Semantic-Aware Priority-Based Resource Allocation for C-V2X Platoons Using Transformer Encoding

Piyush Singh*, Wan-Jen Huang†, Bishmita Hazarika‡, Keshav Singh†, and Trung Q. Duong‡
*Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan
†Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan
‡Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. Johns, Canada
Email: piyush230497@gmail.com, bhazarika@mun.ca, ksingh1980@ieee.org, wanjen.huang@g-mail.nsysu.edu.tw, tduong@mun.ca

***Abstract*—This study presents a semantic-aware resource allocation framework for cellular vehicle-to-everything (C-V2X) platooning systems to support coordinated mobility and safety-critical communication. The framework integrates vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-pedestrian (V2P) links to enhance operational efficiency and reliability. We develop a multi-agent reinforcement learning algorithm for semantic-aware priority-based intelligent resource allocation named SPIRE, that jointly optimizes channel assignment, transmit power, and semantic symbol allocation under stringent real-time constraints. SPIRE incorporates a transformer-based state encoder for spatiotemporal representation, a hybrid dual-critic architecture for stable policy learning, and a priority-weighted reward mechanism to ensure high quality of experience (QoE) and success rate of semantic information transmission (SRS). Extensive simulations show that SPIRE consistently outperforms conventional deep reinforcement learning (DRL) baselines, achieving superior QoE and SRS across varying semantic loads and inter-vehicular spacings. These results demonstrate the framework's robustness and scalability, making it well-suited for future C-V2X deployments in complex urban environments.**

***Index Terms*—Cellular vehicle-to-everything (C-V2X), intelligent transportation systems (ITS), multi-agent reinforcement learning (MARL), semantic communication, vehicular networks.**

I. Introduction

Cellular vehicle-to-everything (C-V2X) communication serves as a critical foundation for intelligent transportation systems (ITS), particularly in enabling vehicle platooning, where ultra-reliable and low-latency messaging is essential for cooperative driving, adaptive cruise control, and emergency response [1]. By leveraging centralized scheduling and spectrum reuse, C-V2X supports high-throughput and low-delay applications. However, resource allocation in highly dynamic vehicular environments remains a challenge due to mobility, channel fading, and frequent topology changes. Conventional algorithms that focus on throughput maximization or delay minimization often fall short in capturing the time-sensitive and context-aware nature of platooning data. Recently, deep reinforcement learning (DRL) and multi-agent reinforcement learning (MARL) have been adopted to enable real-time, decentralized spectrum and power control across vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) links, offering improved adaptability with minimal signaling overhead [2], [3].

Despite the success of DRL and MARL in dynamic network control, most frameworks treat all transmitted data uniformly, ignoring the contextual significance of messages. This is problematic in safety-critical scenarios, where emergency messages should take precedence over routine updates. Semantic communication provides a promising alternative by focusing on the transmission of meaning rather than raw bits, enabling content-aware compression and prioritization [4], [5]. Semantic encoders have demonstrated resilience in low-SNR environments and improved spectrum efficiency through task-specific encoding [6], [7]. However, these systems are generally developed for point-to-point terrestrial links and fail to address the heterogeneous and multi-modal demands of C-V2X networks, which involve V2V, V2I, and vehicle-to-pedestrian (V2P) communications [8], [9]. Moreover, achieving ultra-reliable and low-latency semantic delivery in dense urban environments remains an open challenge, especially when considering multi-agent coordination across varying channel and mobility conditions.

To overcome these limitations, recent work has proposed combining semantic communication with MARL to support real-time resource adaptation under network constraints [10], [11]. These approaches offer decentralized control and some semantic awareness but often neglect message prioritization and task-specific reliability requirements. Balancing traditional physical-layer metrics like SINR and latency with semantic-level indicators such as semantic rate and accuracy is still underexplored. While TD3 and PPO-based DRL models improve convergence and policy robustness, they lack mechanisms to account for message criticality [12]. Similarly, although MARL frameworks scale well in multi-agent settings, they often face convergence issues and do not incorporate semantic objectives directly [13]. In practice, this gap limits the ability to support differentiated services across platoon members and pedestrians in urban mobility systems.

This study addresses these gaps by introducing a semantic-aware resource management framework for C-V2X platooning, which combines a multi-modal semantic encoder/decoder with a distributed MARL algorithm. Each vehicle autonomously optimizes channel assignment, transmit power, and semantic symbol selection based on message importance and link dynamics. We propose the Semantic Priority-based Intelligent Resource Allocation (SPIRE) algorithm, which employs transformer-based state encoding, a hybrid dual-critic architecture, and a priority-weighted reward design to improve Quality of Experience (QoE) and Success Rate of Semantic information transmission (SRS). Through extensive simulations, we show that SPIRE outperforms benchmark DRL approaches including DDPG, TD3, and PPO across varying semantic loads, inter-vehicular spacings, and message priorities. These results validate the framework's robustness, scalability, and practical applicability to next-generation C-V2X deployments. The key contributions of this paper are as follows:

- We propose a comprehensive semantic-aware resource management framework for C-V2X platooning systems, integrating V2V, V2I, and V2P communications to enhance safety and efficiency by optimizing resource allocation for both vehicular and pedestrian nodes in dynamic urban environments.
- We introduce the semantic priority-based intelligent resource allocation (SPIRE) algorithm, which incorporates a transformer-based state encoder for efficient feature extraction, a hybrid dual-critic architecture for stable learning, and a priority-weighted reward design to emphasize the transmission of critical messages. These components ensure high adaptability and performance in heterogeneous communication scenarios.
- Our framework adopts a priority-weighted strategy to jointly optimize Quality of Experience (QoE) and the Success Rate of Semantic transmission (SRS), ensuring that high-priority messages are given preferential treatment to improve system reliability.
- Simulation results demonstrate that SPIRE consistently outperforms baseline DRL methods such as DDPG, TD3, and PPO in terms of QoE and SRS under varying inter-vehicular spacings, semantic loads, and priority levels, validating its robustness and real-world applicability.

II. System Model

We propose a semantic-aware resource management framework tailored for Cellular Vehicle-to-Everything (C-V2X) platooning systems, as shown in Fig. 1. The system comprises N platoons, each consisting of M_n vehicles: one PL and $M_n - 1$ platoon members (PMs). To enhance safety and situational awareness in urban settings, P pedestrian nodes are incorporated, facilitating V2P communication. The framework supports three connectivity types: V2V for intra-platoon coordination, V2I for PLs interfacing with a central BS, and V2P for real-time safety alerts and environmental feedback.

Messages transmitted within the system are categorized by priority—high, common, or low—denoted by factors $q_{n,m}$ for vehicles and q_p for pedestrians. These priorities drive resource allocation strategies and directly impact the QoE, ensuring that critical messages, such as collision warnings, are prioritized over routine updates. The system leverages an Orthogonal Frequency-Division Multiplexing (OFDM) architecture with

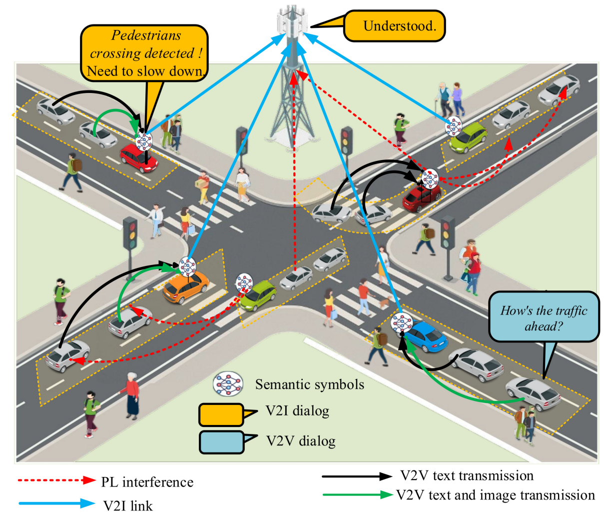

Fig. 1: An illustration of the considered vehicular framework.

K subchannels to mitigate interference and optimize spectral efficiency, adapting to the dynamic and dense communication demands of C-V2X environments.

A. Network Topology

The network topology encompasses N platoons, where each platoon includes a PL responsible for coordination and $M_n - 1$ PMs, yielding a total of $M = \sum_{n=1}^{N} M_n$ vehicles. Pedestrian nodes, represented as $\mathcal{P} = \{1, 2, \ldots, P\}$, are strategically distributed in high-risk zones such as crosswalks, intersections, and pedestrian-heavy areas. The BS serves as the central hub for V2I communication, enabling global resource management and data aggregation. V2V links ensure low-latency intra-platoon exchanges, while V2P links provide direct, short-range communication for immediate safety responses.

B. Communication Framework

The communication framework adopts a semantic-aware approach to maximize the efficiency of data transmission by focusing on the meaning of messages rather than raw bit streams. For text-based messages, such as status updates or alerts, the Deep Semantic Communication (DeepSC [6]) model is employed. For multi-modal data integrating text and images—e.g., sensor readings paired with visual context—the Multi-User Deep Semantic Communication (MU-DeepSC [14]) model is utilized. The semantic encoding process for single-modal communication is defined as:

$$X_T = \mathcal{C}_{\beta_T}\big(\mathcal{S}_{\alpha_T}(S_T)\big), \tag{1}$$

where S_T is the source text, $\mathcal{S}_{\alpha_T}$ is the semantic encoder extracting key meaning, and $\mathcal{C}_{\beta_T}$ is the channel encoder preparing data for transmission. For multi-modal communication, the semantic transmission is defined as:

$$X_I = \mathcal{C}_{\beta_I}(\mathcal{S}_{\alpha_I}(S_I)), \tag{2}$$

where S_I is the input image, encoded semantically by $\mathcal{S}_{\alpha_I}(\cdot)$ and mapped to symbols via channel encoder $\mathcal{C}_{\beta_I}(\cdot)$.

This framework reduces bandwidth usage by prioritizing semantically significant information, adapting to channel conditions, and user requirements.

C. Transmission Model

The transmission model is built on an OFDM system with a subchannel set $\mathcal{K} = \{1, 2, \ldots, K\}$, enabling flexible resource allocation across V2V, V2I, and V2P links. For text-based V2V communication over subchannel k, the received signal at a platoon member is given by

$$Y_{T,n,m}[k] = \rho_{n,k} h_{n,m}[k] X_{T,n}[k] + I_{T,n,m}[k] + \chi_{T,n,m} \,, \quad (3)$$

where $\rho_{n,k} \in \{0, 1\}$ denotes subchannel assignment, $h_{n,m}[k]$ is the complex channel gain incorporating path loss and fading, $I_{T,n,m}[k]$ is the interference from other transmitters, and $\chi_{T,n,m}$ is additive white Gaussian noise (AWGN) with variance σ^2. The signal-to-interference-plus-noise ratio (SINR) is computed as

$$SINR_{T,n,m}[k] = \frac{\rho_{n,k}\, p_{T,n}[k]\, h_{n,m}[k]}{I_{T,n,m}[k] + \sigma^2} \,, \quad (4)$$

where $p_{T,n}[k]$ is the transmit power, and interference is modeled as:

$$I_{T,n,m}[k] = \sum_{n' \neq n} \rho_{n',k}\, \beta_{n',k}\, p_{T,n'}[k]\, h_{n',m}[k] \,, \quad (5)$$

with $\beta_{n',k}$ representing power control factors. Similar formulations apply to V2I and V2P links, adjusted for their respective channel characteristics—e.g., higher path loss in V2I due to BS distance or LOS dominance in V2P. Power and subchannel allocation are optimized to balance throughput and reliability under QoE constraints.

D. Receiver Model

The receiver model focuses on reconstructing semantically meaningful information from the received signals. For text data, the process involves channel decoding followed by semantic decoding $\hat{S}_T = \mathcal{D}_{\alpha_T}\left(\mathcal{C}^{-1}_{\beta_T}(Y_{T,n,m}[k])\right)$. In multi-modal operation, text and image streams are first decoded separately via DeepSC and MU-DeepSC, then fused to yield a joint semantic output (e.g., pairing a textual warning with an obstacle image). The receiver mitigates channel effects through equalization and error correction, and dynamically adjusts decoding thresholds using priority factors $q_{n,m}$ and q_p to ensure that high-priority messages are delivered with maximal reliability.

E. Performance Metrics

The system's is evaluated using three key metrics:

1) Semantic Accuracy: It measures the fidelity of reconstructed data, defined as

$$\xi_{SM} = \Psi_{SM}\big(u_T, SINR_{T,n,m}[k]\big) \,, \quad (6)$$

where Ψ_{SM} is a similarity function and u_T is the transmitted data size.

2) Semantic Rate: It quantifies the rate of meaningful information transfer, given by

$$\phi_n = \frac{W\, \tilde{H}_{SM}}{u_T} \,, \quad (7)$$

where W is the bandwidth and $\tilde{H}_{SM}$ is the semantic entropy.

3) Quality of Experience (QoE): It integrates rate and accuracy with priority weighting:

$$QoE_n^m = q_{n,m}\Big[\omega_q\, Score_R\big(\phi_n^m\big) + \big(1 - \omega_q\big)\, Score_A\big(\xi_n^m\big)\Big], \quad (8)$$

where $\omega_q \in [0, 1]$ balances rate (Score$_R$) and accuracy (Score$_A$). These metrics collectively inform resource optimization, ensuring the system meets the diverse demands of C-V2X platooning applications.

III. Problem Statement

The objective of this work is to optimize semantic-aware resource allocation in C-V2X platooning systems by maximizing the overall QoE for vehicular (intra-platoon V2V and inter-platoon V2I) and pedestrian (V2P) communications, while maintaining a high SRS to ensure timely and reliable message exchange. The optimization problem is formulated as

$$\begin{aligned}
(\mathcal{P}):\quad & \max_{\{\beta_n, \beta_n^p, \rho_n, p_n, u_m, u_p\}} \sum_{n \in \mathcal{N}} \Bigg\{ \sum_{m \in \mathcal{M}_n} \Big[q_{n,m}\, QoE_n^m\Big] \\
& \qquad + \sum_{p \in \mathcal{P}} \Big[q_p\, QoE_p\Big] + \lambda\, \Pr\Big(\Phi_n\Big) \Bigg\} \\
\text{s.t.}\quad & (C.1)\quad \sum_{n \in \mathcal{N}} \Big(\beta_{n,k} + \beta_{n,k}^p\Big) \le 1, \quad \forall k \in \mathcal{W}, \\
& (C.2)\quad \sum_{k \in \mathcal{W}} \beta_{n,k} \le 1, \sum_{k \in \mathcal{W}} \beta_{n,k}^p \le 1 \quad \forall n \in \mathcal{N}, \\
& (C.3)\quad \begin{array}{l} u_T^m \in \{0, 1, \ldots, u_{T,max}\}, \\ u_I^m \in \{0, 1, \ldots, u_{I,max}\}, \quad \forall m \in \mathcal{M}_n, \end{array} \\
& (C.4)\quad \begin{array}{l} u_T^p \in \{0, 1, \ldots, u_{T,max}\}, \\ u_I^p \in \{0, 1, \ldots, u_{I,max}\}, \quad \forall p \in \mathcal{P}, \end{array} \\
& (C.5)\quad 0 \le p_m \le p_{max}, \quad \forall m \in \mathcal{M}_n, \quad \forall k \in \mathcal{W}, \\
& (C.6)\quad \begin{array}{l} Score_{R,m} \ge G_{th}, \\ Score_{A,m} \ge G_{th}, \quad \forall m \in \mathcal{M}_n, \quad \forall k \in \mathcal{W}, \end{array} \\
& (C.7)\quad 0 \le \sum_{k=1}^{K} \sum_{n=1}^{N} \Bigg(\sum_{m \in \mathcal{M}_n} B_s[m,k] + \sum_{p \in \mathcal{P}} B_s^p[p,k] \Bigg) \\
& \qquad\qquad \le B_{\max}^s, \\
& (C.8)\quad \sum_{t=1}^{T} \Bigg(\sum_{m \in \mathcal{M}_n} \beta_{n,k}\, \phi_k^d[m,t] + \sum_{p \in \mathcal{P}} \beta_{n,k}^p\, \phi_k^d[p,t] \Bigg) \\
& \qquad\qquad \ge \frac{B_s}{\Delta T}, \quad \forall n \in \mathcal{N}.
\end{aligned} \quad (9)$$

where $q_{n,m}$ and q_p are priority factors for vehicles and pedestrians, respectively, and λ weights the probability of timely semantic transmission $\Pr(\Phi_n)$. Constraints (C.1)–(C.8) ensure interference management, subchannel usage limits, semantic symbol ranges, power limits, quality thresholds, data load caps, and timely transmission, respectively.

Given the non-convex and mixed-integer nature of this problem, we propose a novel multi-agent reinforcement learning approach, SPIRE, detailed in the next section.

IV. Proposed Solution Framework

We propose a multi-agent reinforcement learning framework, **SPIRE** (*Semantic Priority-based Intelligent REsource allocation*), to solve the non-convex MINLP in (9), which involves discrete channel assignments, SINR-based non-linear constraints, and coupled V2V/V2I/V2P interactions under real-time constraints. To manage complexity, SPIRE applies continuous relaxation and decentralized learning, featuring:

1) **Transformer-Based State Encoding:** A Transformer encoder captures spatiotemporal features from high-dimensional inputs (e.g., channel conditions, interference, semantic loads, priorities).
2) **Hybrid Dual-Critic Architecture:** Twin global critics mitigate overestimation, while local critics handle agent-specific evaluations.
3) **Priority-Weighted Reward Design:** Predefined priority factors $q_{n,m}$ (for vehicle m in platoon n) and q_p (for pedestrians) scale QoE-based rewards to emphasize high-priority messages.
4) **Clipped Surrogate Objective:** Inspired by PPO, this constrains policy updates to stabilize learning.

A. State, Action, and Reward Design

State: The state of platoon leader (PL) n at time t is defined as:

$$s_n^t = \Big(h_{n,BS}[k],\, \{h_{n,m}[k]\}_{m\in\mathcal{M}_n},\, I_T^t[k],\, I_I^t[k],\, B_s^t,\, \phi_d^t,\, \{q_{n,m}\}_{m\in\mathcal{M}_n}\Big), \tag{10}$$

where $h_{n,BS}[k]$ and $\{h_{n,m}[k]\}$ represent the channel gains from PL to the base station and follower vehicles over subchannel k, respectively. $I_T^t[k]$ and $I_I^t[k]$ are interference levels for text and image modalities, B_s^t is the semantic transmission load, ϕ_d^t the instantaneous semantic rate, and $\{q_{n,m}\}$ are priority factors for vehicle messages. For pedestrian nodes, similar structures apply with q_p. The state is processed by a Transformer encoder to extract spatiotemporal features as $\tilde{s}_n^t = \text{TransformerEncoder}(s_n^t)$.

Action: Each PL selects $a_n^t = (\beta_{n,k},\, \beta_{n,k}^p,\, \rho_n,\, p_n,\, u_m,\, u_p)$, where $\beta_{n,k}$ and $\beta_{n,k}^p$ are binary channel assignments for vehicle and pedestrian communications, ρ_n indicates the mode (V2V or V2I), p_n is the transmit power, and u_m, u_p are the semantic symbol counts for vehicles and pedestrians.

Algorithm 1 SPIRE Algorithm

Initialize: Global critics $Q_g^{\psi_1}$, $Q_g^{\psi_2}$, target critics $Q_g^{\psi_1'}$, $Q_g^{\psi_2'}$.
Initialize: For each agent n, actor π_{θ_n}, local critic $Q_n^{\phi_n}$, and corresponding target networks.
Initialize: Replay buffer $\mathcal{B}$, learning rates η_π, η_Q, and soft update rate τ.
for each episode **do**
  Reset the environment; initialize platoons, pedestrian nodes, and channel conditions.
  for each time step t **do**
    for each agent $n \in \mathcal{N}$ **do**
      Observe the raw state s_n^t.
      Compute encoded state:

$$\tilde{s}_n^t = \text{TransformerEncoder}(s_n^t). \tag{12}$$

      Select action:

$$a_n^t = \pi_{\theta_n}(\tilde{s}_n^t). \tag{13}$$

    end for
    Execute joint action; observe next states s_n^{t+1} and obtain local rewards r_n^t via Eq. (11).
    Store $(s_n^t, a_n^t, r_n^t, s_n^{t+1})$ in $\mathcal{B}$.
    if size($\mathcal{B}$) $> B_{th}$ **then**
      Sample a mini-batch from $\mathcal{B}$.
      for each global critic $j = 1, 2$ **do**
        Compute the target value using Eq. (14).
        Update the global critic by minimizing the loss in Eq. (15).
      end for
      Update each local critic similarly.
      Update each actor using the policy loss in Eq. (16).
      Update target networks via Eqs. (18) and (19).
    end if
  end for
end for

Reward: The local reward for agent n at time t captures both the QoE and the timeliness of semantic transmission, weighted by the priority factor is given by:

$$r_n^t = w_1 \cdot \Big(q_{n,m}\, QoE_n^t\Big) - w_2 \cdot \left(1 - \frac{1}{1 + \exp\Big(-\alpha\Big(\Phi_n^t - \frac{B_s}{\Delta T}\Big)\Big)}\right), \tag{11}$$

where w_1 and w_2 are weight factors which balances QoE and timeliness, with $\Phi_n^t = \sum_{\tau=1}^{t}\sum_{m\in\mathcal{M}_n} \beta_{n,k}\, \phi_k^d[m,\tau]$ as the cumulative semantic information. The global reward is the average over all agents: $r^g = \frac{1}{N}\sum_{n\in\mathcal{N}} r_n^t$.

B. Proposed SPIRE Algorithm

Critic Networks: Twin global critics $Q_g^{\psi_1}$ and $Q_g^{\psi_2}$ estimate joint state-action values. The target value is:

$$y^g = r^g + \gamma \min\{Q_g^{\psi_1'}(s', a'), Q_g^{\psi_2'}(s', a')\}, \tag{14}$$

where s', a' are next states/actions, and γ is the discount factor. Loss for critic j:

$$L_g(\psi_j) = \mathbb{E}[(Q_g^{\psi_j}(s,a) - y^g)^2], \tag{15}$$

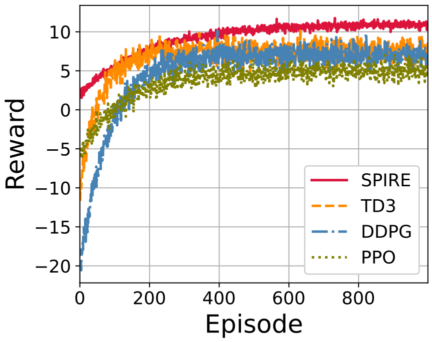

Fig. 2: Reward vs. episodes.

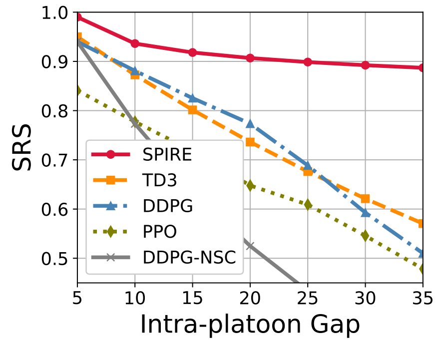

Fig. 3: SRS vs. intra-platoon gap.

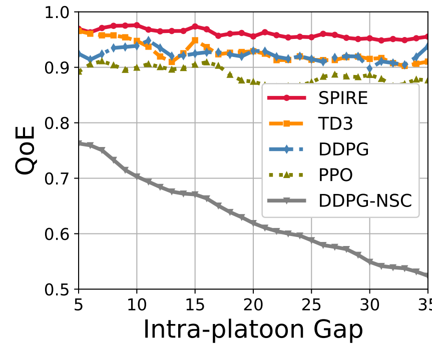

Fig. 4: QoE vs. intra-platoon gap.

with expectations taken over mini-batches. Local critics $Q_n^{\phi_n}$ use the same formulation.

Actor Networks: Each actor $\pi_{\theta_n}(\tilde{s}_n)$ outputs an action from encoded state $\tilde{s}_n$. The policy loss:

$$L_\pi(\theta_n) = \mathbb{E}\left[\min(r_n(\theta_n)A_n, \text{clip}(r_n(\theta_n), 1-\epsilon, 1+\epsilon)A_n)\right], \tag{16}$$

where A_n is the advantage and

$$r_n(\theta_n) = \frac{\pi_{\theta_n}(a_n|s_n)}{\pi_{\theta_n^{old}}(a_n|s_n)} \tag{17}$$

is the probability ratio; ϵ is the PPO clip parameter.

Target Updates:

$$\theta_n' \leftarrow \tau\theta_n + (1-\tau)\theta_n', \quad \phi_n' \leftarrow \tau\phi_n + (1-\tau)\phi_n', \tag{18}$$

$$\psi_j' \leftarrow \tau\psi_j + (1-\tau)\psi_j', \quad j = 1, 2, \tag{19}$$

with $\tau \in (0, 1)$ as the soft update rate. The training loop is presented in **Algorithm 1**.

C. Discussion

SPIRE transforms the intractable MINLP into a solvable MARL formulation through Transformer-based encoding, dual-critic architecture, and priority-weighted reward shaping. It effectively aligns resource allocation with semantic priorities and communication constraints in dynamic C-V2X platooning environments.

V. Simulation Results

In this section, we will discuss the simulation setup and evaluate the performance of the proposed scheme.

A. Simulation setup

The performance of the SPIRE framework was evaluated in a cellular-based C-V2X urban scenario using an Intel Core i5-12600K (3.70 GHz, 10 cores), 32GB RAM, and NVIDIA RTX 4070 Ti GPU on Windows 11. The single-modal text dataset used for semantic extraction is the European Parliament corpus, containing around 2 million sentences and 53 million words. After pre-processing, sentence lengths are limited to 1–20 words. The dataset is split 90% for training and 10% for testing. For multi-modal learning, the CLEVR dataset is used, comprising 70,000 training images with 699,989 questions and 15,000 test images with 149,991 questions. The SAC parameter settings are provided in **Table I**.

TABLE I: Network parameters

Parameter	Value	Parameter	Value
W	180 kHz	K	10
N	5	$p_{\max}$	30 dBm
M	6	σ^2	-114 dBm
P	20	B_s	$1000-6000$ suts
u_T	$5-30$ sym.	G_{th}	0.5
ω_q, ω_p	U(0,1)	q_H, q_C, q_L	$1, 0.7, 0.3$
γ	0.99	τ	0.005
Buffer Size	10^6	Mini-batch	64
η_π	0.0001	η_Q	0.001
α	0.1	ϵ	0.2
Episodes	500	Iter./Ep.	100
Noise	$\mathcal{N}(0, 0.2)$	Carrier Freq.	2 GHz

B. Performance evaluation and benchmarks

To evaluate the effectiveness of the proposed SPIRE framework, we assess QoE and SRS in heterogeneous C-V2X networks that support both vehicular (V2V/V2I) and pedestrian (V2P) communications. SPIRE is compared against the following baselines:

- **DDPG-NSC:** A multi-agent Deep Deterministic Policy Gradient (DDPG) method without semantic awareness, relying on conventional bit-based transmission.
- **DDPG:** A semantic-aware multi-agent DDPG algorithm for resource allocation.
- **TD3:** A twin delayed deep deterministic policy gradient (TD3)-based algorithm that incorporates semantic information.
- **PPO:** A multi-agent Proximal Policy Optimization (PPO) approach used for baseline comparison in terms of policy stability and performance.

C. Results

Fig. 2 illustrates the reward convergence of SPIRE against baseline methods. SPIRE achieves the fastest and most stable convergence, with a sharp rise in reward within the first 200 episodes and minor fluctuations thereafter. This is attributed to its transformer-based state encoding and clipped surrogate updates, which accelerate learning while maintaining stability. TD3 attains a high final return but suffers from significant reward oscillations due to delayed updates and exploration noise. DDPG converges more gradually, reaching a lower but stable plateau near episode 500. PPO, constrained by first-order

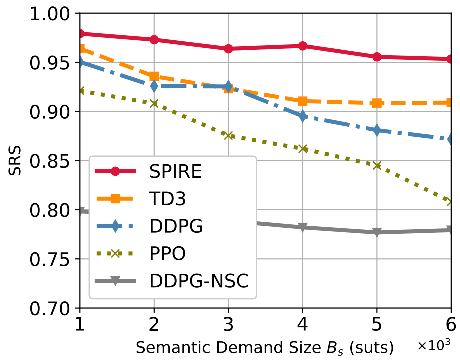

Fig. 5: SRS vs. demand size.

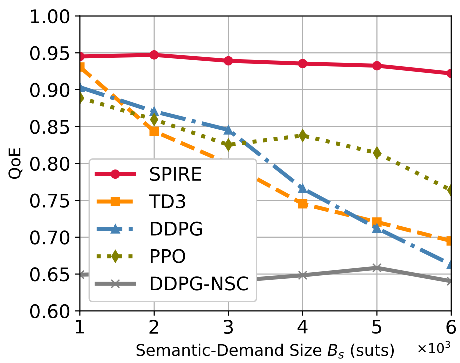

Fig. 6: QoE vs. demand size.

on-policy updates, shows the slowest improvement and lowest final reward.

Fig. 3 and 4 show how increasing inter-vehicular spacing impacts the SRS and QoE. SPIRE demonstrates the strongest robustness, with SRS decreasing only slightly from near 1.0 at 5 m to 0.89 at 35 m, and consistently high QoE across all distances. This is due to its priority-weighted reward and transformer-enhanced state encoding. TD3 maintains moderate performance with a linear SRS drop and slight QoE decline. DDPG performs well at short distances but degrades more sharply beyond 15 m. PPO is the most sensitive to spacing, with significant drops in both SRS and QoE due to limited adaptability. The bit-based DDPG-NSC baseline performs worst, exhibiting sharp degradation across both metrics, emphasizing the necessity of semantic-aware strategies for reliable performance under varying intra-platoon distances.

Fig. 5 and 6 analyze the effect of increasing semantic demand B_s on SRS and QoE. SPIRE shows the highest resilience, with SRS dropping marginally from near 100% to above 95% and QoE from 0.95 to 0.92 as B_s increases from 1×10^3 to 6×10^3 suts. TD3 follows with a moderate decline (SRS to 91%, QoE to 0.70), while DDPG experiences steeper drops (SRS to 87%, QoE to 0.66). PPO is more severely affected, with SRS falling to 81% and QoE to 0.76. DDPG-NSC remains flat around 80% SRS and 0.65 QoE, underscoring the importance of semantic-aware allocation under increasing information loads.

VI. Conclusions

In this study, we proposed a semantic-aware resource management framework for C-V2X platooning systems, introducing the SPIRE algorithm to address the demands of vehicular and pedestrian communications. By integrating transformer-based state encoding, a hybrid dual-critic architecture, and a priority-weighted reward mechanism, SPIRE dynamically optimizes channel assignment, power allocation, and semantic symbol selection to ensure high QoE and SRS across V2V, V2I, and V2P links. Simulation results demonstrate that SPIRE consistently outperforms baseline DRL methods such as DDPG, TD3, and PPO, maintaining superior performance across varying inter-vehicular spacings, semantic loads, and message priorities. Its ability to prioritize critical information and adapt to dynamic urban conditions highlights its advantages over conventional bit-based approaches in terms of resilience and reliability.

References

[1] M. Parvini, M. R. Javan, N. Mokari, B. Abbasi, and E. A. Jorswieck, "AoI-aware resource allocation for platoon-based C-V2X networks via multi-agent multi-task reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 9880–9896, 2023.

[2] P. Singh, B. Hazarika, K. Singh, C. Pan, W.-J. Huang, and C.-P. Li, "DRL-based federated learning for efficient vehicular caching management," *IEEE Internet Things J.*, vol. 11, no. 21, pp. 34 156–34 171, 2024.

[3] B. Hazarika, K. Singh, S. Biswas, S. Mumtaz, and C.-P. Li, "Multi-agent DRL-based task offloading in multiple RIS-aided IoV networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 1175–1190, 2024.

[4] X. Luo, H.-H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 210–219, 2022.

[5] J. M. Gimenez-Guzman, I. Leyva-Mayorga, and P. Popovski, "Semantic V2X communications for image transmission in 6G systems," *IEEE Netw.*, vol. 38, no. 6, pp. 48–54, 2024.

[6] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2663–2675, 2021.

[7] Z. Weng and Z. Qin, "Semantic communication systems for speech transmission," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2434–2444, 2021.

[8] K. Sehla, T. M. T. Nguyen, G. Pujolle, and P. B. Velloso, "Resource allocation modes in C-V2X: From LTE-V2X to 5G-V2X," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8291–8314, 2022.

[9] R. Q. Malik, K. N. Ramli, Z. H. Kareem, M. I. Habelalmatee, A. H. Abbas, and A. Alamoody, "An overview on V2P communication system: Architecture and application," in *Proc. IEEE IICETA*, 2020, pp. 174–178.

[10] B. Hazarika, K. Singh, C.-P. Li, A. Schmeink, and K. F. Tsang, "RADiT: Resource allocation in digital twin-driven UAV-aided internet of vehicle networks," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 11, pp. 3369–3385, 2023.

[11] B. Hazarika, K. Singh, S. Biswas, and C.-P. Li, "DRL-based resource allocation for computation offloading in IoV networks," *IEEE Trans. Ind. Inform.*, vol. 18, no. 11, pp. 8027–8038, 2022.

[12] P. Singh, B. Hazarika, K. Singh, W.-J. Huang, and C.-P. Li, "Augmented multiagent drl for multi-incentive task prioritization in vehicular crowdsensing," *IEEE Internet Things J.*, vol. 11, no. 22, pp. 36 140–36 156, 2024.

[13] Y.-C. Chung, H.-Y. Chang, R. Y. Chang, and W.-H. Chung, "Deep reinforcement learning-based resource allocation for cellular V2X communications," in *Proc. IEEE VTC-Spring*, 2023, pp. 1–7.

[14] H. Xie, Z. Qin, X. Tao, and K. B. Letaief, "Task-oriented multi-user semantic communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2584–2597, 2022.