

Article

PPO-Based Hybrid Optimization for RIS-Assisted Semantic Vehicular Edge Computing

Wei Feng ^{1,†}, Jingbo Zhang ^{1,†}, Qiong Wu ^{1,*} , Pingyi Fan ²  and Qiang Fan ³

¹ School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China; jingbozhang@stu.jiangnan.edu.cn (J.Z.)

² State Key Laboratory of Space Network and Communications, Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

³ Qualcomm, San Jose, CA 95110, USA

[†] These authors contributed equally to this work.

Abstract

To support latency-sensitive Internet of Vehicles (IoV) applications amidst dynamic environments and intermittent links, this paper proposes a Reconfigurable Intelligent Surface (RIS)-aided semantic-aware Vehicle Edge Computing (VEC) framework. This approach integrates RIS to optimize wireless connectivity and semantic communication to minimize latency by transmitting semantic features. We formulate a comprehensive joint optimization problem by optimizing offloading ratios, the number of semantic symbols, and RIS phase shifts. Considering the problem's high dimensionality and non-convexity, we propose a two-tier hybrid scheme that employs Proximal Policy Optimization (PPO) for discrete decision-making and Linear Programming (LP) for offloading optimization. The simulation results have validated the proposed framework's superiority over existing methods. Specifically, the proposed PPO-based hybrid optimization scheme reduces the average end-to-end latency by approximately 40% to 50% compared to Genetic Algorithm (GA) and Quantum-behaved Particle Swarm Optimization (QPSO). Moreover, the system demonstrates strong scalability by maintaining low latency even in congested scenarios with up to 30 vehicles.

Keywords: vehicular edge computing; reconfigurable intelligent surface; semantic communication; task offloading

1. Introduction

The Internet of Vehicles (IoV) is driving the intelligent transformation of transportation systems by enabling real-time connectivity between vehicles, infrastructure, and pedestrians to support autonomous driving. However, the massive high-frequency data generated by onboard sensors is challenging in traditional central cloud architectures, especially to meet low-latency and high-reliability requirements [1–3]. To address this, Vehicular Edge Computing (VEC) allows vehicles to offload tasks to edge servers (ESs) at roadside units (RSUs), which significantly reduces latency and energy consumption by shortening communication distances [4], effectively alleviating core network loads.

Nevertheless, urban obstacles and multipath fading often degrade vehicle–RSU links [5], potentially violating Quality of Service (QoS) standards [6]. Thus, reconfigurable Intelligent Surfaces (RISs) offer a solution by dynamically adjusting electromagnetic signal properties to reroute signals, boost strength, and suppress interference [7]. Research



Academic Editor: Hector E. Nistazakis

Received: 22 January 2026

Revised: 15 February 2026

Accepted: 24 February 2026

Published: 25 February 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

on RIS-aided V2X has optimized these critical links [8,9]. As investigated in [10], RIS empowers networks to proactively manipulate the wireless channel rather than passively adapting to it. Specifically, each passive element on the RIS can independently impose a controlled phase shift on the incident electromagnetic waves. By optimizing the phase-shift matrix, RIS reshapes the propagation environment to compensate for high-frequency attenuation and urban shadows. This mechanism is particularly crucial for semantic-aware VEC, as it ensures high-fidelity transmission of semantic features which are highly sensitive to channel quality.

Concurrently, semantic communication has emerged as a paradigm shift by moving beyond traditional Shannon entropy limits to focus on the effective transmission of meaning. A foundational DeepSC framework proposed by [11] has employed deep learning to jointly optimize source and channel encoding. Unlike bit-level transmission, it extracts and transmits only task-relevant semantic features, thereby reducing data volume while maintaining high information integrity even in a low signal-to-interference-plus-noise ratio (SINR) environments. This approach ensures that minor decoding errors do not compromise information integrity [12], while the reduced data volume significantly lowers latency [13].

Integrating semantic communication with RIS-aided VEC becomes a critical development direction. RIS enhances signal coverage against attenuation, while semantic communication minimizes transmission overhead and enables intelligent resource scheduling [14,15], thus delivering high adaptability under extreme conditions [16]. However, high vehicle mobility causes rapid channel variations, making traditional static optimization methods ineffective for such non-convex problems. Consequently, artificial intelligence [17], specifically deep reinforcement learning (DRL) [18], has gained prominence for its ability to solve complex dynamic problems.

The integration of these components presents several challenges, including complex semantic extraction, time-varying link quality, and difficulty of jointly optimizing offloading ratios, RIS phase shifts, and the number of semantic symbols. To address these issues, this paper proposes a novel optimization framework for an RIS-aided, semantic-aware VEC system. The Proximal Policy Optimization (PPO) algorithm is employed to optimize RIS phase shifts and the number of semantic symbols, while Linear Programming (LP) is used to allocate offloading ratios. The main contributions are summarized as follows:

- We propose an RIS-aided semantic-aware VEC system, in which a three-path semantic task partitioning framework is designed. Each vehicular task is adaptively divided into three parts: local execution, a vehicle-to-infrastructure (V2I) task offloaded to the RSU, and a vehicle-to-vehicle (V2V) task offloaded to a service vehicle (SV). A link-level RIS enhancement mechanism is introduced, enabling the RIS to improve the quality of either the V2I and V2V link based on real-time semantic similarity, mobility, and channel conditions.
- We formulate a comprehensive joint optimization problem involving offloading rates, the number of semantic symbols, and RIS phase shifts in order to minimize transmission latency. To address the non-convexity of this problem, we propose a two-layer collaborative hybrid framework: PPO is employed to make discrete decisions on the number of semantic symbols and RIS phase shifts, while LP is utilized to optimize the offloading ratio.
- Extensive simulation results have demonstrated that the proposed method outperforms other traditional schemes significantly in terms of end-to-end latency.

The remainder of this paper is organized as follows: Section 2 reviews the related work on DRL-based task offloading, RIS-aided VEC systems, and the latest advancements in semantic communications. Section 3 describes the RIS-aided semantic-aware VEC

system model and formulates the joint optimization problem for minimizing system latency. Section 4 details the proposed two-layer collaborative hybrid framework, which integrates the PPO algorithm for discrete configuration and LP for continuous task offloading. The computational complexity of the proposed scheme is also analyzed in this section. Section 5 presents the simulation setup and discusses the performance results of the proposed method in comparison with several benchmarks. Finally, Section 6 concludes this paper and discusses future research directions.

2. Related Work

In this section, we first review DRL-based Mobile Edge Computing (MEC) studies, then summarize existing research on RIS-aided VEC task offloading, and finally discuss semantic communication-assisted VEC work.

Several studies have applied DRL to VEC [19], addressing critical issues such as pricing, security [20], and task dependencies [21]. Wang et al. [22] proposed an Actor-Critic-based task offloading model that utilizes a partial offloading strategy to effectively reduce system costs in multi-vehicle scenarios. Their approach accelerates policy convergence, offering a novel solution for real-time VEC optimization. From an incentive mechanism's perspective, Wu et al. [23] introduced a pricing-driven framework and developed Pricing-Driven Resource Allocation (PDRA) and DRL-based Pricing-Driven Dynamic Resource Allocation (DPDDRA) algorithms. These methods enable scalable and efficient resource management adaptable to rapid mobility and channel variations. Addressing tasks with strong dependencies and strict latency constraints, Zeng et al. [24] proposed a joint optimization method using Double Deep Q-Network (DDQN). By comprehensively modeling task priorities and dependencies, they achieve collaborative optimization of offloading decisions, bandwidth, and computing resources within a multi-stage decision-making process. Furthermore, Ren et al. [25] integrated blockchain and Asynchronous Advantage Actor-Critic (A3C) into the VEC architecture. Through a two-layer Software-Defined Network (SDN) comprising domain and regional controls, they realized secure and trusted service migration, significantly reducing latency and energy consumption while improving throughput. In summary, while these DRL-based studies demonstrate significant adaptability, they predominantly rely on traditional communication links, largely neglecting the potential of RIS-aided environments and semantic-aware models.

The integration of RIS in VEC has been explored for energy efficiency [10,26] and security [27], yet often neglects semantic aspects. Wang et al. [28] proposed an RIS-aided NOMA VEC architecture, employing a two-stage offloading scheme to effectively reduce vehicle energy consumption. Addressing coverage blind spots, Xie et al. [29] utilized dual-roadside RISs and achieved near-optimal real-time offloading via multi-dimensional parameter optimization. In aerial-assisted networks, Rzig et al. [30] developed a HAPS-UAV cooperative framework, using iterative algorithms to balance offloading success rates and energy usage. Similarly, Xiao et al. [31] maximized the minimum weighted throughput in an RIS-UAV MEC system through a synchronous offloading scheme. Collectively, while validating the benefits of RIS in VEC, these works remain confined to traditional links, lacking integration with semantic communication and dynamic link-level RIS selection.

Recent studies have significantly advanced semantic VEC through AI-driven mechanisms, covering areas from efficient offloading [32] and privacy-preserving frameworks [33] to robust transmission in complicated conditions. Specifically, Yang et al. [34] utilized a diffusion-based multi-agent reinforcement learning algorithm to enhance exploration capabilities, effectively reducing system latency and costs. To improve transmission reliability, Du et al. [35] integrated Large Multimodal Models (LMMs) into vehicular networks, enhancing task accuracy in low-SINR environments through task-oriented semantic mech-

anisms. Addressing model maintenance, Zheng et al. [36] proposed a privacy-preserving personalized federated learning framework to overcome privacy and mobility bottlenecks. Furthermore, Chen et al. [37] developed a semantic-aware multimodal model using an MAPPO algorithm to jointly optimize offloading and resource allocation, thereby boosting the Quality of Experience (QoE).

Despite these advancements, existing research often optimizes semantic extraction and resource allocation in isolation. But there are few comprehensive studies that jointly model semantic communication with physical link characteristics, particularly neglecting synergy between RIS assistance and multipath offloading strategies. Mao et al. [38] proposed an RIS-enhanced semantic-aware sensing, communication, computation, and control (SC3) network framework, focusing on closed-loop control optimization in industrial IoT scenarios. In contrast, this paper focuses on the highly dynamic scenarios of the IoV, proposing a unique three-path semantic task partitioning framework, and specifically models the randomness introduced by vehicle mobility, which is fundamentally different from the fixed-device scheduling in industrial scenarios. Zhao et al. [39] investigated traffic reshaping and channel reconfiguration in RIS-aided semantic NOMA networks, with an emphasis on optimizing decoding order to reduce energy consumption. In comparison, this paper introduces V2V cooperative semantic offloading while leveraging SVs as auxiliary computing nodes, which makes more efficient use of idle computing resources in the IoV than pure access point (AP) offloading. Xie et al. [40] studied the fairness of semantic communication systems assisted by STAR-RIS in order to enhance semantic similarity by maximizing the minimum SINR. Different from their work, this paper optimizes the physical-layer channel via RIS, but also directly takes the number of semantic symbols as a decision variable for cross-layer optimization. This three-dimensional joint optimization scheme, which deeply couples the physical-layer phase matrix with the application-layer semantic compression depth, offers greater flexibility than beamforming alone when handling latency-sensitive tasks.

To the best of our knowledge, this is among the first works to jointly consider offloading allocation, RIS phase shift optimization, and the number of semantic symbols in an RIS-aided semantic-aware VEC system.

3. System Model and Problem Formulation

In recent years, with the continuous growth of wireless communication demands, a small bit error in traditional communication systems may lead to a significant semantic error. In contrast, in semantic communication, errors from the decoding process do not affect the integrity of the transmitted information. For example, the words “youth” and “Adolescent” can be mapped to the same semantic symbols, even though their spellings are completely different. Therefore, we utilize the pre-trained DeepSC model [11] to improve communication efficiency by focusing on the meaning of the transmitted data rather than just the raw data itself. The semantic communication framework is illustrated in Figure 1, which consists of semantic encoding, channel encoding, channel decoding, semantic decoding, the channel, and noise. All vehicles and RSUs share the same semantic knowledge base and are equipped with the DeepSC system. However, it is important to note that maintaining knowledge base consistency in highly dynamic IoV environments presents a significant challenge. In this study, we focus on the joint optimization of resource allocation under a static knowledge base as a foundational step. The potential overhead and latency associated with real-time, dynamic knowledge base updates—necessitated by varying semantic tasks—are considered a limitation of the current framework and will be explored in future research. We assume that the transmitted task features are consistent with those in [11], and thus directly reuse the encoding results of the DeepSC semantic

communication model. Furthermore, we combine these encoding results with relevant information such as SINR to derive the relevant variables for semantic communication, eliminating additional data compression and storage steps originally required by the DeepSC model [41].

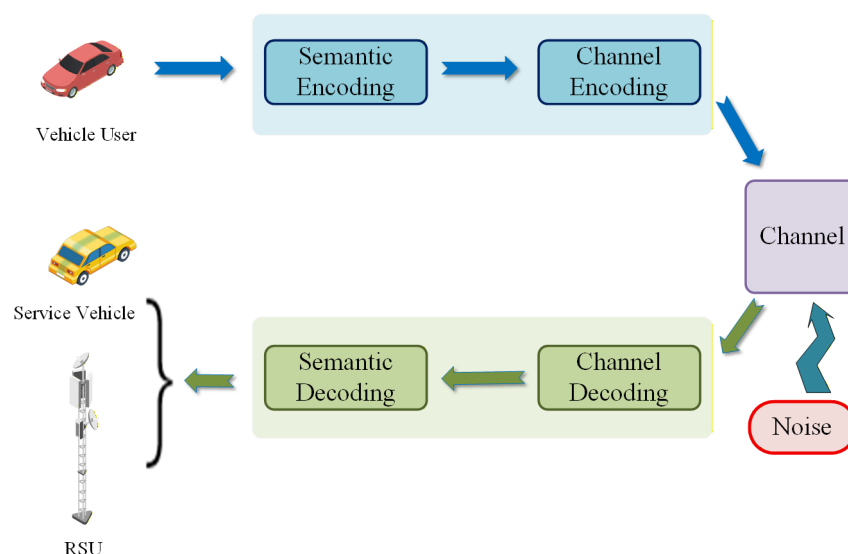


Figure 1. Framework of semantic communication system.

3.1. Scenario

As shown in Figure 2, the RIS-aided semantic-aware VEC system comprises K vehicle users $\mathcal{K} = \{1, \dots, K\}$ and edge nodes $\mathcal{J} = \{0, \dots, J\}$. Here, 0 denotes the MEC-enabled RSU, while $1 \sim J$ are SVs. Any vehicle other than your own may serve as an SV. Time is discretized into slots $\mathcal{T} = \{1, \dots, T\}$ of duration Δt . The RSU is deployed in a relatively open area near the center and equipped with a large number of multiple-input multiple-output (MIMO) antennas. Due to the high vehicle mobility, each vehicle is equipped with an omnidirectional antenna. In each time slot Δt , vehicles are uniformly distributed on the road. Regarding task characteristics, the task arrivals of vehicle users follow a Poisson distribution, and the text-based tasks generated by vehicle users have a fixed size. In practical IoV scenarios, text-based tasks may vary in length. However, the adopted DeepSC-based semantic encoder maps textual information into a controllable number of semantic symbols, which mitigates the impact of bit-level length variations. The integration with DeepSC is realized by mapping the high-dimensional features extracted by the semantic encoder into a variable number of semantic symbols v . This allows the framework to adaptively perform ‘semantic compression’ based on the real-time channel quality $h_{k,j}^t$ and SINR $\gamma_{k,j}^t$. Specifically, we utilize the validated semantic parameter mapping from [42] to ensure that the relationship between v , SINR, and semantic similarity δ accurately reflects the performance of a physical DeepSC implementation. Furthermore, the proposed two-tier hybrid framework exhibits inherent robustness: the PPO agent can adaptively tune the semantic symbol count $v_{k,j}^t$ based on real-time feedback, while the LP solver provides instantaneous optimal offloading ratios to compensate for the computational fluctuations, ensuring the system remains within the latency thresholds. Each vehicle user adaptively partitions its task into three parts: local execution, V2I semantic offloading to the RSU, and V2V semantic offloading to the SV. For the target selection of V2V task offloading, a vehicle user intends to select the nearest service vehicle.



Figure 2. System model diagram.

3.2. RIS-Based Semantic Offloading Model

In our proposed model, we assume that within time slot Δt , the position of the k -th vehicle user is (x_k^t, y_k^t, z_k^t) , the position of the j -th edge node is (x_j^t, y_j^t, z_j^t) , and the fixed-position RIS is located at $(\text{RIS}_x, \text{RIS}_y, \text{RIS}_z)$. Thus, we can obtain the channel gain between the k -th vehicle user and the j -th edge node:

$$\mathbf{h}_{k,j}^t = \sqrt{l(d_{k,j}^t)^{-\eta_{k,j}}} \mathbf{g}_{k,j}^t, \quad (1)$$

where l denotes the path loss at reference distance $d_0 = 1$ m; $d_{k,j}^t$ represents the distance between the k -th vehicle and j -th edge node, with $\eta_{k,j}$ as the path loss exponent; and $\mathbf{g}_{k,j}^t$ accounts for the small-scale multipath fading.

To mitigate the severe path loss during signal propagation, we deploy a large RIS on the surface of buildings adjacent to the road, as illustrated in Figure 2, in order to reconstruct the wireless propagation environment through intelligent reflection. Specifically, the RIS consists of a large number of passive reflecting elements. In addition, considering that the links between them are line-of-sight (LoS), the channel between the two can be modeled as a static channel. Therefore, the channel gain between the RIS and the j -th edge node can be obtained as follows:

$$\mathbf{h}_{r,j} = \sqrt{l(d_{r,j})^{-\eta_{r,j}}} \sqrt{\frac{K}{1+K}} \mathbf{h}_{r,j}^{\text{LoS}}, \quad (2)$$

where $d_{r,j}$ is the RIS-to-edge distance, $\eta_{r,j}$ is the path loss exponent, and K is the Rician factor. Since both the RIS and the RSU are fixedly deployed in the urban IoV environment, this ensures the existence of a dominant LoS component. This provides a stable “virtual LoS link” to compensate for the severe multipath fading and occlusion commonly seen in direct V2I and V2V links. The LoS component $\mathbf{h}_{r,j}^{\text{LoS}}$ is defined as follows:

$$\mathbf{h}_{r,j}^{\text{LoS}} = \left[1, e^{-j\frac{2\pi}{\lambda} D_r \sin(\theta_{r,j})}, \dots, e^{-j\frac{2\pi}{\lambda} (N-1) D_r \sin(\theta_{r,j})} \right]^T, \quad (3)$$

with carrier wavelength λ , element spacing D_r , departure angle $\theta_{r,j}$, and total number of reflecting elements N . Similarly, the channel gain between user k and the RIS is:

$$\mathbf{h}_{k,r}^t = \sqrt{l(d_{k,r}^t)^{-\eta_{k,r}}} \sqrt{\frac{K}{1+K}} \mathbf{h}_{k,r}^{t,\text{LoS}} \quad \forall k \in \mathcal{K}, \forall t \in \mathcal{T}. \quad (4)$$

Here, $d_{k,r}^t$ and $\eta_{k,r}$ denote the time-varying user-RIS distance and path loss exponent, respectively. The LoS component is:

$$\mathbf{h}_{k,r}^{t,\text{LoS}} = \left[1, e^{-j\frac{2\pi}{\lambda} D_r \sin(\theta_{k,r}^t)}, \dots, e^{-j\frac{2\pi}{\lambda} (N-1) D_r \sin(\theta_{k,r}^t)} \right]^T, \quad (5)$$

where $\theta_{k,r}^t$ denotes the angle of arrival (AoA) of the signal from the k -th vehicle user to the RIS at time slot t .

Based on the above RIS-assisted communication propagation characteristics, we further analyze the communication quality of the system. Assume that the system spectrum is divided into B orthogonal resource blocks (RBs), denoted by the set $\mathcal{B} = \{1, 2, \dots, B\}$. Vehicle-to-roadside link resources are scarce and show frequency selectivity. As a result, multiple links may reuse the same RB and cause co-channel interference. For any communication link (k, j) , if it is allocated to b -th resource block for transmission, i.e., $b \in \mathcal{B}$, SINR between the k -th vehicle user and the j -th edge node at time slot t can be expressed as follows:

$$\gamma_{k,j}^{t,(b)} = \frac{p_{k,j} \left| \mathbf{h}_{k,j}^t + \left(\mathbf{h}_{r,j}^{(m)} \right)^H \Theta^t \mathbf{h}_{k,r}^t \right|^2}{\sum_{(k',j') \neq (k,j)} \alpha_{k',j'}^{(b)} p_{k',j'} \left| \mathbf{h}_{k',j'}^t \right|^2 + \sigma^2}. \quad (6)$$

Here, $p_{k,j}$ denotes the transmission power from the k -th vehicle user to the j -th edge node, and the term $\mathbf{h}_{k,j}^t + \left(\mathbf{h}_{r,j}^{(m)} \right)^H \Theta^t \mathbf{h}_{k,r}^t$ represents the signal component reflected by the RIS. We define the diagonal phase shift matrix of the RIS as $\Theta^t = \text{diag}(\mu_1 e^{j\theta_1^t}, \mu_2 e^{j\theta_2^t}, \dots, \mu_N e^{j\theta_N^t})$ for all $n \in [1, N]$, where $\mu_n \in [0, 1]$ represents the transmission coefficient amplitude and phase shift for each element n . Due to hardware constraints, the phase shift can only be selected from a finite discrete set $\theta_n^t \in \Phi = \left\{ 0, \frac{2\pi}{2^q}, \dots, \frac{2\pi(2^q-1)}{2^q} \right\}$, where q is the control bit for the phase shift discretization degree. $\sum_{(k',j') \neq (k,j)} \alpha_{k',j'}^{(b)} p_{k',j'} \left| \mathbf{h}_{k',j'}^t \right|^2$ represents the aggregated interference from other links (k', j') sharing the same resource block b , where $\alpha_{k',j'}^{(b)} \in \{0, 1\}$ is a binary spectrum reuse variable (set to 1 if and only if link (k', j') occupies resource block b). σ^2 is the additive white Gaussian noise power.

Based on the analysis of the physical layer channel and SINR, we further construct a semantic-aware transmission model. This paper adopts DeepSC as the underlying text transmission framework. At the transmitter, the k -th vehicle user encodes its text task into a semantic symbol sequence using a DeepSC encoder and transfers it to the edge node j (RSU or SV) over a wireless channel; the receiver decodes it using a DeepSC receiver to recover the original text information. Unlike traditional communication systems focusing on bit transmission rate, our system focuses on the effective transmission efficiency of semantic information. Therefore, we adopt the semantic rate as the core performance metric, with the unit of semantic units per second (suts/s). At time slot t , the semantic rate $R_{k,j}^t$ between the k -th user and edge node j is defined as follows:

$$R_{k,j}^t = \frac{W \cdot I_k}{L_k \cdot v_{k,j}^t} \delta_{k,j}^t \quad (7)$$

where W , I_k , and L_k denote the bandwidth, average semantic information, and word length per sentence, respectively. The semantic similarity $\delta_{k,j}^t \in [0, 1]$ is a non-linear function of the number of semantic symbols $v_{k,j}^t$ and SINR $\gamma_{k,j}^{t,(b)}$, expressed as $\delta_{k,j}^t = f(v_{k,j}^t, \gamma_{k,j}^{t,(b)})$. This mapping relationship $f(\cdot)$ is derived from the pre-trained DeepSC model. By adopting the validated semantic parameter table from [42], we reuse the verified and reliable data, thus reducing the repetition of experimental work.

The semantic rate formula represents the effective throughput of information by coupling the physical layer bandwidth W with the application layer's semantic richness. Unlike bit-rate, which treats all bits equally, this model prioritizes the 'semantic units'. By optimizing $v_{k,j}^t$ as a decision variable, the system can maintain a high semantic similarity δ_{th} even when RIS phase shifts Θ^t cannot fully mitigate deep fades.

3.3. Computational Model

As illustrated in the three-path framework, the total task volume D_k generated by vehicle k is split into three parallel components with offloading coefficients ρ_k^{loc} , ρ_k^{RSU} , and ρ_k^{SV} . These paths operate concurrently, and the end-to-end latency is determined by the bottleneck path. Each path's delay is modeled separately to reflect the distinct communication and computation characteristics of local, V2I, and V2V resources.

- Local computing: The local processing delay of tasks depends on the vehicle's computing capability. Denote f_k^{loc} as local CPU frequency of k -th vehicle user, C as CPU cycles per bit of data, and D_k as original bit data size. The local computing delay T_k^{loc} is expressed as follows:

$$T_k^{loc} = \frac{\rho_k^{loc} D_k C}{f_k^{loc}}. \quad (8)$$

- RSU-assisted computing: Task offloading to RSU involves two delays: semantic data transmission delay and RSU computing delay. In the transmission phase, DeepSC converts raw data to semantic symbol streams. A conversion factor H maps raw data D_k to semantic task queue length (sentence count) $Q_k = D_k/H$. The V2I link transmission delay $T_{k,0}^{tr}$ is:

$$T_{k,0}^{tr} = \frac{\rho_k^{RSU} Q_k I_k}{R_{k,0}^t} = \frac{\rho_k^{RSU} Q_k \cdot (v_{k,0}^t L_k)}{W \cdot \delta_{k,0}^t}. \quad (9)$$

In the computing phase, RSU evenly allocates total computing capacity F^{RSU} , U_0 connected vehicles, and thus the RSU processing delay $T_{k,0}^{com}$ is expressed as follows:

$$T_{k,0}^{com} = \frac{\rho_k^{RSU} D_k C}{F^{RSU}/U_0}. \quad (10)$$

Note that the return delay is neglected due to small computation results.

- SV-assisted computing: For V2V offloading, vehicles collaborate with the nearest SV. The transmission delay $T_{k,j}^{tr}$ (where $j \neq 0$) is:

$$T_{k,j}^{tr} = \frac{\rho_k^{SV} Q_k I_k}{R_{k,j}^t} = \frac{\rho_k^{SV} Q_k \cdot (v_{k,j}^t L_k)}{W \cdot \delta_{k,j}^t}, j \neq 0. \quad (11)$$

For computing resource allocation: j -th SV serves U_j vehicles by sharing its computing capacity F_j^{SV} . A service threshold $U_{\max} \approx \lfloor K/J \rfloor$ is set to avoid overload while exces-

sive requests are offloaded to the next nearest SV. Under normal load, SV computing delay $T_{k,j}^{\text{com}}$ (where $j \neq 0$) is:

$$T_{k,j}^{\text{com}} = \frac{\rho_k^{\text{SV}} D_k C}{F_j^{\text{SV}} / U_j}, j \neq 0. \quad (12)$$

Return delay is neglected, consistent with V2I offloading.

Since local execution, V2I offloading and V2V offloading run in parallel, the total task delay of vehicle k is the maximum of the three paths:

$$T_k^{\text{Total}} = \max \{ T_k^{\text{loc}}, T_{k,0}^{\text{tr}} + T_{k,0}^{\text{com}}, T_{k,j}^{\text{tr}} + T_{k,j}^{\text{com}} \}. \quad (13)$$

3.4. Problem Formulation

The aim of this is to improve VEC system efficiency by minimizing the overall delay. Specifically, we jointly optimize the offloading ratios, the number of semantic symbols, and the RIS phase shifts. The resulting optimization problem, $\mathcal{P}1$, is defined as follows:

$$\begin{aligned} \mathcal{P}1 : \quad & \min_{\rho_k^{\text{loc}}, \rho_k^{\text{RSU}}, \rho_k^{\text{SV}}, v_{k,j}^t, \theta_n^t} \sum_{k \in \mathcal{K}} T_k^{\text{Total}} \\ \text{s.t.} \quad & \text{C1 : } \rho_k^{\text{loc}}, \rho_k^{\text{RSU}}, \rho_k^{\text{SV}} \in [0, 1], \forall k \\ & \text{C2 : } \rho_k^{\text{loc}} + \rho_k^{\text{RSU}} + \rho_k^{\text{SV}} = 1, \forall k \\ & \text{C3 : } \delta_{k,j}^t \geq \delta_{\text{th}}, \forall k, j, t \\ & \text{C4 : } \theta_n^t \in \Phi, \forall n \\ & \text{C5 : } v_{k,j}^t \in \{1, \dots, v_{k,j}^{\text{max}}\} \\ & \text{C6 : } T_k^{\text{loc}} \leq T_k^{\text{max}}, \forall k \\ & \text{C7 : } T_{k,0}^{\text{tr}} + T_{k,0}^{\text{com}} \leq T_k^{\text{max}}, \forall k \\ & \text{C8 : } T_{k,j}^{\text{tr}} + T_{k,j}^{\text{com}} \leq T_k^{\text{max}}, \forall k, j, j \neq 0. \end{aligned} \quad (14)$$

In Problem $\mathcal{P}1$, Constraint C1 defines the range of task offloading proportions, while constraint C2 ensures the task is fully processed. Constraint C3 guarantees the semantic similarity of all links (k, j) is no less than the threshold. Constraint C4 restricts RIS phase shifts to the feasible configuration set. Constraint C5 denotes that the number of semantic symbols is a finite discrete value. Finally, constraints C6, C7 and C8 ensure each task part meets the delay requirement, where T_k^{max} denotes the maximum delay threshold.

4. Problem Formulation and Solution

Since problem $\mathcal{P}1$ involves both continuous offloading ratio variables and the number of semantic symbols, as well as RIS phase shift variables, the objective function is non-convex. If DRL is adopted to optimize them simultaneously, it will lead to an excessively large action space and state space, thereby prolonging the training time and increasing the uncertainty of the training results. To address this, this paper adopts a two-layer collaborative hybrid solution framework. The upper layer utilizes the PPO algorithm to learn the discrete policies for RIS phase shift selection and the number of semantic symbols selection. The lower layer employs LP to achieve the optimal offloading ratio given the upper-layer actions, thus minimizing the end-to-end latency under the given RIS phase shifts and the number of semantic symbols choices. Specifically, to effectively solve $\mathcal{P}1$, we decouple it into two mutually nested sub-problems and process them using this two-layer framework. First, given the number of semantic symbols and RIS phase shift variables, the task offloading ratios are optimized to minimize the delay. Then, based on the feedback

of the optimal task offloading ratios, the optimal strategies for the number of semantic symbols and RIS phase shifts are searched.

4.1. PPO Algorithm

We employ PPO to optimize the discrete variables, i.e., the number of semantic symbols $v_{k,j}^t$ and phase shifts θ_n^t . With a given continuous offloading ratio ρ , $\mathcal{P}1$ is reformulated as sub-problem $\mathcal{P}2$:

$$\begin{aligned} \mathcal{P}2: \quad & \min_{v_{k,j}^t, \theta_n^t} \sum_{k \in \mathcal{K}} T_k^{\text{total}} \\ \text{s.t.} \quad & C3, C4, C5, C7, C8. \end{aligned} \quad (15)$$

This process is modeled as a Markov Decision Process (MDP), where the RSU agent interacts with the environment via state s , action a , and reward r .

(1) The state at slot t is:

$$s_t = \{L_t, \theta_{k,r}^t, \theta_n^t, \gamma_{k,j}^{t,(b)}\}, \quad (16)$$

where L_t denotes the coordinate positions of all vehicles, including the real-time 3D coordinates of vehicle users ($k \in \mathcal{K}$) and SVs ($j \in \mathcal{J}$); $\theta_{k,r}^t$ represents the AoA of the signal from the k -th vehicle user to the RIS at time slot t ; θ_n^t denotes the phase shift coefficients of the RIS; and $\gamma_{k,j}^{t,(b)}$ is the SINR of the b -th RB at the current time slot.

(2) Action space: The agent's action space is:

$$a_t = \{v_{k,j}^t, \theta_n^t\}, \quad (17)$$

where $v_{k,j}^t$ is the number of semantic symbols decision and θ_n^t is the discrete phase shifts decision.

(3) Reward function: Upon executing action a_t , the lower-layer LP solver computes the minimum system latency T_k^{total} . The reward is defined as the negative total latency:

$$r_t = -T_k^{\text{total}}. \quad (18)$$

We model the policy $\pi_\psi(a_t|s_t)$ as a Multivariate Gaussian Distribution to handle the high-dimensional action space efficiently. Although the optimization variables (RIS phase shifts and the number of semantic symbols) are inherently discrete, directly searching in a high-dimensional discrete space is computationally prohibitive. Therefore, we adopt a continuous relaxation strategy during the training phase. Specifically, the Actor network outputs continuous actions within the range $[-1, 1]$, which are subsequently mapped to valid discrete values. For a given state s_t , the network outputs the mean vector $\mu(s_t)$ and maintains a standard deviation vector σ . The continuous action sampling formula is:

$$a_t = \tanh(\mu(s_t) + \sigma \odot z_t), \quad z_t \sim \mathcal{N}(0, I), \quad (19)$$

where $\mathcal{N}(0, I)$ denotes the standard normal distribution, and $\odot$ represents element-wise multiplication. The tanh function constrains the raw actions a_t to the interval $[-1, 1]$. To bridge the gap between the continuous output and discrete physical constraints, we perform the following mapping operations. For the semantic symbol count $v_{k,j}^t$, the continuous action $a_v \in [-1, 1]$ is linearly scaled and mapped to the nearest integer:

$$v_{k,j}^t = \text{round}\left(\frac{a_v + 1}{2} \cdot (v_{\max} - 1)\right) + 1. \quad (20)$$

Similarly, for the RIS phase shift θ_n^t , the action $a_\theta \in [-1, 1]$ is mapped to the nearest feasible discrete phase in set Φ :

$$\theta_n^t = \arg \min_{\phi \in \Phi} |\phi - \pi(a_\theta + 1)|. \quad (21)$$

The adoption of continuous relaxation followed by a rounding operation provides a practical balance between training efficiency and hardware constraints. By enabling smooth gradient updates in a continuous action space, this strategy effectively circumvents the instability inherent in high-dimensional discrete optimization. This hybrid approach leverages the efficient gradient updates of continuous PPO while ensuring that the executed actions strictly satisfy the discrete constraints of the physical system. The Critic network $V_\phi(s_t)$ (parameterized by ϕ) is used to estimate the value of the current state to assist in policy update. To reduce the variance in gradient estimation, the algorithm uses the advantage function A_t to evaluate the relative superiority of actions.

The advantage function A_t is calculated as the difference between the discounted cumulative return G_t and the state value estimation:

$$G_t = \sum_{l=0}^{T-t-1} \gamma^l r_{t+l}, \quad (22)$$

$$A_t = G_t - V_\phi(s_t), \quad (23)$$

where γ is the discount factor. In practical calculations, A_t is usually standardized to further improve the numerical stability of training. The core of the PPO algorithm lies in limiting the difference between the old and new policies. The probability ratio $r_t(\psi)$ is defined as the ratio of the probability of the new policy π_ψ to the old policy $\pi_{\psi_{old}}$ under the current action: $r_t(\psi) = \frac{\pi_\psi(a_t|s_t)}{\pi_{\psi_{old}}(a_t|s_t)}$. To prevent performance collapse caused by excessively large policy update steps, PPO introduces a clipping operation. The loss function $L^{CLIP}(\psi)$ of the Actor is defined as follows:

$$L^{CLIP}(\psi) = \mathbb{E}_t[\min(r_t(\psi)A_t, \text{clip}(r_t(\psi), 1 - \epsilon, 1 + \epsilon)A_t)]. \quad (24)$$

In the formula, the $\text{clip}(\cdot)$ function restricts the probability ratio to $[1 - \epsilon, 1 + \epsilon]$. The Critic network is updated by minimizing the Mean Square Error (MSE) between the predicted value and the true return:

$$L^{VF}(\phi) = \mathbb{E}_t[(V_\phi(s_t) - G_t)^2]. \quad (25)$$

Considering the policy improvement, value fitting, and exploration ability comprehensively, the total loss function L^{total} is constructed as follows:

$$L^{total} = -L^{CLIP}(\psi) + c_1 L^{VF}(\phi) - c_2 S[\pi_\psi](s_t), \quad (26)$$

where c_1 is the value loss coefficient; and c_2 is the entropy coefficient; $S[\pi_\psi]$ denotes the entropy of the policy distribution, which encourages the agent to maintain high randomness in the early stage of training and prevents premature convergence to local optima. Finally, by using the Adam optimizer to minimize the above total loss function, the algorithm can end-to-end learn the optimal RIS configuration and traffic steering strategy.

4.2. LP Algorithm

After the PPO agent outputs action a_t , semantic parameter and RIS properties are fixed, and V2I, V2V link channels, transmission rates become constants; this removes the non-linear coupling between RIS configuration and offloading proportion in P1.

We then use LP to optimize the continuous offloading proportion ρ for minimal delay, transforming $\mathcal{P}1$ into $\mathcal{P}3$:

$$\begin{aligned} \mathbf{P3}: \quad & \min_{\rho} \sum_{k \in \mathcal{K}} T_k^{\text{total}} \\ \text{s.t.} \quad & C1, C2, C6, C7, C8. \end{aligned} \quad (27)$$

Based on the system's total latency, we first define the unit data processing time coefficients for the three paths: $\mu_k^{\text{loc}} = \frac{D_k C}{f_k^{\text{loc}}}$, $\mu_k^{\text{RSU}} = \frac{Q_k I_k}{R_{k,0}(\mathcal{A}_t)} + \frac{D_k C}{F^{\text{RSU}}/U_0}$, $\mu_k^{\text{SV}} = \frac{Q_k I_k}{R_{k,j}(\mathcal{A}_t)} + \frac{D_k C}{F_j^{\text{SV}}/U_j}$. Since constraints C6, C7, C8 relate to ρ , they can be simplified as follows:

$$\begin{aligned} \rho_k^{\text{loc}} \mu_k^{\text{loc}} - T_k^{\text{total}} &\leq 0, \\ \rho_k^{\text{RSU}} \mu_k^{\text{RSU}} - T_k^{\text{total}} &\leq 0, \\ \rho_k^{\text{SV}} \mu_k^{\text{SV}} - T_k^{\text{total}} &\leq 0. \end{aligned} \quad (28)$$

To solve $\mathcal{P}3$ efficiently with mature numerical optimization algorithms, we convert its physical model into the standard matrix-form optimization problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A} \mathbf{x} \leq \mathbf{b}, \quad \mathbf{A}_{\text{eq}} \mathbf{x} = \mathbf{b}_{\text{eq}}, \quad \mathbf{lb} \leq \mathbf{x} \leq \mathbf{ub}. \end{aligned} \quad (29)$$

We define the augmented optimization vector $\mathbf{x} \in \mathbb{R}^4$, which includes three offloading proportion variables and one delay auxiliary variable $\mathbf{x} = [\rho_k^{\text{loc}}, \rho_k^{\text{RSU}}, \rho_k^{\text{SV}}, T_k]^T$. The goal is to minimize the fourth term of $\mathbf{x}$. Therefore, the coefficient vector $\mathbf{c}$ is set as $\mathbf{c} = [0, 0, 0, 1]^T$.

Based on the simplified expressions of constraints C6, C7, C8, we construct the inequality constraint matrix $\mathbf{A}$ and vector $\mathbf{b}$:

$$\mathbf{A} = \begin{bmatrix} \mu_k^{\text{loc}} & 0 & 0 & -1 \\ 0 & \mu_k^{\text{RSU}} & 0 & -1 \\ 0 & 0 & \mu_k^{\text{SV}} & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (30)$$

Constraint C2 requires the sum of offloading proportions to be 1, corresponding to the equality constraint matrix and vector:

$$\mathbf{A}_{\text{eq}} = [1, 1, 1, 0], \quad \mathbf{b}_{\text{eq}} = [1]. \quad (31)$$

Constraint C1 and the maximum delay threshold T_k^{max} form the variable bounds:

$$\mathbf{lb} = [0, 0, 0, 0]^T, \quad \mathbf{ub} = [1, 1, 1, T_k^{\text{max}}]^T. \quad (32)$$

Through the above transformation, the problem is reformulated as a standard convex optimization problem, which is solved using the interior-point method to obtain the globally optimal solution $\mathbf{x}^*$. The minimum system total delay $T_{\text{sys}} = \sum_{k \in \mathcal{K}} T_k^{\text{Total}}$ is calculated based on $\mathbf{x}^*$, which is fed back to the upper-layer PPO algorithm as a reward signal to guide the update of the number of semantic symbols and RIS configuration policy.

The lower-layer LP solver plays a critical role in mitigating the impact of task size fluctuations and queueing effects. For each time slot t , the LP optimizer receives the actual task demand D_k^t and the reshaped channel gains $h_{k,j}^t$ from the upper-layer RIS configuration. By formulating the sub-problem as a standard linear program with $4K$ variables, including auxiliary delay variables T_k , the solver yields the optimal offloading ratios $\rho_k^{\text{loc}}, \rho_k^{\text{RSU}}, \rho_k^{\text{SV}}$ instantaneously. If a vehicle experiences a sudden increase in task volume or a signal blockage due to mobility, the LP solver automatically re-balances the workload among the

three available computational paths. This dynamic compensation mechanism ensures that even under non-ideal conditions such as service vehicle movement or transient traffic bursts, the end-to-end delay is minimized and kept within the safety-critical threshold T_k^{max} .

4.3. Algorithm Implementation and Training Phase

Based on the constructed loss function and optimization objectives, we designed a complete training process for the PPO algorithm to achieve the joint optimization of RIS configuration, the number of semantic symbols and task scheduling.

The training process follows the online reinforcement learning paradigm. In the interaction phase, the Actor network first outputs the action movement towards RIS phase shifts and the number of semantic symbols at current state. Subsequently, the optimizer within the environment reconstructs the channel environment for this action decision and solves the LP problem to obtain the optimal relay decision and minimum transmission delay, which are then fed back as reward signals. In the model update phase, trajectory data is sampled periodically from the experience replay buffer to calculate the advantage function, and the PPO clipped objective function constraint strategy is used for parameter update. This approach not only ensures the training stability but also coordinates the optimization of physical layer resource allocation and hyperparameter selection in the semantic layer, thus ultimately maximizing the long-term cumulative reward of the system. The complete execution steps of the two-layer collaborative optimization solution framework proposed in this paper are presented in Algorithm 1.

Algorithm 1 PPO Algorithm Training Phase

Require: Maximum training episodes E_{max} , Maximum time steps per episode T_{max} , Network parameter update interval T_{update} ; PPO hyperparameters: learning rate α , clipping coefficient ϵ , discount factor γ

Ensure: Trained Actor network parameters ψ^* , Trained Critic network parameters ϕ^*

- 1: Randomly initialize the weight parameters of Actor network and Critic network.
 - 2: Initialize the experience replay buffer $\mathcal{D}$ for storing trajectory data. Initialize the global time step counter $t_{step} = 0$.
 - 3: **for** $episode = 1$ **to** E_{max} **do**
 - 4: Reset the environment and observe the initial state s_1 .
 - 5: **for** $t = 1$ **to** T_{max} **do**
 - 6: Input the current state s_t into the Actor network. Generate the joint action a_t via Equation (19).
 - 7: Call the LP solver to obtain the optimal task offloading ratio p^* and the corresponding minimum system delay t_{total} .
 - 8: Calculate the reward r_t and get the next state s_{t+1} .
 - 9: Store the transition tuple (s_t, a_t, r_t, s_{t+1}) in the experience buffer $\mathcal{D}$.
 - 10: Update the current state $s_t \leftarrow s_{t+1}$ and increment the time step counter t_{step} .
 - 11: **if** $t_{step} > T_{update}$ **then**
 - 12: Based on the buffer $\mathcal{D}$, calculate the discounted cumulative return and advantage function via Equations (22) and (23).
 - 13: **for** $k = 1$ **to** K **do**
 - 14: Calculate the probability ratio of the new and old policies.
 - 15: Calculate the total loss function via Equation (26).
 - 16: Minimize the total loss using Adam optimizer and update the parameters of Actor and Critic networks.
 - 17: **end for**
 - 18: **end if**
 - 19: **end for**
 - 20: **end for**
-

4.4. Testing Phase

In the testing phase, operations such as network parameter adjustments and gradient updates are not required. Instead, the trained model is directly employed to obtain the optimal scheme for RIS configurations and task offloading ratios. Note that each test round consists of 200 test time slots, and the final result is derived from the average performance across all time slots within that round. The pseudo-code for the testing phase is shown in Algorithm 2.

Algorithm 2 PPO Algorithm Testing Phase

Require: Trained Actor network parameters ψ^* , number of test episodes E_{test} , maximum time steps per episode T_{max}

Ensure: Average system delay

```

1: for episode = 1 to  $E_{test}$  do
2:   Reset the environment and obtain the initial state  $s_1$ .
3:   for  $t = 1$  to  $T_{max}$  do
4:     Input the current state  $s_t$  into the Actor network.
5:     Turn off exploration noise and output the deterministic action  $a_t = \tanh(\mu(s_t))$ .
6:     Parse the action  $a_t$  to get RIS phase shifts and the number of semantic symbols.
7:     Solve Problem P3 to obtain the optimal unloading ratio  $\rho^*$  and the minimum system delay  $T_{sys}^*$ .
8:     Observe the next state  $s_{t+1}$ .
9:     Update the state  $s_t \leftarrow s_{t+1}$ .
10:   end for
11: end for

```

4.5. Complexity Analysis

To substantiate the practical feasibility of the proposed two-layer optimization framework, a rigorous analysis of the computational complexity is indispensable. The overall system complexity per time slot is characterized by the synergy between the upper-layer PPO inference and the lower-layer LP optimization.

The computational overhead of the upper-layer decision-making predominantly stems from the forward propagation of the Actor network during the online inference phase. Assuming the Actor network comprises L fully connected layers with n_l neurons in the l -th layer, the complexity of processing a single state input is defined as $\mathcal{O}\left(\sum_{l=1}^{L-1} n_l \cdot n_{l+1}\right)$. We denote this neural network inference complexity as $NN_{inference}$. Given that the network architecture—characterized by the number of layers and neurons—is fixed upon the completion of offline training and is typically maintained at a compact scale, $NN_{inference}$ represents a constant and minimal computational demand. This ensures that the vehicle users, even with limited onboard processing power, can generate optimal RIS phase shifts θ_n^t and semantic symbol configurations $v_{k,j}^t$ in near real time.

Following the PPO-based discrete decision-making, lower-layer optimization involves solving a standard linear programming problem to determine the continuous offloading ratios ρ_k^{loc} , ρ_k^{RSU} , and ρ_k^{SV} . By formulating this sub-problem as a standard LP form with $4K$ variables (including three offloading proportions and one auxiliary delay variable for each of the K vehicles), we employ the interior-point method for resolution. The computation complexity of this stage is approximately $\mathcal{O}(K^{3.5})$. Consequently, the total computation complexity of the proposed framework for each time step is expressed as $\mathcal{O}(NN_{inference} + K^{3.5})$. This decoupled architecture effectively circumvents the exponential complexity incurred by high-dimensional non-convex joint optimization, thereby guaranteeing the low-latency requirements of safety-critical vehicular communications.

5. Simulation Results and Analysis

Simulations are implemented via Python 3.8 and PyTorch 1.12.1 in an urban scenario where an RIS is deployed to assist vehicles suffering from severe attenuation. The setup comprises 1 RSU at $(-10, 150, 25)$, 1 RIS at $(10, 175, 25)$. Each vehicle's speed is fixed at 20 m/s and assumed to remain constant thereafter, which is used for theoretical modeling and training setup. In actual simulation, vehicle speed remains constant while inter-vehicle spacing correlates with vehicle count. Vehicle users generate a task load of 0.4 Mbit per time slot (generated via Poisson distribution). Key simulation environment parameters are shown in Table 1. These parameters are carefully selected to align with the 5G-V2X communication standards and practical vehicular hardware specifications. Specifically, the semantic communication settings are based on the pre-trained DeepSC model to ensure compatibility, while the computing capacities of RSUs and SVs are set to reflect typical edge server performance in urban IoV scenarios.

Table 1. Simulation environment parameter table.

Parameters	Value	Parameters	Value
$\eta_{r,j}$	2.2	$p_{k,j}$	0.2 W
N	6×6	σ^2	1.44×10^{-10} W
K	15	$\eta_{k,j}$	3.5
$\eta_{k,r}$	2.2	W	360 kHz
C	1000 cycles/bit	I_k	100
f_k^{loc}	2 GHz	F^{RSU}	6 GHz
F_j^{SV}	2 GHz	L_k	20
H	1200 bit	δ_{th}	0.9
α	0.0003 (actor)/0.001 (critic)	γ	0.6
ϵ	0.2	E_{max}	5000

We compare our proposed PPO method with the following methods:

- Genetic Algorithm (GA) [43]: GA simulates natural selection and genetic mechanisms, searching for the global optimal solution within the solution space through selection, crossover, and mutation operations.
- Quantum-behaved Particle Swarm Optimization (QPSO): QPSO is a heuristic benchmark adopting the QPSO algorithm [44]. Compared with the standard Particle Swarm Optimization (PSO) algorithm, QPSO eliminates the velocity vector and utilizes wave functions from quantum mechanics to describe the motion state of particles. It has been proven to possess stronger global search capabilities and fewer control parameters.

Figure 3 illustrates the trend of average cumulative reward of the proposed algorithm during the training process. In this figure, the horizontal axis represents the number of training episodes, while the vertical axis denotes the average system reward achieved within each episode. It can be observed that the training process exhibits a distinct convergence trend, demonstrating favorable convergence characteristics and learning efficiency of PPO algorithm.

Figure 4 depicts the impact of vehicle transmit power on the average total system delay under the experimental setting of 15 vehicles and an RIS size of 6×6 . As illustrated, with the vehicle transmit power increasing from 0.1 W to 0.3 W, the system delay of all algorithms monotonically decreases. This is attributed to the fact that increased transmit power effectively improves the SINR at the receiver, thereby enhancing channel capacity and reducing data transmission latency. Notably, the proposed PPO algorithm significantly outperforms GA and QPSO benchmark algorithms across the entire power range. Specifically, the proposed algorithm achieves a performance improvement of approximately 40%

to 50%. Traditional heuristic algorithms often fluctuate and get trapped in local optima. In contrast, the PPO algorithm generates a smoother and continuously descending curve. This fully validates the robustness of the DRL framework in handling high-dimensional continuous action spaces, as well as its superiority in maximizing system efficacy through precise control of RIS phase shifts and the number of semantic symbols.

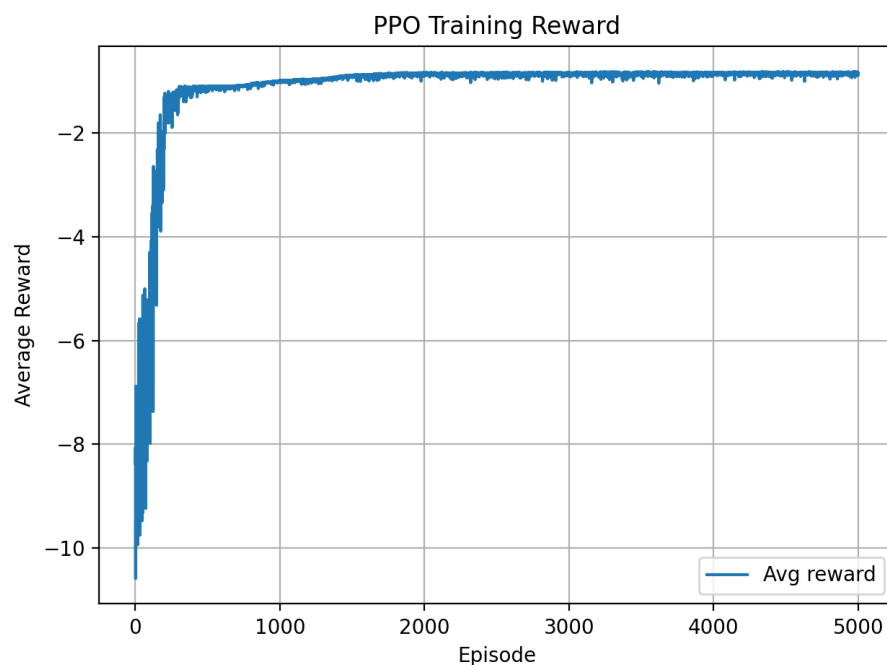


Figure 3. Reward convergence plot of PPO.

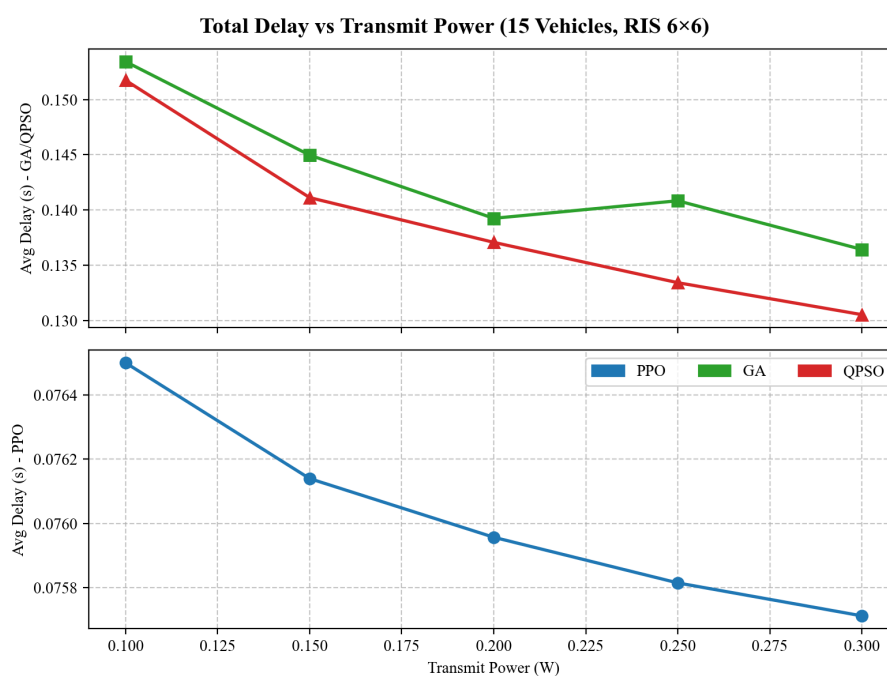


Figure 4. Performance comparison of average delay under varying vehicle transmission powers.

Figure 5 analyzes the communication delay components for V2V and V2I links under different transmit powers. When the transmit power increases from 0.1 W to 0.3 W, SINR is also improved accordingly, thus leading to reduced delays for all algorithms on both links. The proposed PPO algorithm achieves the lowest delay, especially in the V2V link, where it stabilizes at around 0.030 s, providing over 45% performance gain. For the V2I link, PPO also maintains the best performance. These results confirm that the PPO agent effectively mitigates co-channel interference caused by spectrum reuse while enhancing the desired signal strength, thus satisfying the low-latency requirements of IoV.

Figure 6 utilizes boxplots to evaluate algorithm stability with varying vehicle densities. The outliers beyond the Interquartile Range (IQR) highlight instances where algorithms are trapped in poor local optima. When the vehicle count reaches 30, GA and QPSO show elongated boxes and frequent outliers. This indicates significant instability and stochastic fluctuations in high-dimensional spaces. In contrast, PPO maintains a compact distribution with minimal outliers, demonstrating its robustness in preventing extreme latency and ensuring fairness during congestion.

Figure 7 illustrates the system delay under dynamic vehicle densities. As observed, increasing vehicle density exacerbates spectrum competition and co-channel interference, leading to higher delays across all schemes. However, the proposed PPO algorithm demonstrates superior scalability compared to the benchmarks. Even in a congested scenario with 30 vehicles, PPO maintains a low delay with a minimal growth rate, whereas GA and QPSO exhibit rapid performance degradation. This confirms the robustness of the DRL-based approach in terms of resource allocation within interference-limited environments.

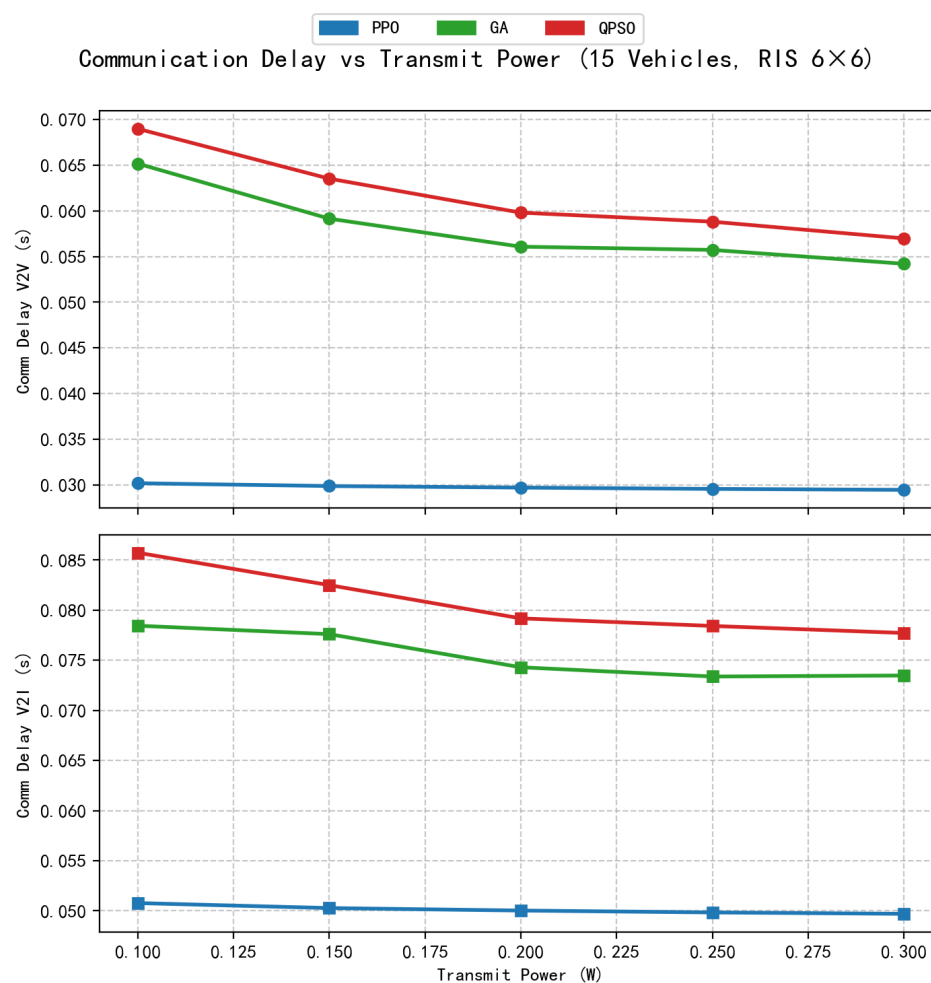


Figure 5. Average delay of V2V and V2I links versus vehicle transmission power.

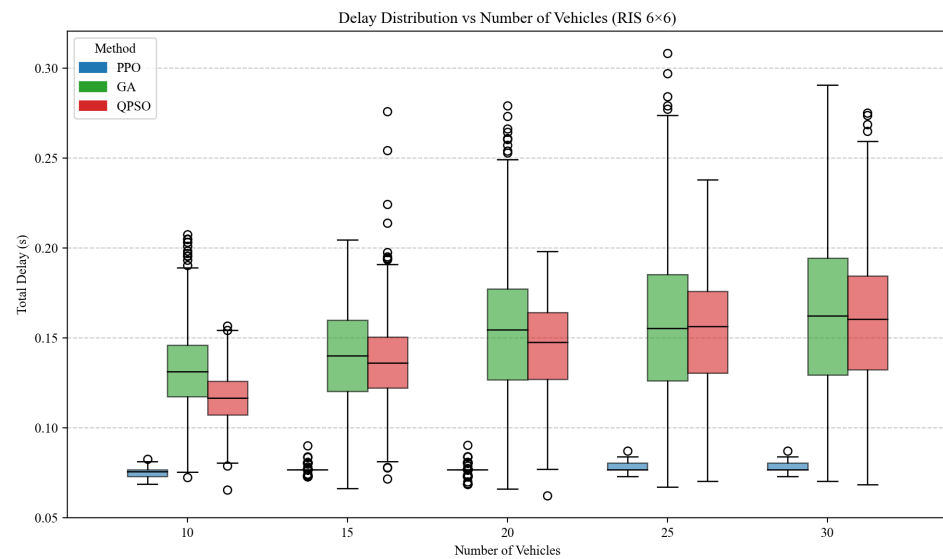


Figure 6. Boxplot of average delay under different numbers of vehicles.

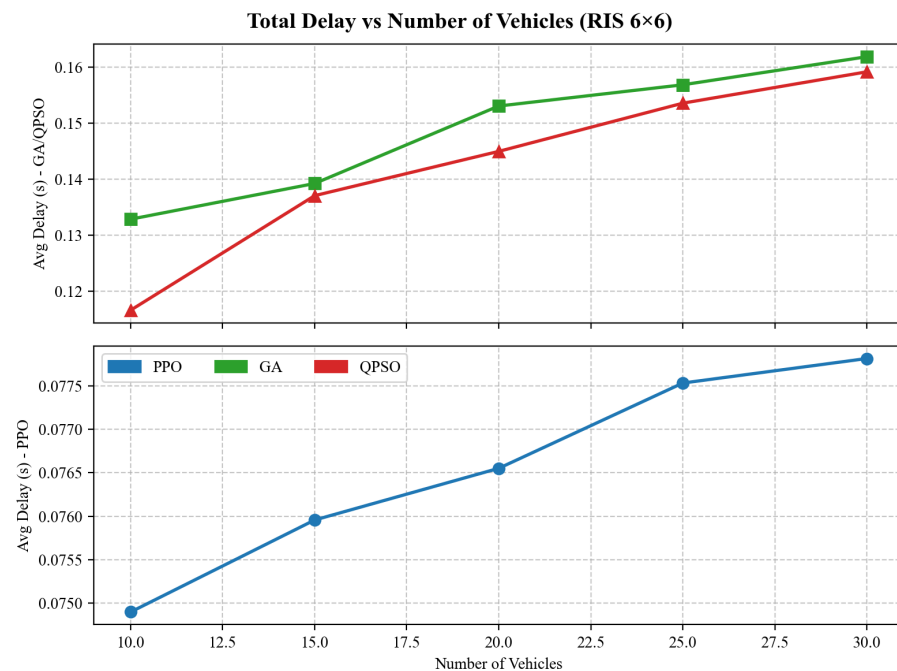


Figure 7. Average delay versus number of vehicles.

Figure 8 illustrates the impact of RIS scale on system delay. As the number of elements increases from 4×4 to 10×10 , system delay decreases for all schemes due to enhanced passive beamforming gains. However, PPO consistently outperforms GA and QPSO, which struggle with the “curse of dimensionality” in large-scale phase optimization. This demonstrates PPO’s superior ability to handle high-dimensional continuous action spaces and fully exploit large-scale RIS hardware.

Figure 9 employs boxplots to evaluate the algorithm stability with varying RIS scales. The expanding Interquartile Ranges (IQRs) and frequent outliers observed in GA and QPSO indicate severe instability and entrapment in local optima as dimensionality increases. Conversely, PPO exhibits compact distributions with minimal outliers, demonstrating its robustness in overcoming the “curse of dimensionality” associated with large-scale RIS optimization.

In summary, the simulation results with varying vehicle densities and RIS scales provide a comprehensive ablation analysis of the proposed framework. The RIS component is essential for reconfiguring the physical channel, as evidenced by the latency reduction observed when the number of reflecting elements increases. The semantic communication module enables semantic compression, which is the key driver behind the 40–50% performance improvement over traditional bit-level optimization schemes [45]. Finally, the two-tier hybrid architecture (PPO-LP) ensures computation efficiency and stability. By delegating the continuous offloading task to the LP solver, the PPO agent can effectively handle the complex discrete search space of RIS phase shifts without succumbing to the curse of dimensionality.

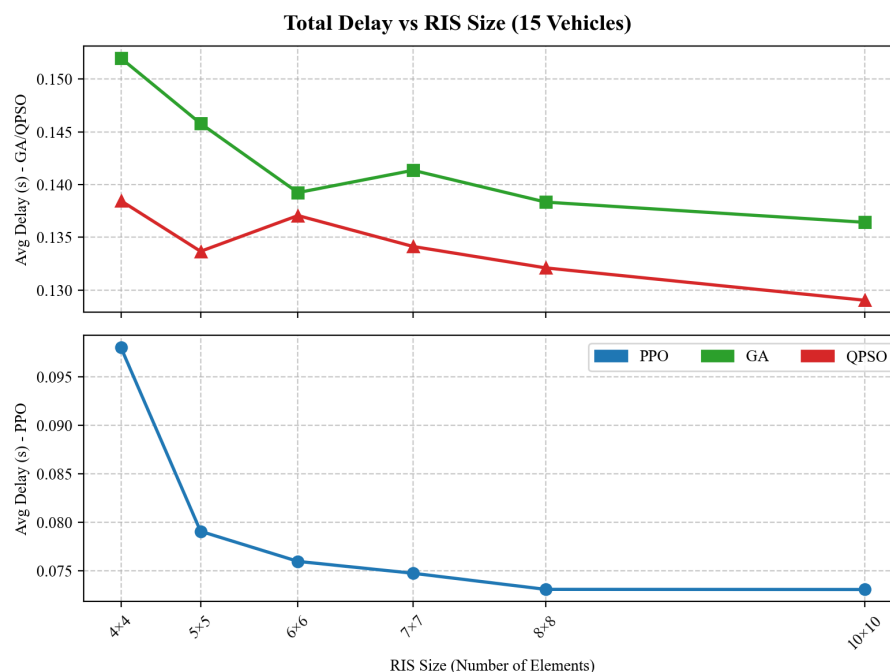


Figure 8. Average delay versus number of RIS-reflecting elements.

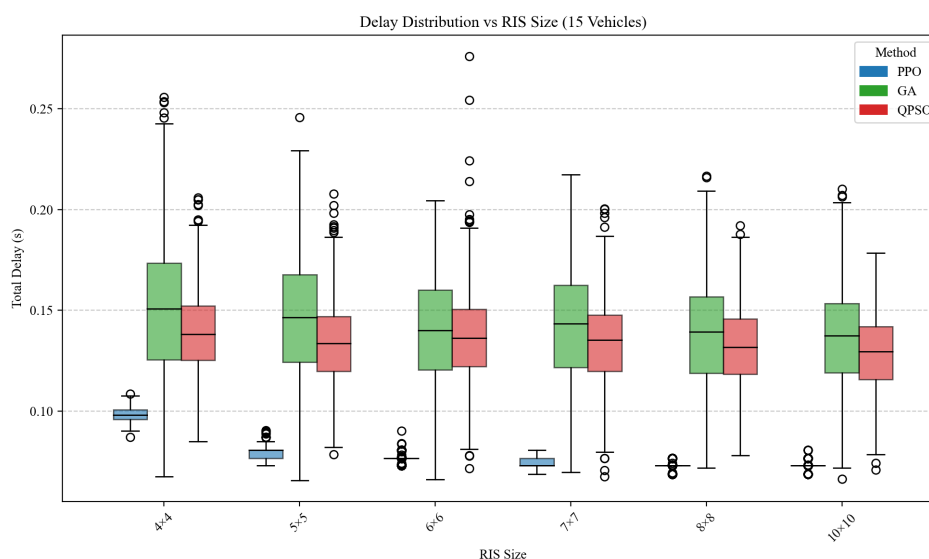


Figure 9. Boxplot of total delay under different numbers of RIS-reflecting elements.

6. Conclusions

In this paper, we have proposed a novel RIS-assisted semantic-aware VEC framework to tackle the conflicting requirements of ultra-low latency and high reliability in complex

dynamic vehicular networks. By integrating link-level RIS enhancement with semantic feature transmission, we constructed a robust three-path offloading model that significantly mitigates channel fading and reduces transmission overhead. To solve the formulated non-convex joint optimization problem involving coupled discrete and continuous variables, we devised a two-tier hybrid solution framework. Specifically, we employed a continuous relaxation-based PPO algorithm to efficiently search for optimal discrete configurations of RIS phase shifts and the number of semantic symbols while utilizing LP to optimize continuous offloading ratios. Extensive simulation results demonstrate that the proposed approach outperforms varying baselines, reducing the average system latency by approximately 40% to 50% as compared to that observed with the GA and QPSO algorithms. Crucially, our method exhibits superior robustness and scalability in high-density scenarios, which effectively overcomes the “curse of dimensionality” in large-scale RIS optimization. Future work will focus on investigating the updating mechanisms of a dynamic semantic knowledge base and exploring multi-RIS cooperative coverage to further enhance system adaptability in wider areas.

Author Contributions: Conceptualization, W.F.; methodology, J.Z. and W.F.; software, J.Z. and W.F.; validation, Q.W. and Q.F.; formal analysis, P.F. and W.F.; investigation, J.Z.; resources, W.F. and Q.W.; data Curation, W.F.; writing—original draft, W.F. and J.Z.; writing—review and editing, W.F., Q.W. and P.F.; visualization, J.Z.; supervision, W.F. and Q.W.; project administration and funding acquisition, Q.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the National Natural Science Foundation of China under Grant 61701197; in part by the Basic Research Program of Jiangsu under Grant BK20252084; in part by the National Key Research and Development Program of China under Grant 2021YFA1000500(4) and in part by the 111 Project under Grant B23008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are disclosed via GitHub (<https://github.com/qiongwu86/PPO-Based-Hybrid-Optimization-for-RIS-Assisted-Semantic-Vehicular-Edge-Computing.git>) (accessed on 23 February 2026).)

Conflicts of Interest: Author Qiang Fan was employed by the company Qualcomm. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Jain, G.; Das, R.; Raj, D.; Mohapatra, H. Exploring Low-Latency Communication Protocols for Time-Sensitive Industrial Automation and Autonomous Vehicle Systems. In Proceedings of the 2025 International Conference on Computing, Intelligence, and Application (CIACON), Durgapur, India, 18–19 July 2025; pp. 1–7. <https://doi.org/10.1109/CIACON65473.2025.11189493>.
2. Huang, Y.F.; Liao, J.W.J.; Huang, W.; Lin, F.L.; Chen, S.K.; Sheu, Y.H. Applications of Fuzzy Inference Machine in Low Latency Vehicle Networks. In Proceedings of the 2025 IEEE International Conference on Consumer Electronics—Taiwan (ICCE-Taiwan), Kaohsiung, Taiwan, 16–18 July 2025; pp. 425–426. <https://doi.org/10.1109/ICCE-Taiwan66881.2025.11207945>.
3. Erbat, M.M.; Schiele, G. A novel dynamic service function chaining to enable Ultra-Reliable Low Latency Communication in NFV. In Proceedings of the 2023 17th International Conference on Telecommunications (ConTEL), Graz, Austria, 11–13 July 2023; pp. 1–8. <https://doi.org/10.1109/ConTEL58387.2023.10198947>.
4. Xie, Y.; Wu, Q.; Fan, P.Y.; Cheng, N.; Chen, W.; Wang, J.Z.; Letaief, K.B. Resource Allocation for Twin Maintenance and Task Processing in Vehicular Edge Computing Network. *IEEE Internet Things J.* **2025**, *12*, 32008–32021. <https://doi.org/10.1109/JIOT.2025.3576582>.
5. Jiao, L.; Wang, P.; Alipour-Fanid, A.; Zeng, H.C.; Zeng, K. Enabling Efficient Blockage-Aware Handover in RIS-Assisted mmWave Cellular Networks. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 2243–2257. <https://doi.org/10.1109/TWC.2021.3110522>.

6. Gu, X.H.; Zhang, G.A.; Ji, Y.C.; Duan, W.; Wen, M.W.; Ding, Z.G.; Ho, P.H. Intelligent Surface Aided D2D-V2X System for Low-Latency and High-Reliability Communications. *IEEE Trans. Veh. Technol.* **2022**, *71*, 11624–11636. <https://doi.org/10.1109/TVT.2022.3189627>.
7. Naaz, F.; Nauman, A.; Khurshaid, T.; Kim, S.-W. Empowering the Vehicular Network with RIS Technology: A State-of-the-Art Review. *Sensors* **2024**, *24*, 337. <https://doi.org/10.3390/s24020337>.
8. Ozcan, Y.U.; Ozdemir, O.; Kurt, G.K. Reconfigurable Intelligent Surfaces for the Connectivity of Autonomous Vehicles. *IEEE Trans. Veh. Technol.* **2021**, *70*, 2508–2513. <https://doi.org/10.1109/TVT.2021.3060667>.
9. Han, S.; Luo, G.; Qu, F.; Lestas, M.; Wang, F.-Y. Empowering Vehicle Connectivity—The SOTA and Future Prospects of Reconfigurable Intelligent Surfaces in Mobile Communications: A Review. *IEEE Sens. J.* **2025**, *25*, 32021–32037. <https://doi.org/10.1109/JSEN.2025.3586678>.
10. Qi, K.; Wu, Q.; Fan, P.; Cheng, N.; Fan, Q.; Wang, J. Reconfigurable Intelligent Surface Assisted VEC Based on Multi-Agent Reinforcement Learning. *IEEE Commun. Lett.* **2024**, *28*, 2427–2431. <https://doi.org/10.1109/LCOMM.2024.3451182>.
11. Xie, H.; Qin, Z.; Li, G.Y.; Juang, B.H. Deep Learning Enabled Semantic Communication Systems. *IEEE Trans. Signal Process.* **2021**, *69*, 2663–2675. <https://doi.org/10.1109/TSP.2021.3071210>.
12. Gündüz, D.; Qin, Z.J.; Aguerri, I.E.; Dhillon, H.S.; Yang, Z.H.; Yener, A.; Wong, K.K.; Chae, C.-B. Beyond Transmitting Bits: Context, Semantics, and Task-Oriented Communications. *IEEE J. Sel. Areas Commun.* **2023**, *41*, 5–41.
13. Shao Z, Wu Q, Fan P, et al. Semantic-Aware Spectrum Sharing in Internet of Vehicles Based on Deep Reinforcement Learning. *IEEE Internet Things J.* **2024**, *11*, 38521–38536. <https://doi.org/10.1109/JIOT.2024.3448538>.
14. Zhao, Z.X.; Yang, Z.H.; Huang, C.W.; Wei, L.; Yang, Q.Q.; Zhong, C.J.; Xu, W.; Zhang, Z.Y. A Joint Communication and Computation Design for Distributed RIS-Assisted Probabilistic Semantic Communication in IIoT. *IEEE Internet Things J.* **2024**, *11*, 26568–26579. <https://doi.org/10.1109/JIOT.2024.3409271>.
15. Du, H.Y.; Wang, J.C.; Niyato, D.; Kang, J.W.; Xiong, Z.H.; Zhang, J.S.; Shen, X.S. Lightweight Wireless Sensing Through RIS and Inverse Semantic Communications. In Proceedings of the 2023 IEEE Wireless Communications and Networking Conference (WCNC), Glasgow, UK, 26–29 March 2023; pp. 1–6. <https://doi.org/10.1109/WCNC55385.2023.10119005>.
16. Jiang, P.W.; Wen, C.-K.; Jin, S.; Li, G.Y. RIS-Enhanced Semantic Communications Adaptive to User Requirements. *IEEE Trans. Commun.* **2024**, *72*, 4134–4148. <https://doi.org/10.1109/TCOMM.2024.3369697>.
17. Kiran, B.R.; Sobh, I.; Talpaert, V.; Mannion, P.; Al Sallab, A.A.; Yogamani, S.; Pérez, P. Deep Reinforcement Learning for Autonomous Driving: A Survey. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 4909–4926. <https://doi.org/10.1109/TITS.2021.3054625>.
18. Li, S.C.; Huang, Q.R.; Chen, H.B.; Jiang, R.H.; Dong, M.X.; Ota, K.; Quek, T.Q.S. DRL-Based Joint Task Offloading and Resource Allocation in Air-Ground Integrated Vehicular Edge Computing Network With Energy Harvesting. *IEEE Trans. Veh. Technol.* **2025**, 1–15. <https://doi.org/10.1109/TVT.2025.3621730>.
19. Gu, X.Y.; Wu, Q.; Fan, P.Y.; Fan, Q.; Cheng, N.; Chen, W.; Letaief, K.B. DRL-Based Resource Allocation for Motion Blur Resistant Federated Self-Supervised Learning in IoV. *IEEE Internet Things J.* **2025**, *12*, 7067–7085. <https://doi.org/10.1109/JIOT.2024.3492326>.
20. Xu, X.; Wu, Q.; Fan, P.Y.; Wang, K.Z.; Cheng, N.; Chen, W.; Letaief, K.B. Enhanced Velocity-Adaptive Scheme: Joint Fair Access and Age of Information Optimization in Vehicular Networks. *IEEE Trans. Mob. Comput.* **2026**, *25*, 3488–3505. <https://doi.org/10.1109/TMC.2025.3617145>.
21. Chu, J.H.; Wu, Q.; Fan, P.Y.; Chen, W.; Wang, K.Z.; Cheng, N.; Letaief, K.B. V2X-Assisted Distributed Computing and Control Framework for Connected and Automated CAVs under Ramp Merging Scenario. *IEEE Trans. Mob. Comput.* **2026**, 1–18. <https://doi.org/10.1109/TMC.2026.3650774>.
22. Wang, B.X.; Tu, D.; Wang, J. DRL-Based Partial Task Offloading for Multiple Vehicles in VEC Networks. In Proceedings of the 2024 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Toronto, ON, Canada, 19–21 June 2024; pp. 1–6. <https://doi.org/10.1109/BMSB62888.2024.10608336>.
23. Wu, S.J.; Yang, L.; Li, J.J.; Guo, H.Z.; Ahmad, I.; Da Costa, D.B.; Jiang, H.B.; Niyato, D. DRL-Based Pricing-Driven for Task Offloading and Dynamic Resource in Vehicle Edge Computing. *IEEE Trans. Mob. Comput.* **2025**, *24*, 10389–10404. <https://doi.org/10.1109/TMC.2025.3569817>.
24. Zeng, J.A.; Kuang, Z.F.; Chen, R.; Liu, A.F. Delay-Sensitive Dependent Tasks Offloading and Resource Allocation in VEC: A Deep-Reinforcement Learning Approach. *IEEE Internet Things J.* **2025**, *12*, 34190–34203. <https://doi.org/10.1109/JIOT.2025.3577449>.
25. Ren, Y.L.; Chen, X.Y.; Guo, S.; Guo, S.Y.; Xiong, A. Blockchain-Based VEC Network Trust Management: A DRL Algorithm for Vehicular Service Offloading and Migration. *IEEE Trans. Veh. Technol.* **2021**, *70*, 8148–8160. <https://doi.org/10.1109/TVT.2021.3092346>.
26. Zhang, M.Y.; Su, Z.; Xu, Q.C.; Qi, Y.H.; Fang, D.F. Energy-Efficient Task Offloading in UAV-RIS-Assisted Mobile Edge Computing with NOMA. In Proceedings of the IEEE INFOCOM 2024—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Vancouver, BC, Canada, 20 May 2024; pp. 1–6. <https://doi.org/10.1109/INFOCOMWKSHPS61880.2024.10620850>.

27. Wahid, A.; Mirza, M.A.; Ahmed, M.; Sheraz, M.; Chuah, T.C.; Lee, I.E.; Khan, W.U. Toward Secure and Scalable Vehicular Edge Computing With Zero-Energy RIS Using DRL. *IEEE Access* **2024**, *12*, 129330–129346. <https://doi.org/10.1109/ACCESS.2024.3457853>.
28. Wang, X.T.; Yi, M.; Chen, M.J.; Liu, Z.Q.; Vasilakos, A.V.; Song, H.H.; Farouk, A. Joint Task Offloading and Resource Allocation in RIS-assisted NOMA-VEC Intent-based Networking. *IEEE Internet Things J.* **2025**, *1*. <https://doi.org/10.1109/JIOT.2025.3620606>.
29. Xie, Y.B.; Shi, L.; Li, Z.H.; Ding, X.; Fan, Y.Q. Efficient Task Offloading in Double Roadside RIS-Assisted Vehicular Edge Computing Networks Using Deep Reinforcement Learning. *IEEE Trans. Veh. Technol.* **2025**, *74*, 11353–11365. <https://doi.org/10.1109/TVT.2025.3543664>.
30. Rzig, I.; Jaafar, W.; Jebalia, M.; Tabbane, S. Energy-Efficient Vehicular Task Offloading Using Multi-Mode MEC and RIS-Equipped Aerial Platforms. *IEEE Open J. Commun. Soc.* **2025**, *6*, 7604–7619. <https://doi.org/10.1109/OJCOMS.2025.3606340>.
31. Xiao, H.; Hu, X.Y.; Wang, W.J.; Su, Z.; Wong, K.-K.; Yang, K. STAR-RIS and UAV Combination in MEC Networks: Simultaneous Task Offloading and Communications. *IEEE Trans. Commun.* **2025**, *73*, 6169–6184. <https://doi.org/10.1109/TCOMM.2025.3535895>.
32. Zhang, H.J.; Wang, H.Y.; Li, Y.B.; Long, K.P.; Leung, V.C.M. Toward Intelligent Resource Allocation on Task-Oriented Semantic Communication. *IEEE Wirel. Commun.* **2023**, *30*, 70–77. <https://doi.org/10.1109/MWC.008.2200504>.
33. Peng, J.C.; Xing, H.L.; Chen, X.Y.; Li, Y.; Cui, Y.H.; Zheng, D.Y.; Ale, L.; Feng, L. Security Enhanced Computation Offloading for Collaborative Inference at Semantic-Communication-Empowered Edge. *IEEE Trans. Mob. Comput.* **2025**, *24*, 8071–8088. <https://doi.org/10.1109/TMC.2025.3555298>.
34. Yang, Y.; Ma, W.Q.; Sun, W.; He, J.H.; Fu, Y.R.; Yuen, C.; Zhang, Y. Diffusion-Based Multi-Agent Reinforcement Learning for Semantic Vehicular Edge Computing. *IEEE Trans. Serv. Comput.* **2025**, *18*, 3668–3681. <https://doi.org/10.1109/TSC.2025.3618082>.
35. Du, B.X.; Du, H.Y.; Niyato, D.; Li, R.D. Task-Oriented Semantic Communication in Large Multimodal Models-Based Vehicle Networks. *IEEE Trans. Mob. Comput.* **2025**, *24*, 9822–9836. <https://doi.org/10.1109/TMC.2025.3564543>.
36. Zheng, G.H.; Ni, Q.; Navaie, K.; Pervaiz, H.; Zarakovitis, C. A Distributed Learning Architecture for Semantic Communication in Autonomous Driving Networks for Task Offloading. *IEEE Commun. Mag.* **2023**, *61*, 64–68. <https://doi.org/10.1109/MCOM.002.2200765>.
37. Chen, X.; Feng, D.; Jiang, W.; Luo, Q.; Chen, G.; Sun, Y. Online Multi-Task Offloading for Semantic-Aware Edge Computing Systems. *arXiv* **2025**, arXiv:2407.11018.
38. Mao, S.; Yuen, C.; Liu, L.; Xiao, M.; Yu, S.; Zhang, N. RIS-Enhanced Semantic-Aware Sensing, Communication, Computation, and Control for Internet of Things. *IEEE Trans. Wirel. Commun.* **2026**, *25*, 2231–2246. <https://doi.org/10.1109/TWC.2025.3595550>.
39. Zhao, S.H.; Long, Y.S.; Li, L.H.; Gu, B.; Gong, S.M.; Xiong, Z.H. Joint Traffic Reshaping and Channel Reconfiguration in RIS-Assisted Semantic NOMA Communications. In Proceedings of the 2025 IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA), Shenyang, China, 10–12 October 2025; pp. 9–14. <https://doi.org/10.1109/ISPA67752.2025.00011>.
40. Xie, P.; Li, F.; Zhang, M.C.; Quan, W.; Zhu, J.L.; Cheng, N. STAR-RIS Assisted Information Transmission Based on Fairness in Semantic Communication Systems. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 17007–17020. <https://doi.org/10.1109/TWC.2024.3449414>.
41. Liu, S.; Yang, H.L.; Zheng, M.T.; Xiao, L.; Xiong, Z.H.; Niyato, D. UAV-Enabled Semantic Communication in Mobile Edge Computing Under Jamming Attacks: An Intelligent Resource Management Approach. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 17493–17507. <https://doi.org/10.1109/TWC.2024.3454073>.
42. Yan, L.; Qin, Z.J.; Zhang, R.; Li, Y.Z.; Li, G.Y. Resource Allocation for Text Semantic Communications. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 1394–1398. <https://doi.org/10.1109/LWC.2022.3170849>.
43. Baker, B.M.; Ayeche, M.A. A genetic algorithm for the vehicle routing problem. *Comput. Oper. Res.* **2003**, *30*, 787–800. [https://doi.org/10.1016/S0305-0548\(02\)00051-5](https://doi.org/10.1016/S0305-0548(02)00051-5).
44. Fu, Y.G.; Ding, M.Y.; Zhou, C.P.; Hu, H.P. Route Planning for Unmanned Aerial Vehicle (UAV) on the Sea Using Hybrid Differential Evolution and Quantum-Behaved Particle Swarm Optimization. *IEEE Trans. Syst. Man, Cybern. Syst.* **2013**, *43*, 1451–1465. <https://doi.org/10.1109/TSMC.2013.2248146>.
45. Zheng, M.T.; Yang, H.L.; Liu, S.; Lin, K.L.; Xiao, L.; Han, Z. Reliable Semantic Communication With QoE-Driven Resource Scheduling for UAV-Assisted MEC. *IEEE Trans. Veh. Technol.* **2025**, *74*, 11484–11489. <https://doi.org/10.1109/TVT.2025.3542775>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.