communication channel, where the transmitted position information of each follower vehicle is maliciously manipulated by an unknown multiplicative attack signal, resulting in corrupted state information available to the controller.

2.2. Vehicular Communication Network

The inter-vehicle communication topology of the platoon is described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, where $\mathcal{V} = \{1, 2, \dots, n\}$ denotes the node set consisting of n follower vehicles. The set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents the directed communication links and $\mathcal{A} = [c_{i,j}]$ is the adjacency matrix of $\mathcal{G}$, where $c_{i,j} = 1$ indicates information transmission from vehicle j to vehicle i , whereas $c_{i,j} = 0$ indicates no communication interactions between the two vehicles. Then we define the graph Laplacian matrix $\mathcal{L} = \mathcal{D} - \mathcal{A}$, where degree matrix $\mathcal{D} = \text{diag}(\delta_1, \delta_2, \dots, \delta_n)$ with $\delta_i = \sum_{j=1, j \neq i}^n c_{i,j}$. Direct access to the leader information is characterized by the pinning coefficient $\mathcal{B} = \text{diag}(b_1, b_2, \dots, b_n)$, where $b_i > 0$ if the leader can transfer information to follower i , and $b_i = 0$ otherwise. Notably, the augmented directed graph containing the leader and all follower vehicles has a directed spanning tree rooted at the leader. Based on Fig. 1, a directed vehicular communication network is considered in this paper, where each follower vehicle exchanges state information with its neighboring vehicle and AV-1 and AV-4 have access to the leader information.

2.3. Control Architecture

The platoon consists of one leader and n follower AVs connected through a directed communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$. Each follower constructs the distributed tracking errors using local state estimates, information from neighboring vehicles, and the available leader trajectory x_r . Corrupted position measurements are first processed by the projection based compensation mechanism and then supplied to the fuzzy logic based composite observer $(\hat{x}_i, \hat{v}_i, \hat{a}_i)$ for joint estimation of vehicle states (x_i, v_i, a_i) , unknown uncertainties (f_i^v, f_i^a) , and external uncertainties d_i and cyberattacks ψ_i . Based on these estimates, the backstepping procedure, tracking differentiators, and actor-critic-identifier structure generate the optimal virtual controllers and actual control input. Direct information exchange may occur between neighboring vehicles according to the prescribed communication topology and no central coordinator is required during the control process. The control architecture is therefore a graph-based distributed leader-following vehicular optimal platoon control framework.

2.4. Problem Statement

Consider a high-order multi-vehicle dynamical system including n AVs. The dynamics of each vehicle are governed by the high-order system, which is subject to unknown uncertainties (f_i^v, f_i^a) , and external uncertainties d_i and cyberattacks ψ_i . Notably, this paper only addresses sensor data tampering among cyberattacks, excluding DoS, replay, communication delay, and actuator attacks. The control architecture diagram is shown in Fig. 1. Under these conditions, the control objectives of this paper are formulated as follows:

1) Design a distributed RL-enabled optimal platoon controller such that all followers accurately track the prescribed leader position, velocity, and acceleration trajectories while maintaining the desired platoon spacing.

2) Validate the robustness and effectiveness of the proposed optimal platoon control framework through simulation experiments.

3 Optimal Control Design and Stability Analysis

Following the graph-based distributed optimal platoon control architecture described in Section 2, we now introduce the optimal controller design based on the actor-critic-identifier structure, a projection-based compensation mechanism, a fuzzy logic based composite observer and a tracking differentiator. Then we outline the stability analysis based on the Lyapunov-based stability theory.

3.1. Optimal Control Design

In this paper, to deal with the cyberattacks in optimal control process, we introduce the following projection-based compensation mechanism.

Projection-based compensation mechanism. In the vehicular dynamical model, the cyberattack injection $\psi_i(t) = \varphi_i(t)x_i(t)$ is established by an unknown coefficient $\varphi_i(t)$. Then an inverse compensation factor, defined as $\tau_i(t) = \frac{1}{1+\varphi_i(t)}$, is introduced to attenuate the influence of this unknown multiplicative distortion. We assume that τ_i satisfies $0 < \underline{\tau}_i \leq \tau_i \leq \bar{\tau}_i$. Provided that $\varphi_i(t) \neq -1$ and that both the attack coefficient and its time derivative remain bounded, $\tau_i(t)$ and its derivative are also bounded. Notably, based on (Zhao and Yang, 2024), the condition $\varphi_i(t) \neq -1$ ensures that the measured position information is not completely eliminated, and the bounded derivative assumption of $\varphi_i(t)$ reflects the practical limitation that a physical attacker cannot generate infinitely fast variations in the injected signal due to limited computational resources, communication bandwidth, and attack implementation capability. Since the compensation factor is unavailable in practice, an adaptive estimator $\hat{\tau}_i$ is constructed, with the corresponding estimation error defined as $\bar{\tau}_i = \tau_i - \hat{\tau}_i$. Therefore, the error between the compensated position and the actual position is determined by $\bar{\tau}_i(x_i + \varphi_i x_i)$. Hence, as long as $\bar{\tau}_i$ can be guaranteed to be bounded and gradually enter a small neighborhood, the impact of cyber attacks on the subsequent observer and controller can be confined within an adjustable range. To avoid parameter drift under persistent cyberattacks, we design a continuously differentiable convex boundary function $\kappa(\hat{\tau}_i)$ as

$$\kappa(\hat{\tau}_i) = \frac{(m_i + 1)\hat{\tau}_i^2 - \hat{\tau}_i^{max 2}}{m_i \hat{\tau}_i^{max 2}}, \quad (4)$$

where $m_i > 0$ denotes the attack-tolerance parameter, which controls the tolerance of the projection boundary and $\hat{\tau}_i^{max} > 0$ denotes the prescribed norm bound of the estimate, which gives the maximum

allowable range of the estimate. When the estimate lies in the interior region, or when the update direction points inward, the projection operator $\mathcal{E}_i[\xi_i] = \xi_i - \frac{\kappa\kappa^T}{\|\kappa\|^2}\xi_i$ preserves the original adaptive direction. When the estimate is near the boundary and the adaptive direction continues to point outward, the projection operator weakens its normal component, thereby preventing $\hat{e}_i$ from leaving the feasible region. The explicit form of ξ_i is determined through the Lyapunov stability analysis.

Fuzzy logic based composite observer. The unknown nonlinear dynamics $f_i^v(v_i)$ and $f_i^a(v_i, a_i)$ are represented by fuzzy logic systems with ideal constant fuzzy weights M_i^{v*} and M_i^{a*} , fuzzy basis vectors Π_i^v and Π_i^a , and bounded approximation errors ε_i^v and ε_i^a . Then with $\varpi_i = \hat{e}_i(x_i + \varphi_i x_i) - \hat{x}_i$, we design the fuzzy logic based composite observer as

$$\hat{x}_i(t) = \hat{v}_i + \mu_i^v \varpi_i, \quad (5)$$

$$\hat{v}_i(t) = \hat{a}_i + \mu_i^v \varpi_i, \quad (6)$$

$$\hat{a}_i(t) = \frac{1}{\lambda_i} u_i + M_i^{vT} \Pi_i^v + M_i^{aT} \Pi_i^a + \mu_i^a \varpi_i + \hat{d}_i, \quad (7)$$

$$\hat{d}_i = \mu_i^a \varpi_i, \quad (8)$$

where μ_i^v , μ_i^v , μ_i^a and μ_i^a are all designed observer gains. The proposed observer structure introduces external disturbances as augmented states and employs a fuzzy logic system to approximate the unknown internal nonlinear dynamics of the vehicle. Therefore, it enables the simultaneous estimation of system states, model uncertainties, and external disturbances within a unified framework. For compact analysis in Lyapunov function, with a unique symmetric positive definite matrix $P_{0,i}$, we define $\hat{X}_i = [\hat{x}_i, \hat{v}_i, \hat{a}_i, \hat{d}_i]^T$. The fuzzy logic based composite observer is designed to reconstruct reliable vehicle information from compensated measurements under uncertain and adversarial environments. By incorporating fuzzy approximation and disturbance estimation mechanisms, the observer provides accurate state feedback for the subsequent backstepping and actor-critic-identifier based optimal control design.

Second-order tracking differentiator. During the recursive backstepping control procedure, the derivative of the first virtual controller $\beta_{1,i}$ is required in the design of the second step, while the derivative of the second virtual controller $\beta_{2,i}$ subsequently appears in the construction of the actual control input. Through the backstepping method, virtual controllers $\beta_{1,i}$, $\beta_{2,i}$ and the actual controller u_i can be derived under different constraint conditions. Since controller design is not the primary focus of this work, the specific controller forms are not explicitly specified here. Since the virtual controllers generally involve observed states, fuzzy adaptive weights, neighboring information, and RL policies simultaneously, their direct analytical differentiation would introduce numerous higher-order derivatives and complicated coupling terms. To reduce the computational burden induced by recursive differentiation and improve the transient smoothness of the virtual control signals, a second-order tracking differentiator ($s = 1, 2$) is introduced as

$$\dot{\eta}_{s,i} = \rho_{s,i}, \quad (9)$$

$$\dot{\rho}_{s,i} = -2\epsilon_{s,i} r_{s,i} \rho_{s,i} - r_{s,i}^2 (\eta_{s,i} - \beta_{s,i}), \quad (10)$$

where $\eta_{s,i}$ denotes the smooth tracking signal of the virtual controller $\beta_{s,i}$, and $\rho_{s,i}$ represents the estimated derivative of $\eta_{s,i}$. The parameters $\epsilon_{s,i} > 0$ and $r_{s,i} > 0$ denote the damping coefficient and tracking bandwidth, respectively. Then we define the approximation errors of tracking differentiator as $e_{1,s,i} = \eta_{s,i} - \beta_{s,i}$ and $e_{2,s,i} = \rho_{s,i} -$

$\dot{\beta}_{s,i}$. For compact analysis in Lyapunov function, with a unique symmetric positive definite matrix $P_{k,i}$, we also define $E_{k,i} = [e_{1,s,i}, e_{2,s,i}]^T$.

Based on the compensated state estimates provided by the fuzzy logic based composite observer, a recursive error transformation is introduced to handle the distributed platoon tracking objective. Let ℓ_i denote the prescribed longitudinal offset of AV- i to the leader vehicle, with $\ell_0 = 0$ and $\ell_i = i \times \ell^*$. The distributed position error of AV- i is defined as

$$z_{1,i} = \sum_{j=1}^n c_{i,j} [(\hat{x}_i + \ell_i) - (\hat{x}_j + \ell_j)] + b_i [(\hat{x}_i + \ell_i) - x_r]. \quad (11)$$

The first term in Eq.(11) measures the deviation between AV- i and its neighboring vehicles, while the second term measures the deviation from the leader whenever the leader information is directly available to AV- i . Based on the backstepping method, the first virtual controller is denoted by $\beta_{1,i}$. The tracking differentiator introduced previously is employed to generate the smoothed signal $\eta_{1,i}$ and the derivative estimate $\rho_{1,i}$. The second backstepping error is then defined by

$$z_{2,i} = \hat{v}_i - \eta_{1,i}. \quad (12)$$

This error describes the difference between the estimated vehicle velocity and the filtered output of the position layer virtual controller. Similarly, the second virtual controller $\beta_{2,i}$ is processed by another tracking differentiator, whose outputs are denoted by $\eta_{2,i}$ and $\rho_{2,i}$. The acceleration layer error is defined as

$$z_{3,i} = \hat{a}_i - \eta_{2,i}. \quad (13)$$

The variable $z_{3,i}$ measures the mismatch between the estimated acceleration and the smoothed command generated by the velocity layer. Through this recursive construction, the original third order platoon control problem is transformed into the stabilization of $z_{1,i}$, $z_{2,i}$ and $z_{3,i}$. Then we define

$$z_i = [z_{1,i}, z_{2,i}, z_{3,i}]^T. \quad (14)$$

Actor-critic-identifier structure. The stabilizing controller derived above guarantees the fundamental robustness of the transformed platoon error system. The RL is introduced to improve the closed loop performance by optimizing the policy corrections added to the nominal virtual controllers and physical input. In this architecture, the actor generates the policy corrections implemented by the controller, the critic approximates the optimal long term performance, and the identifier estimates the system dynamics. Firstly, based on the backstepping virtual controller of each vehicle $\bar{u}_i$ and optimal virtual controller of each vehicle u_i^* we design the system performance index function as

$$J_i^*(z_i) = \min_{\bar{u}_i} \int_t^\infty [z_i^T Q_i z_i + \bar{u}_i^T R_i \bar{u}_i] d\chi, \quad (15)$$

where $Q_i = Q_i^T$ and $R_i = R_i^T$ are designed matrices. The matrix Q_i determines the relative penalties assigned to the position, velocity, and acceleration errors. The matrix R_i penalizes the magnitude of the RL policy correction. A larger Q_i places greater emphasis on tracking accuracy, while a larger R_i discourages aggressive control corrections. Computing the time derivative on both sides of Eq.(15) gives the HJB equation. Since this value function is generally unavailable for the nonlinear platoon system, the critic network is introduced to approximate it online. On a prescribed compact set, the optimal value function is represented as

$$J_i^*(z_i) = W_{c,i}^T Y_i(z_i) + \phi_{c,i}(z_i), \quad (16)$$

where $W_{C,i}^*$ is the ideal critic weight vector, $Y_i(z_i)$ denotes a continuously differentiable basis vector and $\phi_{C,i}(z_i)$ represents the bounded value function approximation error. The critic estimate is defined by $\hat{J}_i(z_i) = \hat{W}_{C,i}^T Y_i(z_i)$. The critic evaluates the expected future cost associated with the current tracking error and policy. Its output is used to approximate the value function gradient required by the optimal policy. Then based on the actor-critic structure, we have the actor NN weight and critic NN weight as

$$\hat{W}_{A,i} = -\Gamma_{A,i} [Y_{A1,i}(\hat{W}_{A,i} - \hat{W}_{C,i}) + Y_{A2,i}\hat{W}_{A,i}], \quad (17)$$

$$\hat{W}_{C,i} = -\Gamma_{C,i} \frac{\varrho_i}{(1 + \varrho_i^T \varrho_i)^2} \vartheta_i - \Gamma_{C,i} Y_{C,i} \hat{W}_{C,i}, \quad (18)$$

where matrices $\Gamma_{A,i} > 0$ and $\Gamma_{C,i} > 0$ determine the actor and critic learning rates. And $Y_{A1,i}$, $Y_{A2,i}$ and $Y_{C,i}$ are all basis vectors. ϱ_i denotes the regression vector and ϑ_i represents the Bellman error. In Eq.(17), the first term causes the actor weights to follow the policy information inferred by the critic. The second term provides regularization and prevents unbounded actor weight growth under residual approximation errors or insufficient excitation.

In the process of differentiating with respect to z_i , we define the transformed dynamics in a unified form as $F_i = [F_{1,i}, F_{2,i}, F_{3,i}]^T$, where $F_{1,i}$ contains the effects of the neighboring velocity information, the leader velocity, and the observer innovation signals. $F_{2,i}$ collects the estimated velocity channel nonlinearity, the observer correction term, and the derivative estimate supplied by the first tracking differentiator. $F_{3,i}$ contains the estimated acceleration channel nonlinearity, the disturbance estimate, the observer correction, and the derivative estimate of the second virtual controller. An identifier is then introduced to approximate the dynamics as $F_i = W_{I,i}^{*T} Y_{I,i} + \phi_{I,i}$, where $W_{I,i}^{*T}$ is the ideal identifier weight vector, $Y_{I,i}$ denotes a continuously differentiable basis vector and $\phi_{I,i}$ represents the identifier approximation residual. We design the identifier NN weight as

$$\hat{W}_{I,i} = \Gamma_{I,i} z_i^T - \Gamma_{I,i} Y_{I,i} \hat{W}_{C,i}. \quad (19)$$

Notably, the key point of the optimal control design process lies in the coupling design among cyberattack compensation, uncertainty observation, recursive control transformation, and online policy optimization, rather than the isolated application of individual techniques.

3.2. Stability Analysis

This part provides a rigorous stability analysis based on the Lyapunov method. Theorem 1 summarizes the main stability result, and its proof is subsequently provided to validate the effectiveness of the proposed design.

Theorem 1. Consider the third-order vehicular dynamical system Eq.(1)-Eq.(3) with uncertainties $d_i(t)$ and cyberattacks $\psi_i(t)$. Based on the projection-based compensation mechanism Eq.(4), the fuzzy logic based composite observer Eq.(5)-Eq.(8), the tracking differentiator Eq.(9) and Eq.(10) and actor-critic-identifier structure Eq.(15)-Eq.(19), the vehicle fleet is able to track the prescribed position, velocity and acceleration trajectories, while all closed-loop errors remain SGUUB.

Proof: To analyse the complete closed-loop system, consider the composite Lyapunov function

$$V = \sum_{i=1}^n (V_{O,i} + V_{D,i} + V_{Z,i} + V_{A,i} + V_{C,i} + V_{I,i}), \quad (20)$$

where

$$V_{O,i} = \frac{1}{2} \tilde{X}_i^T P_{O,i} \tilde{X}_i + \frac{1}{2\gamma_{\tau,i}} \tilde{\tau}_i^2, \quad (21)$$

$$V_{D,i} = \sum_{s=1}^2 \frac{1}{2} E_{k,i}^T P_{k,i} E_{k,i}, \quad (22)$$

$$V_{Z,i} = \frac{1}{2} z_i^T z_i, \quad (23)$$

$$V_{A,i} = \frac{1}{2} \tilde{W}_{A,i}^T \Gamma_{A,i}^{-1} \tilde{W}_{A,i}, \quad (24)$$

$$V_{C,i} = \frac{1}{2} \tilde{W}_{C,i}^T \Gamma_{C,i}^{-1} \tilde{W}_{C,i}, \quad (25)$$

$$V_{I,i} = \frac{1}{2} \text{tr}(\tilde{W}_{I,i}^T \Gamma_{I,i}^{-1} \tilde{W}_{I,i}). \quad (26)$$

Then based on the optimal controller design, including the projection-based compensation mechanism Eq.(4), the fuzzy logic based composite observer Eq.(5)-Eq.(8), the tracking differentiator Eq.(9) Eq.(10) and actor-critic-identifier structure Eq.(15)-Eq.(20), we have

$$\begin{aligned} \dot{V} \leq & -\delta_X \sum_{i=1}^n \|\tilde{X}_i\|^2 - \delta_\tau \sum_{i=1}^n \tilde{\tau}_i^2 - \delta_O \sum_{i=1}^n (\|M_i^v\|^2 + \|M_i^a\|^2) \\ & - \delta_T \sum_{i=1}^n \|E_{k,i}\|^2 - \delta_Z \sum_{i=1}^n \|z_i\|^2 - \delta_A \sum_{i=1}^n \|\tilde{W}_{A,i}\|^2 \\ & - \delta_C \sum_{i=1}^n \|\tilde{W}_{C,i}\|^2 - \delta_I \sum_{i=1}^n \|\tilde{W}_{I,i}\|^2 + \Delta, \end{aligned} \quad (27)$$

where $\delta_X, \delta_\tau, \delta_O, \delta_T, \delta_Z, \delta_A, \delta_C, \delta_I$ are strictly positive. The constant Δ collects all irreducible bounded residuals in the closed-loop system. Its principal components arise from the cyberattack variation rate, the external disturbance derivative, the fuzzy approximation errors, the identifier residual, the tracking differentiator residual, and the value-function approximation error. The negative quadratic terms in Eq.(27) can consequently be related to the Lyapunov function itself. Hence, for an appropriate constant $\delta > 0$. Then we have

$$\dot{V} \leq -\delta V + \Delta. \quad (28)$$

And Eq.(28) satisfies

$$0 \leq V(t) \leq e^{-\delta t} V(0) + \frac{\Delta}{\delta}. \quad (29)$$

The first term describes the exponentially decaying influence of the initial errors. The second term describes the residual error induced by bounded cyberattack variations, disturbances, approximation errors, and learning residuals. As time increases, the contribution of the initial conditions vanishes and $V(t)$ approaches a neighborhood whose size is determined by $\frac{\Delta}{\delta}$. This result proves that all closed-loop errors are SGUUB.

In Theorem 1, we have established that the proposed RL enabled optimal platoon control framework guarantees that all closed-loop errors are SGUUB under simultaneous uncertainties and cyberattacks. Specifically, the composite Lyapunov function in Eq.(20) incorporates the tracking errors, observer errors, attack compensation errors, tracking differentiator errors, and neural-network approximation errors. By applying Lyapunov stability analysis, it is proved that the proposed controller can maintain bounded tracking performance while effectively attenuating the effects of unknown nonlinearities, external disturbances, and corrupted measurements. Compared with existing platoon control methods that mainly focus on stability or attack compensation separately, the proposed framework integrates projection-based cyberattack compensation, fuzzy logic based composite observation, second-order tracking

differentiation, and actor-critic-identifier based optimal learning into a unified control architecture. Moreover, the proposed design avoids repeated analytical differentiation of virtual controllers and achieves resilient platoon tracking performance under complex uncertain environments.

4 Simulation Results

To evaluate the effectiveness of the proposed RL-enabled optimal platoon control framework, numerical simulations are conducted in MATLAB for a platoon consisting of one leader and six follower AVs. The results examine trajectory tracking (Fig. 2; Table 1), neural-network learning (Fig. 3), attack resilience (Fig. 4), and inter-vehicle safety (Fig. 5), thereby verifying the theoretical properties established in Section 3. Finally, we discuss the comparison with related works and different control strategies (Table 2; Fig. 6; Fig. 7).

4.1. Parameter Setting

In this paper, we consider the vehicular platoon of one leader and six follower vehicles. For the desired trajectory of the leader vehicle, we adopt a quintic polynomial smooth transition function as

$$S(\sigma) = 10\sigma^3 - 15\sigma^4 + 6\sigma^5, \sigma = \frac{t - t_1}{t_2 - t_1}. \quad (30)$$

In the interval where the velocity transitions from v_1 to v_2 , the desired position and velocity of the leader vehicle are defined as

$$x_r(t) = 150 + \int_0^t v_r(s) ds, \quad (31)$$

$$v_r(t) = v_1 + (v_2 - v_1)S(\sigma). \quad (32)$$

The desired acceleration and jerk are given by the first and second derivatives of this trajectory, respectively. The leader vehicle follows a smooth velocity profile over the 100 s simulation. Its velocity increases from 0 to 12 m/s during the first 8 s, remains constant until 18 s, rises to 22 m/s by 28 s, and is maintained until 38 s. It then accelerates to 30 m/s by 50 s and holds this velocity until 62 s, after which it decelerates to 18 m/s by 74 s. Following an 8 s constant velocity phase, the leader accelerates to 26 m/s by 92 s and maintains this velocity until the end of the simulation. The velocity profile includes repeated acceleration, cruising, and deceleration phases up to 30 m/s, enabling evaluation of platoon tracking performance across a broad operating range. The desired inter-vehicle distance ℓ^* is 19.2 m, consisting of a desired safety gap of 16 m and the vehicle length of 3.2 m. Based on the graph theory, the inter-vehicle communication in vehicular system is shown in Fig. 1, and the parameters are designed as: the adjacency matrix is

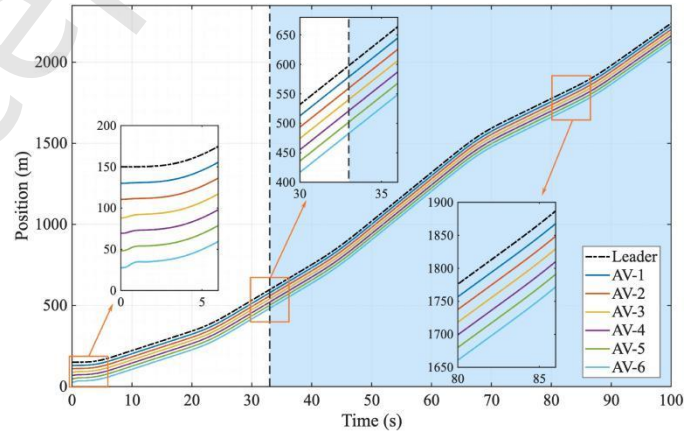
$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

the degree matrix is $\mathcal{D} = \text{diag}(0, 1, 1, 0, 1, 1)$ and the pinning coefficient is $\mathcal{B} = \text{diag}(1, 0, 0, 1, 0, 0)$.

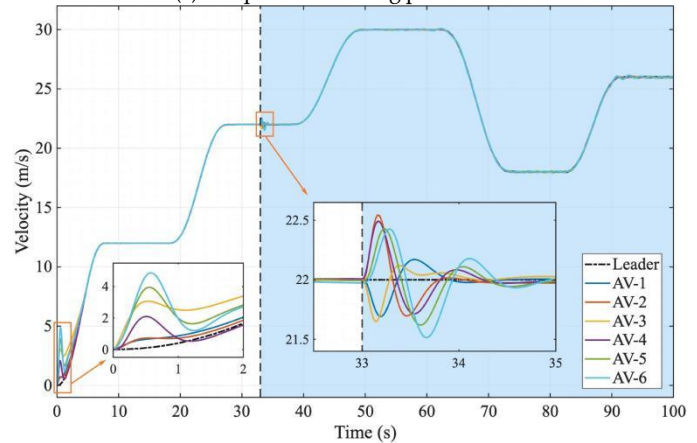
The initial conditions are selected as: initial leader vehicular position $x_r(0) = 150$ m; initial follower vehicular position $x_i(0) = [130, 110.5, 88, 69.4, 48.2, 27.6]^T$; initial velocity and acceleration are $v_i(0) = 0$ m/s and $a_i(0) = 0$ m/s².

In the proposed RL based optimal platoon control framework, the parameter selection follows a hierarchical tuning procedure. First, the observer gains are selected to guarantee sufficiently fast convergence

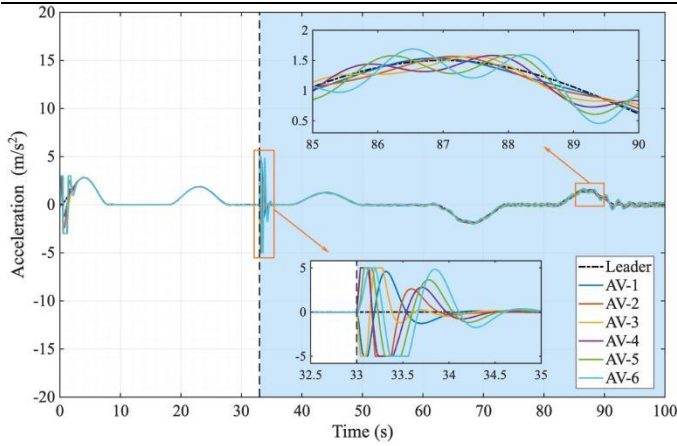
of the state and uncertainty estimation errors while satisfying the stability requirements derived from the Lyapunov analysis. Second, the tracking differentiator parameters are adjusted to balance transient response smoothness and tracking accuracy. Third, the actor-critic-identifier learning rates are selected within a stable range to ensure bounded online adaptation while maintaining satisfactory convergence performance. Notably, the designed parameters need to be positive. Therefore, the control parameters are chosen as: the cyberattack-based gain is $\gamma_{\tau,i} = [541, 556, 572, 581, 603, 617]^T$; the inertial lag parameter is $\lambda_i = [1.01, 1.07, 0.96, 1.01, 1.03, 0.96]^T$; the observer gains are $\mu_i^p = 4\mu_i^*$, $\mu_i^v = 6\mu_i^{*2}$, $\mu_i^a = 4\mu_i^{*3}$ and $\mu_i^d = \mu_i^{*4}$ with the basis gain $\mu_i^* = [10, 10.1, 10.5, 10.7, 10.7, 11]^T$; the differentiator parameters are $\epsilon_{s,i} = 24$ and $r_{s,i} = 20$; the actor-critic-identifier neural-network-based parameters are $\Gamma_{A,i} = \Gamma_{C,i} = [1.2, 0.8, 0.8, 0.6, 0.6, 0.5]^T$ and $\Gamma_{I,i} = [0.8, 1.0, 1.1, 0.9, 0.6, 0.8]^T$. In addition, the duration of the simulation experiment is 100 s, with a time-step of 0.0025 s. It should be emphasized that the parameter settings in this simulation are selected to verify the effectiveness of the proposed control framework under the considered platoon scenario. The objective of parameter selection is to satisfy the theoretical stability requirements and provide a representative evaluation environment, rather than to obtain a universal parameter configuration for all possible vehicle dynamics and operating conditions.



(a) The position tracking performance



(b) The velocity tracking performance



(c) The acceleration tracking performance

Fig. 2. The position, velocity and acceleration tracking performance of AVs under the proposed optimal platoon control framework.

4.2. Results Analysis

Fig. 2 presents the position, velocity, and acceleration tracking performance of the vehicle platoon. As shown in Fig. 2(a), the follower AVs start from different initial positions and gradually form the prescribed platoon configuration. The position trajectories subsequently evolve in parallel with the leader trajectory while maintaining the desired spacing. The enlarged views show that the position responses remain continuous and bounded during the initial transient, the cyberattack occurring at 33 s, and the subsequent control process. Fig. 2(b) shows that all follower AVs accurately track the multiple acceleration, constant-velocity, and deceleration stages of the leader. The initial velocity errors converge rapidly, and the follower velocities remain close to the leader velocity over the entire operating range. After the cyberattack is activated, short-duration oscillations appear in the velocity responses. These oscillations are quickly attenuated, and velocity coordination is subsequently recovered. Fig. 2(c) further shows that the acceleration trajectories remain bounded during the entire simulation. Relatively large adjustments occur during initialization and immediately after the attack, reflecting the compensation action required to restore the platoon states.

Table 1. Quantitative evaluation of ultimate tracking errors.

Statistic	Errors		
	Position (m)	Velocity (m/s)	Acceleration (m/s ²)
Overall mean	0.03966	0.06335	0.21198
10-30 s max	0.03267	0.03111	0.05433
10-30 s mean	0.00785	0.00460	0.01154
40-80 s max	0.06936	0.11727	0.25730
40-80 s mean	0.01598	0.02316	0.04854

To quantitatively validate the ultimate boundedness guaranteed by the Lyapunov analysis, the practical ultimate tracking errors are evaluated under the simulation conditions. As summarized in Table 1, the proposed controller achieves small bounded errors in position, velocity, and acceleration throughout the entire platoon operation, including the pre-attack steady-state interval (10-30 s) and the post-attack recovery stage for all follower vehicles (40-80 s). The obtained ultimate tracking errors remain small and bounded over the

evaluated operating intervals, indicating that the proposed controller effectively prevents persistent error growth under the simultaneous influence of system uncertainties and cyberattacks. After the transient responses, the position, velocity, and acceleration tracking errors are confined within compact neighborhoods around zero and exhibit no tendency toward divergence. This behavior is consistent with the ultimate boundedness predicted by the Lyapunov analysis. Therefore, the quantitative results provide numerical evidence supporting the SGUUB property established in Theorem 1 and further demonstrate the capability of the proposed RL-enabled optimal platoon control framework to preserve accurate and stable platoon tracking under uncertain and adversarial operating conditions.

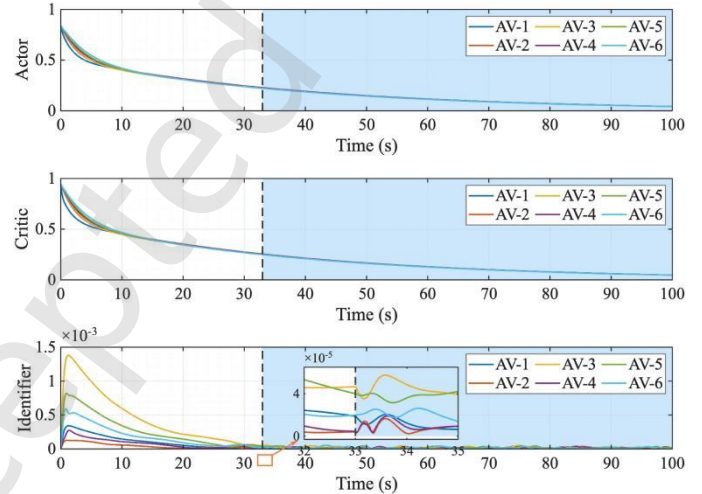


Fig. 3. The weight norms of the actor-critic-identifier NNs.

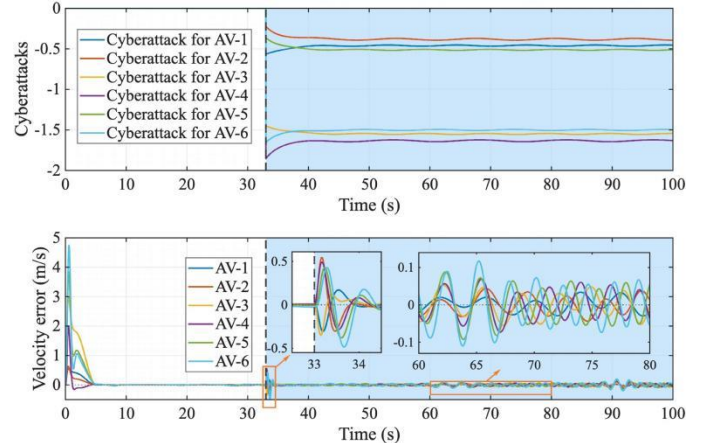


Fig. 4. The cyberattack signals and velocity tracking error of AVs.

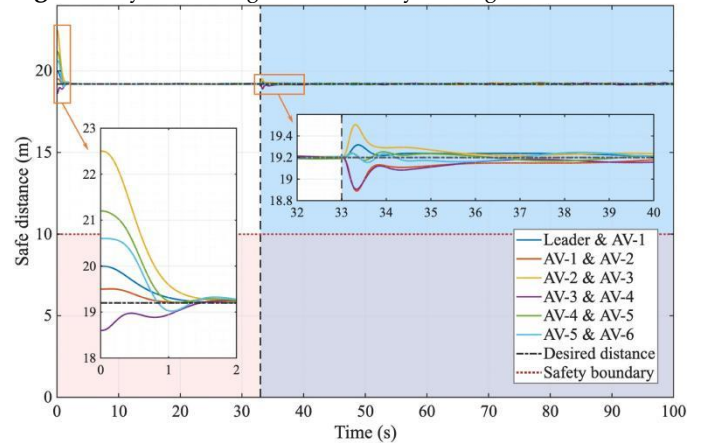


Fig. 5. The inter-vehicle distances between consecutive AVs.

Fig. 3 illustrates the evolution of the actor, critic, and identifier neural-network weight norms. The actor and critic weights decrease smoothly from their initial values and remain bounded throughout the simulation. The identifier weights rapidly approach a small neighborhood during the initial learning stage. A bounded adjustment appears after 33 s because the cyberattacks change the transformed error dynamics. The weights subsequently remain bounded without divergence, confirming the stable online adaptation of the actor–critic–identifier structure.

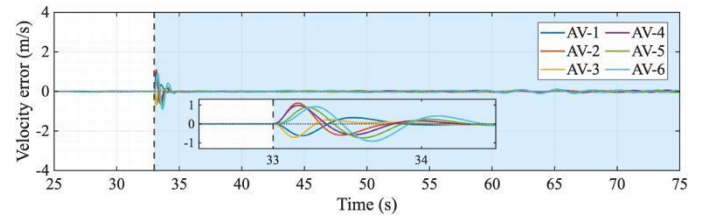
Fig. 4 evaluates the resilience of the proposed framework under time-varying cyberattacks. The upper panel shows that heterogeneous attack signals are injected into the six follower AVs at 33 s. These signals contain different biases and time-varying components while remaining bounded. The lower panel presents the corresponding velocity tracking errors. The errors converge rapidly during the initial stage and exhibit only transient increases after attack activation. Under the projection-based compensation mechanism and fuzzy logic based composite observer, the attack-induced deviations are attenuated and the velocity errors return to small neighborhoods. The later bounded oscillations are associated with changes in the leader trajectory, nonlinear dynamics, and external uncertainties.

Fig. 5 presents the distances between consecutive vehicles. Owing to the non-equilibrium initial positions, the initial inter-vehicle distances differ from the desired value. Under the proposed controller, all spacing trajectories rapidly converge to the prescribed distance of 19.2 m. Small transient deviations occur near the cyberattack instant, after which the distances return to the desired neighborhood. Throughout the simulation, all inter-vehicle distances remain above the 10 m safety boundary. This result demonstrates that the proposed framework maintains collision-free platoon operation during initial convergence, cyberattack activation, and leader-velocity variations.

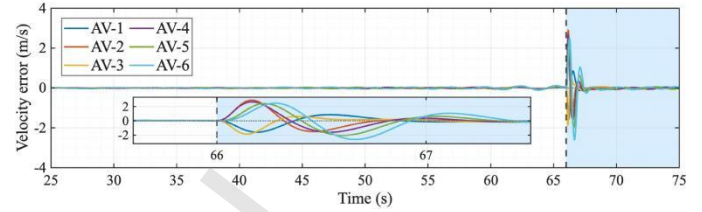
Table 2. Comparison of the proposed method with related works.

References	Considered aspects		
	Optimal control	Uncertainty	Cyberattack
[1] in 2017	✓	✗	✗
[2] in 2018	✗	✓	✗
[3] in 2019	✓	✗	✗
[4] in 2020	✓	✗	✗
[5] in 2021	✓	✗	✓
[6] in 2022	✗	✗	✓
[7] in 2023	✗	✓	✓
[8] in 2024	✗	✓	✗
[9] in 2025	✓	✗	✓
[10] in 2026	✗	✓	✓
This paper	✓	✓	✓

Note: [1] is (Mirzaei and Mirzaeinejad, 2017); [2] is (Ji et al., 2018); [3] is (Ferdowsi et al., 2019); [4] is (Dong et al., 2020); [5] is (Liu et al., 2021b); [6] is (Xie et al., 2022); [7] is (Liu et al., 2023); [8] is (Wen et al., 2024); [9] is (Tajdari et al., 2025); [10] is (Chen et al., 2026).



(a) The velocity tracking performance in the presence of cyberattacks initiated at 33s



(b) The velocity tracking performance in the presence of cyberattacks initiated at 66s

Fig. 6. The comparison of tracking performance of AVs under cyberattacks at different attack instants.

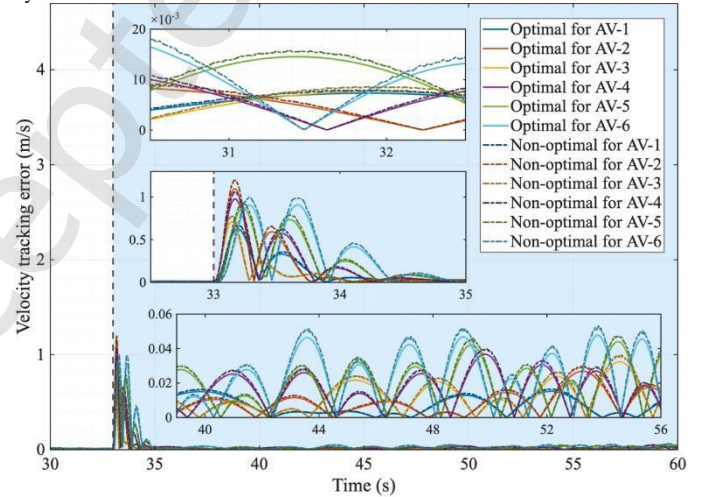


Fig. 7. The comparison of velocity tracking performance between optimal and non-optimal controllers under cyberattacks.

4.3. Comparisons with Existing Literature

Table 2 compares the proposed method with representative studies in terms of optimal control, uncertainties, and cyberattacks.

Early studies primarily concentrated on individual performance objectives, such as optimal motion regulation or uncertainty compensation, under comparatively structured operating conditions (Dong et al., 2020; Mirzaei and Mirzaeinejad, 2017). As connected and autonomous vehicle systems became more dependent on communication and sensing, subsequent research gradually incorporated cyberattack detection, resilient estimation, heterogeneous dynamics, and robust platoon coordination. Recent studies have further combined multiple practical factors, particularly uncertainties and cyberattacks, within unified control frameworks (Liu et al., 2021; Tajdari et al., 2025; Xie et al., 2022). Nevertheless, the simultaneous consideration of RL enabled optimal control, nonlinear uncertainties, and cyberattack resilience remains limited. This progression indicates that current research is moving from isolated performance enhancement toward comprehensive optimization and security-aware control for vehicles operating in uncertain and adversarial environments.

Specifically, compared with (Dong et al., 2020; Mirzaei and Mirzaeinejad, 2017), the proposed optimal control strategy

simultaneously considers system uncertainties and cyberattacks, thereby enhancing the robustness of the optimal control framework. Moreover, compared with (Chen et al., 2026; Liu et al., 2023), this paper further performs a comprehensive optimization of the control process. Hence, the proposed platoon control framework simultaneously integrates RL based optimal control strategy, cyberattack compensation, and uncertainty estimation. The actor-critic-identifier structure optimizes the virtual and actual control policies, while the projection based compensation mechanism and fuzzy logic based composite observer recover reliable state information under unknown nonlinear dynamics, uncertainties, and corrupted measurements. Therefore, the proposed method provides a more comprehensive solution for high-order vehicle platoons operating in uncertain and adversarial environments.

Fig. 6 further demonstrates the robustness of the proposed optimal platoon control framework against cyberattacks occurring at different instants. When cyberattacks are launched at $t = 33$ s and $t = 66$ s, the velocity tracking errors experience short-duration transient fluctuations but rapidly recover to small bounded neighborhoods. This indicates that the proposed projection-based compensation mechanism and fuzzy logic based composite observer can effectively attenuate attack-induced disturbances and maintain stable platoon tracking performance. Therefore, the proposed framework exhibits strong resilience under time-varying cyberattack scenarios. Fig. 7 compares the velocity tracking performance of the optimal and non-optimal controllers under cyberattacks. Notably, optimal control strategy refers to the control framework proposed in this paper, while non-optimal control strategy denotes the backstepping-based adaptive control strategies presented in (Wang et al., 2025; Xie et al., 2022). Before the occurrence of cyberattacks, both controllers achieve small velocity tracking errors, indicating accurate platoon tracking performance. After the cyberattacks are activated, the tracking errors increase sharply due to the corrupted information, but they rapidly converge back to bounded neighborhoods. Compared with the non-optimal controller, the proposed optimal controller consistently achieves smaller tracking errors and reduced oscillations throughout the attack scenario. These results demonstrate the effectiveness of the actor-critic-identifier based optimization mechanism in improving tracking accuracy and enhancing the robustness of the proposed platoon control framework.

5 Conclusion and Future Research Directions

This paper developed an RL enabled optimal platoon control framework for third-order AVs subject to uncertainties and cyberattacks. A projection-based compensation mechanism and a fuzzy logic based composite observer were designed to recover reliable state information and estimate system uncertainties and cyberattacks. Second-order tracking differentiators were incorporated into the optimized backstepping procedure to reduce recursive differentiation, while the RL based actor-critic-identifier structure enabled online optimization of the virtual and actual control policies. Composite Lyapunov analysis established the SGUUB property of all closed-loop errors within the vehicular system. Numerical results further demonstrated accurate position, velocity, and acceleration tracking, bounded learning parameters, and safe inter-vehicle spacing under uncertainties and cyberattacks.

Future work will focus on three practical extensions. First, communication and information transmission delays will be

incorporated into the distributed platoon control framework to improve robustness under realistic vehicle-to-vehicle (V2V) communication conditions. Second, coupled longitudinal and lateral vehicle dynamics will be considered to support coordinated platooning, lane changing, and formation reconfiguration in complex traffic scenarios. Third, constrained control strategies will be developed to explicitly enforce velocity and acceleration limits, thereby improving driving safety, passenger comfort, and the practical feasibility of the proposed framework.

Author contributions

Ziming Wang: Writing – original draft, Formal analysis, Software, Conceptualization. **Karl H. Johansson:** Methodology, Resources, Supervision. **Apostolos I. Rikos:** Writing – review & editing, Validation, Supervision.

Replication and data sharing

The program code within this research can be made accessible upon request via email to the corresponding author.

Acknowledgements

The work of A.I.R. was supported by the Guangdong Provincial Project (Grant No. 2024QN11G109).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

- Besselink, B., Johansson, K. H., 2017. String stability and a delay-based spacing policy for vehicle platoons subject to disturbances. *IEEE Trans. Autom. Control*, 62, 4376-4391.
- Chen, G., Zhou, Q., Ren, H., Li, H., 2023. Sensor-fusion-based event-triggered following control for nonlinear autonomous vehicles under sensor attacks. *IEEE Trans. Autom. Sci. Eng.*, 22, 17411-17420.
- Chen, X., Zhai, C., Chen, C., Ma, Y., Wang, B., Yan, C., Chen, Y., 2026. Adaptive resilient control for vehicle platoons under hybrid cyberattacks, actuator saturation, and parameter uncertainties. *IEEE Internet Things J.*, 13, 6669-6682.
- Dong, F., Zhao, X., Chen, Y. H., 2020. Optimal longitudinal control for vehicular platoon systems: adaptiveness, determinacy, and fuzzy. *IEEE Trans. Fuzzy Syst.*, 29, 889-903.
- Fei, Y., Shi, P., Li, Y., Liu, Y., Qu, X., 2024. Formation control of multi-agent systems with actuator saturation via neural-based sliding mode estimators. *Knowl.-Based Syst.*, 284, 111292.
- Ferdowsi, A., Ali, S., Saad, W., Mandayam, N. B., 2019. Cyber-physical security and safety of autonomous connected vehicles: optimal control meets multi-armed bandit learning. *IEEE Trans. Commun.*, 67, 7228-7244.
- He, X., Hashemi, E., Johansson, K. H., 2021. Distributed control under compromised measurements: resilient estimation, attack detection, and vehicle platooning. *Automatica*, 134, 109953.
- Ioannou, P. A., Chien, C. C., 1993. Autonomous intelligent cruise control. *IEEE Trans. Veh. Technol.*, 42, 657-672.
- Ji, X., Liu, Y., He, X., Yang, K., Na, X., Lv, C., et al., 2018. Interactive control paradigm-based robust lateral stability controller design for autonomous automobile path tracking with uncertain disturbance: a dynamic game approach. *IEEE Trans. Veh. Technol.*, 67, 6906-6920.
- Liu, A., Li, T., Gu, Y., Dai, H., 2021a. Cooperative extended state observer based control of vehicle platoons with arbitrarily small time headway. *Automatica*, 129, 109678.
- Liu, Y., Li, Y., Wang, Y., Zhang, X., Gooi, H. B., Xin, H., 2021b. Robust and resilient distributed optimal frequency control for microgrids against cyber attacks. *IEEE Trans. Ind. Inform.*, 18, 375-386.
- Liu, Y., Xu, L., Cai, G., Yin, G., Yan, F., 2023. Distributed robust platooning control for heterogeneous vehicle group under parametric uncertainty and hybrid attacks. *IEEE Trans. Veh. Technol.*, 72, 5677-5689.
- Liu, Y., Yao, D., Li, H., Lu, R., 2022. Distributed cooperative compound tracking control for a platoon of vehicles with adaptive NN. *IEEE Trans. Cybern.*, 52, 7039-7048.
- Mirzaei, M., Mirzaeinejad, H., 2017. Fuzzy scheduled optimal control of integrated vehicle braking and steering systems. *IEEE/ASME Trans. Mechatron.*, 22, 2369-2379.
- Mousavinejad, E., Vlacic, L., 2022. Secure platooning control of automated vehicles under cyber attacks. *ISA Trans.*, 127, 229-238.
- Narri, V., Alanwar, A., Mårtensson, J., Pettersson, H., Nordin, F., Johansson, K. H., 2025. Situational awareness using set-based estimation and vehicular communication: an occluded pedestrian-crossing scenario. *Commun. Transp. Res.*, 5, 100190.
- Ren, H., Cheng, Z., Qin, J., Lu, R., 2023. Deception attacks on event-triggered distributed consensus estimation for nonlinear systems. *Automatica*, 154, 111100.
- Sacone, S., 2024. Platoon control in traffic networks: new challenges and opportunities. *IEEE Trans. Intell. Transp. Syst.*, 25, 149-183.
- Shladover, S., Desoer, C., Hedrick, J., Tomizuka, M., Walrand, J., Zhang, W., et al., 1991. Automated vehicle control developments in the PATH program. *IEEE Trans. Veh. Technol.*, 40, 114-130.
- Song, X., Sun, P., Ahn, C. K., Song, S., 2024. Reinforcement learning-based event-triggered predefined-time optimal fuzzy control for nonlinear constrained systems. *IEEE Trans. Fuzzy Syst.*, 32, 4646-4659.
- Tajdari, F., Papaioannou, G., Happee, R., 2025. Optimal-coupling-observer av motion control securing comfort in the presence of cyber attacks. *IEEE Trans. Intell. Transp. Syst.*, 27, 1426-1441.
- Wang, Z., Zhang, Y., Zhao, C., Yu, H., 2025. Adaptive event-triggered formation control of autonomous vehicles. <https://doi.org/10.48550/arXiv.2506.06746>.
- Wang, Z., Ge, X., Han, Q.-L., 2026a. Event-triggered V2X communication and control co-design for intelligent electrified transportation systems. *J. Intell. Connect. Veh.*, early accepted.
- Wang, Z., Jiang, G., Zhang, Y., Johansson, K. H., Rikos, A. I., 2026b. Cooperative switched formation control of autonomous vehicles: An event-triggered approach to input saturation and time-delay challenges. <https://doi.org/10.48550/arXiv.2606.11971>.
- Wang, Z., Piao, S., Wang, X., Tsung, F., Johansson, K. H., Rikos, A. I., 2026c. SAFE: Switched adaptive formation via event-triggered control for multi-vehicle systems. *IEEE Syst., Man, Cybern. Lett.*, early accepted.
- Wen, G., Chen, C. P., Ge, S. S., 2020. Simplified optimized backstepping control for a class of nonlinear strict-feedback systems with unknown dynamic functions. *IEEE Trans. Cybern.*, 51, 4567-4580.
- Wen, G., Wan, Y., Zhou, J., Zheng, D., Chen, C. P., 2024. Distributed robust event-triggered platooning control of connected vehicles with uncertain dynamics: a neuro-adaptive approach. *IEEE Trans. Ind. Inform.*, 21, 2570-2579.
- Xie, M., Ding, D., Ge, X., Han, Q.-L., Dong, H., Song, Y., 2022. Distributed platooning control of automated vehicles subject to replay attacks based on proportional integral observers. *IEEE/CAA J. Autom. Sin.*, 11, 1954-1966.
- Yang, Y., Zhan, J., Liu, Y., Wang, Q., 2025. Cross-city transfer learning: Applications and challenges for smart cities and

sustainable transportation. *Commun. Transp. Res.*, 5, 100206.

Zhao, J., Yang, G. H., 2024. Observer-based finite-time fuzzy adaptive resilient control for uncertain nonlinear systems against deception attacks and unknown dead

zones. *IEEE Trans. Autom. Sci. Eng.*, 22, 8762-8772.

Zhao, L., Cao, X., Li, X., 2022. Adaptive sliding-mode control for inertial reference units via adaptive tracking differentiators. *IEEE Trans. Syst. Man Cybern. Syst.*, 53, 3208-3218.



Ziming Wang received the B.Eng. degree in Electronic Information Engineering from Southwest University

in 2023 and the M.Phil. degree in the Trust of Robotics and Autonomous Systems from the Hong Kong University of Science and Technology (Guangzhou) in 2025. He was a visiting scholar at the School of Vehicle and Mobility, Tsinghua University. He is currently a PhD student in the Thrust of Artificial Intelligence from the HKUST-GZ. He is a senior reviewer of the IEEE TITS, IEEE TSMC, IEEE TMC and IEEE/CAA JAS. His research interests include control theory, optimization, reinforcement learning, and their applications in intelligent transportation systems.



Karl Henrik Johansson is Swedish Research Council Distinguished Professor in Electrical Engineering

and Computer Science at KTH Royal Institute of Technology in Sweden and Founding Director of Digital Futures. He earned his MSc degree in Electrical Engineering and PhD in Automatic Control from Lund University. He has held visiting positions at UC Berkeley, Caltech, NTU and other prestigious institutions. His research interests focus on networked control systems and cyber-physical systems with

applications in transportation, energy, and automation networks. For his scientific contributions, he has received numerous best paper awards and various distinctions from IEEE, IFAC, and other organizations. He has been awarded Distinguished Professor by the Swedish Research Council, Wallenberg Scholar by the Knut and Alice Wallenberg Foundation, Future Research Leader by the Swedish Foundation for Strategic Research. He has also received the triennial IFAC Young Author Prize and IEEE CSS Distinguished Lecturer. He is the recipient of the 2024 IEEE CSS Hendrik W. Bode Lecture Prize. His extensive service to the academic community includes being President of the European Control Association, IEEE CSS Vice President Diversity, Outreach & Development, and Member of IEEE CSS Board of Governors and IFAC Council. He has served on the editorial boards of *Automatica*, *IEEE TAC*, *IEEE TCNS* and many other journals. He has also been a member of the Swedish Scientific Council for Natural Sciences and Engineering Sciences. He is Fellow of both the IEEE and the Royal Swedish Academy of Engineering Sciences.



Apostolos I. Rikos is an Assistant Professor at the Artificial Intelligence Thrust of the Information Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China. He is also affiliated with the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China. He received his B.Sc., M.Sc., and Ph.D. degrees in Electrical Engineering from the Department of Electrical and Computer Engineering, University of Cyprus in 2010, 2012, and 2018, respectively. In 2018, he joined the KIOS Research and Innovation Center of Excellence in Cyprus, where he was a Research Lecturer. In 2020, he joined the Division of Decision and Control Systems of KTH Royal Institute of Technology as a Postdoctoral Researcher. In 2023, he joined the Department of Electrical and Computer Engineering, Division of Systems Engineering, at Boston University as a Postdoctoral Associate. His research interests are in the area of distributed optimization and learning, distributed network control and coordination, privacy and security.