Semantic-Aware Spectrum Efficiency for 6G V2X URLLC with Multi-Agent Hierarchical DRL

Piyush Singh*, Bishmita Hazarika†, Keshav Singh‡, Wan-Jen Huang‡, and Trung Q. Duong†

*Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan

†Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. Johns, Canada

‡Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan

Email: piyush230497@gmail.com, bhazarika@mun.ca, ksingh1980@ieee.org, wanjen.huang@g-mail.nsysu.edu.tw, tduong@mun.ca

***Abstract*—In this study, we propose SCF6, a novel semantic communication framework for 6G-enabled vehicular networks tailored to ultra-reliable low-latency communication (URLLC) scenarios. SCF6 integrates semantic encoding/decoding with conventional channel processing, optimizing transmission by focusing on data meaning. Leveraging BERT (bidirectional encoder representations from transformers)-based natural language processing, it ensures high semantic similarity between transmitted and received messages. To maximize semantic spectrum efficiency (SSEE) and success rate (SR) for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications under strict URLLC constraints, we design a multi-agent hierarchical attention-based semantic deep reinforcement learning (MAHAS-DRL) framework. MAHAS-DRL coordinates resource allocation and spectrum sharing, embedding hierarchical attention at both semantic and channel levels to enhance decision-making, optimize power control, and reduce interference. Simulations demonstrate SCF6's superiority over traditional DRL methods in spectrum efficiency, reliability, and latency, proving effective for dynamic urban vehicular networks.**

***Index Terms*—Vehicle-to-Everything (V2X), Ultra-Reliable Low-Latency Communication (URLLC), Multi-Agent Deep Reinforcement Learning (MADRL), Semantic Communication.**

I. Introduction

VEHICULAR networks, particularly vehicle-to-everything (V2X) communication, have become essential in modern transportation systems, enabling real-time data exchange critical for road safety, traffic management, and autonomous driving. V2X includes vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, supporting tasks like collision avoidance, traffic control, and cooperative driving [1]. As vehicular networks advance with 6G, they face heightened demands for ultra-reliable low-latency communication (URLLC), especially in dynamic urban environments where high mobility and fluctuating network conditions complicate stable, low-latency connectivity [1], [2].

Conventional V2X systems often transmit raw data without prioritizing its relevance, leading to inefficiencies, particularly under bandwidth constraints and high data volumes [3], [4]. Semantic communication, which emphasizes transmitting the meaning rather than raw bits, offers a solution to this inefficiency, enhancing data relevance and reducing bandwidth requirements [5]. While semantic-aware approaches have been explored in wireless networks [6], [7], their application in vehicular systems remains limited, with few studies addressing the unique demands of V2X communications under URLLC constraints [8]. Additionally, effective integration of semantic policies with resource and spectrum optimization has not been fully achieved, especially in multi-agent vehicular environments requiring coordinated decision-making.

Deep reinforcement learning (DRL) and multi-agent DRL (MADRL) approaches have been applied to V2X resource allocation and spectrum sharing, demonstrating promise for autonomous decision-making [9], [10]. However, DRL frameworks face challenges in scaling and adapting to dynamic environments and often struggle with the non-stationary nature of multi-agent networks [11], [12]. Recent work suggests that hierarchical attention mechanisms can improve decision-making by prioritizing critical information, yet their application to semantic and channel-level policies in V2X communication remains underexplored [13].

Addressing these gaps, we propose a novel semantic communication framework for 6G vehicular networks. Our approach introduces a hierarchical attention mechanism in MADRL, jointly optimizing semantic and channel-level policies to enhance spectrum efficiency, reliability, and communication relevance under URLLC constraints, making it suitable for dynamic vehicular environments. The primary contributions of this paper are as follows:

- We develop a novel semantic communication framework specifically designed for 6G-enabled vehicular networks, incorporating semantic encoding/decoding alongside traditional channel processing to meet URLLC requirements, ensuring efficient and reliable communication for both V2V and V2I links.
- We employ bidirectional encoder representations from transformers (BERT) to quantify semantic similarity, ensuring communication integrity while defining key performance metrics like high-speed semantic rate, semantic spectrum efficiency evaluation, and success rate to assess efficiency and reliability under URLLC constraints.
- We formulate an optimization problem to maximize both SSEE and SR for V2V and V2I communications while ensuring compliance with strict URLLC requirements, addressing challenges in spectrum sharing, power allocation, and resource management for efficient 6G vehicular networks.

- We propose a multi-agent DRL framework that coordinates decision-making across multiple vehicles to ensure optimal resource allocation and maximize spectrum usage efficiency in dynamic urban vehicular environments.
- The multi-agent DRL framework incorporates a hierarchical attention mechanism with channel-level and semantic-level policies, enabling intelligent coordination among agents by focusing on relevant communication links and environmental features, thereby improving V2V and V2I communication performance through reduced interference, optimized transmission power, and efficient spectrum sharing.
- Finally, we evaluate the effectiveness of our proposed framework by comparing it against popular DRL techniques, both with and without semantic communication, demonstrating the superior performance of our approach.

II. System Model

We introduce a *Semantic Communication Framework for 6G* (SCF6) for vehicular networks, supporting V2V and V2I communication by transmitting semantic information rather than raw data. SCF6 is tailored for 6G URLLC, achieving low latency and high reliability by integrating semantic and channel encoding/decoding for optimized data flow in dynamic vehicular contexts. Fig. 1 illustrates V2X communication, where semantic data exchange enables real-time route adjustments and traffic responses, supported by a shared knowledge base and source-channel joint coding. Our network model consists of N vehicles, a central base station (BS) at intersections, and M roadside units (RSUs) for coverage. Vehicles, equipped with SCF6 transceivers, follow a spatial Poisson distribution, moving at constant velocity v_i, and make intersection turns based on probabilities p_l, p_r, and p_s, where $p_l + p_r + p_s = 1$. Boundary-crossing vehicles re-enter from the opposite side to maintain density. Each SCF6 transceiver has a semantic encoder/decoder and a channel encoder/decoder to optimize semantic information transmission. At transmission, each vehicle generates a sentence $S_i[w] = [s_{i,1}[w], s_{i,2}[w], \ldots, s_{i,l_i}[w]]$ of l_i words for the w-th sub-band, processed by a semantic encoder to extract meaning, encoding it as $X_i[w] = \text{sem-enc}_\alpha(S_i[w])$. where $X_i[w] \in \mathbb{R}^{u_q \cdot l_i}$, and u_q represents the average semantic symbols per word. This semantic vector is then channel-encoded for transmission as $X'_i[w] = \text{ch-enc}_\beta(X_i[w])$. With $N_s = u_q \cdot l_i$ semantic symbols, the total semantic information encoded is I_i, with per-symbol information is

$$\frac{I_i}{N_s} = \frac{I_i}{u_q l_i}\,. \tag{1}$$

A. 6G Communication Model

The transmitted semantic signal $X'_i[w]$ in SCF6 faces wireless impairments like path loss, shadowing, fast fading, and interference. In 6G vehicular networks, URLLC demands

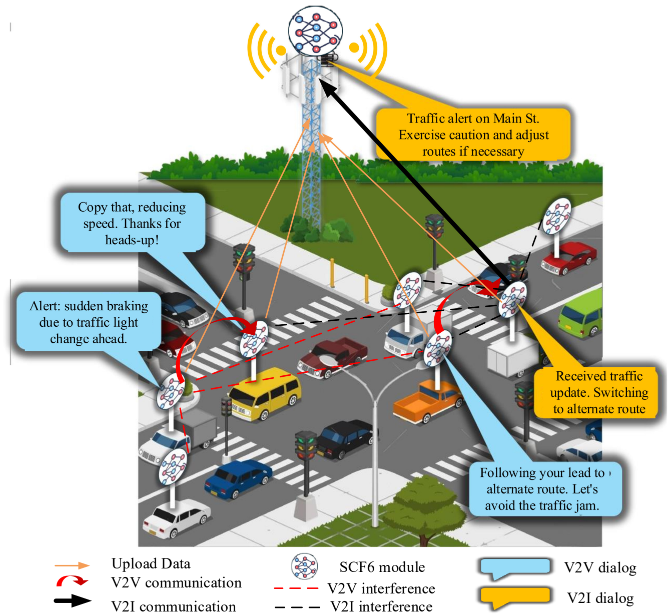

Fig. 1: An illustration of the considered vehicular framework.

SCF6 meet strict low-latency and high-reliability standards essential for vehicular applications. To meet the latency requirement, the transmission time T_{URLLC} must not exceed the maximum allowable delay $\tau_{\max}$, requiring SCF6 to optimize encoding and decoding to minimize delays. The reliability constraint ensures the packet error rate (PER) remains within acceptable limits, expressed as:

$$\text{PER}_{\text{URLLC}} = \mathbb{P}(\text{SINR} < \gamma_{\text{URLLC}}) \leq \epsilon_{\text{URLLC}}\,, \tag{2}$$

where γ_{URLLC} is the minimum required signal-to-interference-plus-noise ratio (SINR) and ϵ_{URLLC} (typically 10^{-5}) is the maximum permissible PER. Achieving these constraints requires that SINR levels exceed γ_{URLLC}, which SCF6 manages through optimal power allocation ($P_{\text{V2I}}[w]$ and P_q) and efficient spectrum sharing ($b_q[w]$). The received signal at the destination (BS or vehicle) is modeled as:

$$Y_i[w] = H_i[w]X'_i[w] + I_i[w] + N_i[w]\,, \tag{3}$$

where $H_i[w]$ represents channel gain on sub-band w, $I_i[w]$ is interference, and $N_i[w]$ denotes additive white Gaussian noise (AWGN). Channel gain $H_i[w]$ combines path loss, shadowing, and fast fading, expressed as $H_i[w] = s_{s_i}[w] \cdot l_{s_i}$, where $s_{s_i}[w]$ models small-scale fading and l_{s_i} accounts for large-scale path loss and shadowing. The path loss depends on whether the link is line-of-sight (LOS) or non-line-of-sight (NLOS), determined by the minimum of the horizontal and vertical distances between transmitter and receiver:

$$PL(d) = \begin{cases} PL_{\text{LOS}}(d), & \text{if } \min(d_{\text{h}}, d_{\text{v}}) < 7, \\ PL_{\text{NLOS}}(d_{\text{h}}, d_{\text{v}}), & \text{if } \min(d_{\text{h}}, d_{\text{v}}) \geq 7. \end{cases} \tag{4}$$

Shadowing, representing slow fading due to obstacles, is updated dynamically as vehicles move:

$$\mathcal{S}_{\text{new}} = \exp\left(-\frac{\Delta d}{d_{\text{dec}}}\right) \cdot \mathcal{S}_{\text{old}} + \sqrt{1 - \exp\left(-\frac{2\Delta d}{d_{\text{dec}}}\right)} \cdot X, \quad (5)$$

where Δd is the distance between successive positions, d_{dec} is the decorrelation distance, and $X \sim \mathcal{N}(0, 3)$. Fast fading $s_{s_i}[w]$ is modeled as a Rayleigh fading process, with $s_{s_i}[w] = |h_q(t)|$, where $h_q(t) \sim \mathcal{CN}(0, 1)$. The SINR for the w-th V2I link is calculated as [14]

$$\gamma_{\text{V2I}}[w] = \frac{P_{\text{V2I}}[w] h_{w,B}[w]}{\sigma^2 + \sum_{q=1}^{Q} b_q[w] P_q h_{q,B}[w]}, \quad (6)$$

where P_q is the transmission power of the q-th V2V link, $h_{q,B}[w]$ is the channel gain from the q-th V2V transmitter to the BS, and $b_q[w]$ indicates sub-band sharing.

For the q-th V2V link, the SINR is:

$$\gamma_{\text{V2V}}[w] = \frac{P_{\text{V2V}}[w] h_q[w]}{\sigma^2 + P_{\text{V2I}}[w] h_{w,q}[w] + \sum\limits_{q'=1, q' \neq q}^{Q} b_{q'}[w] P_{q'} h_{q',q}[w]}, \quad (7)$$

where $P_{\text{V2I}}[w]$ is the V2I transmission power, $h_{w,q}[w]$ is the channel gain from the V2I transmitter to the q-th V2V receiver, $h_{q',q}[w]$ is the channel gain from the q'-th V2V transmitter to the q-th receiver, and σ^2 is the noise power.

B. Receiver Model and Semantic Similarity Calculation

Upon receiving $Y_i[w]$, the receiver performs channel decoding to recover $\tilde{X}_i[w] = \text{ch-dec}_\nu(Y_i[w])$, where $\text{ch-dec}_\nu(\cdot)$ is the channel decoding function with parameter ν. Next, semantic decoding reconstructs the sentence $\hat{S}_i[w] = \text{sem-dec}_\mu(\tilde{X}_i[w])$, where $\text{sem-dec}_\mu(\cdot)$ is the semantic decoding function parameterized by μ.

To evaluate semantic reconstruction accuracy, we compute the cosine similarity between the original sentence $S_i[w]$ and the reconstructed sentence $\hat{S}_i[w]$ using BERT embeddings:

$$\xi = \frac{B(S_i[w]) \cdot B(\hat{S}_i[w])}{\|B(S_i[w])\| \cdot \|B(\hat{S}_i[w])\|}, \quad (8)$$

where $B(\cdot)$ represents the BERT embedding similar to [15], and $\|\cdot\|$ denotes the Euclidean norm. The mutual information (MI) between the transmitted symbol $X_i[w]$ and received symbol $Y_i[w]$ quantifies shared information and can be expressed as the Kullback-Leibler (KL) divergence between the joint distribution and the product of marginals:

$$\mathcal{I}(X_i[w]; Y_i[w]) = D_{\text{KL}}(p(X_i[w], Y_i[w]) \parallel p(X_i[w]) p(Y_i[w])), \quad (9)$$

where $p(X_i[w])$ and $p(Y_i[w])$ are marginal densities, and $p(X_i[w], Y_i[w])$ is the joint density. The KL divergence can be approximated using a neural network via its dual form as $D_{\text{KL}}(P \parallel Q) \geq \mathbb{E}_P[T] - \log \mathbb{E}_Q\left[e^T\right]$, where T is a function approximated by a neural network. The lower bound on the mutual information is then expressed as:

$$\mathcal{I}(X_i[w]; Y_i[w]) \geq L_{\text{MI}}(X_i[w], Y_i[w]; \alpha, \beta, T), \quad (10)$$

where L_{MI} is the lower bound on mutual information, approximated using a neural network with parameters α and β, representing the weights and biases of the neural network used to approximate T.

C. Performance Metrics

Then we evaluate the system performance in this study using two key metrics:

1) Semantic Spectrum Efficiency Evaluation (SSEE): SSEE evaluates the efficiency of semantic information transmission over the available bandwidth under URLLC constraints as:

$$\text{SSEE} = \frac{\text{HSR}}{B} = \frac{I}{B \cdot T_{\text{URLLC}}} \cdot \xi, \quad (11)$$

where B is the bandwidth in Hz. SSEE, in *suts/s/Hz*, reflects bandwidth efficiency under URLLC time limits, where HSR is the rate of successful semantic information transmission:

$$\text{HSR} = \frac{I}{T_{\text{URLLC}}} \cdot \xi, \quad (12)$$

where I is the average semantic information per sentence (in *suts*), T_{URLLC} is the maximum transmission time (in seconds), and ξ is the semantic similarity metric. HSR, in *suts/s*, ensures rapid, reliable semantic transmission under 6G URLLC.

2) Success Rate (SR): SR measures the reliability of V2V communication under URLLC, defined as:

$$\text{SR}_{\text{V2V}}[w] = \mathbb{P}\left(\text{SINR}_{\text{V2V}}[w] \geq \gamma_{\text{URLLC}}\right), \quad (13)$$

where γ_{URLLC} is the minimum SINR for URLLC.

III. Problem Formulation

Our objective is to optimize semantic-aware spectrum sharing in 6G vehicular networks by maximizing the SSEE for V2I links and the SR for V2V links under URLLC constraints. To enhance semantic communication performance, we aim to minimize the cross-entropy loss L_{CE} between $S_i[w]$ and $\hat{S}_i[w]$ for accurate sentence reconstruction:

$$L_{\text{CE}}(S_i[w], \hat{S}_i[w]) = -\sum_{l=1}^{l_i} q(s_{i,l}[w]) \log p(s_{i,l}[w]), \quad (14)$$

where $q(s_{i,l}[w])$ is the true probability of the l-th word in $S_i[w]$, and $p(s_{i,l}[w])$ is the predicted probability in $\hat{S}_i[w]$. We also aim to maximize mutual information $I(X_i[w]; Y_i[w])$ to increase shared information and enhance reliability, as in (9).

Joint optimization of the parameters α, β, μ, and ν of the semantic and channel encoders/decoders, to minimize L_{CE} and maximize $I(X_i[w]; Y_i[w])$, improves semantic similarity ξ and preserves the transmitted message's meaning under URLLC. Thus, the optimization problem is expressed as:

$$
\begin{aligned}
(\mathcal{P}): \quad & \max_{\{b_q[w], P_q, u_q\}} \quad \sum_{w=1}^{W} \mathrm{SSEE}[w] + \sum_{w=1}^{W} \sum_{q=1}^{Q} b_q[w] \cdot \mathrm{SR}[q,w] \\
\text{s.t.} \quad & (C.1) \quad b_q[w] \in \{0,1\}, \quad \forall q \in Q, \forall w \in W, \\
& (C.2) \quad \sum_{w=1}^{W} b_q[w] \leq 1, \quad \forall q \in Q, \\
& (C.3) \quad 0 \leq P_q \leq P_{\max,\mathrm{V2V}}, \quad \forall q \in Q, \\
& (C.4) \quad 0 \leq P_{\mathrm{V2I}}[w] \leq P_{\max,\mathrm{V2I}}, \quad \forall w \in W, \\
& (C.5) \quad \xi_{q,w} \geq \xi_{\min}, \text{and} \quad \mathrm{SINR}_{\mathrm{V2V}}[q,w] \geq \gamma_{\mathrm{URLLC}}, \\
& (C.6) \quad T_{\mathrm{URLLC}} \leq \tau_{\max}, \\
& (C.7) \quad \mathrm{SINR}_{\mathrm{V2I}}[w] \geq \gamma_{\mathrm{URLLC}}, \quad \forall w \in W.
\end{aligned} \tag{15}
$$

In this formulation, decision variables $b_q[w]$, P_q, and u_q are defined as follows: $b_q[w]$ is a binary variable indicating if the q-th V2V link shares sub-band w with a V2I link ($b_q[w] = 1$) or not ($b_q[w] = 0$). The objective function maximizes SSEE for V2I links and SR for V2V links, considering spectrum sharing via $b_q[w]$. Constraint $(C.1)$ ensures $b_q[w]$ is binary, while $(C.2)$ limits each V2V link q to one sub-band w, reflecting hardware constraints. Constraints $(C.3)$ and $(C.4)$ cap transmission powers P_q and $P_{\mathrm{V2I}}[w]$ within allowable limits. Constraint $(C.5)$ enforces both a minimum semantic similarity $\xi_{\min}$ and the minimum SINR requirement γ_{URLLC} for V2V links, ensuring both semantic fidelity and reliability. Constraint $(C.6)$ limits transmission time T_{URLLC} to $\tau_{\max}$, meeting URLLC latency requirements. Finally, constraint $(C.7)$ ensures the SINR for V2I links meets the URLLC threshold γ_{URLLC}, maintaining required reliability.

IV. Proposed Solution Framework

To optimize semantic-aware spectrum sharing in 6G vehicular networks, we propose the **M**ulti-**A**gent **H**ierarchical **A**ttention-Based **S**emantic **D**eep **R**einforcement **L**earning (MAHAS-DRL) framework. MAHAS-DRL employs a hierarchical policy structure to coordinate decision-making across vehicle agents, aiming to maximize semantic spectrum efficiency and success rates in V2I and V2V communications.

In this framework, each vehicle functions as an agent in a partially observable Markov decision process, where agents use local information (e.g., channel conditions, semantic tasks) to make coordinated decisions for optimizing global objectives. The hierarchical policy includes two levels:

1. *Semantic-Level Policy*: Determines the number of semantic symbols to transmit, optimizing semantic similarity, spectrum use, and transmission delay.
2. *Channel-Level Policy*: Manages spectrum sharing and power control, optimizing channel selection and transmission power to minimize interference and ensure reliable links.

The agent's local state at time t includes channel and semantic data as $s_q(t) = (H_q(t), Z_q(t), SD_q(t), T_q(t), \tau_q(t))$, where $H_q(t)$ represents channel gains, $Z_q(t)$ includes SINR and semantic symbol data, $SD_q(t)$ is remaining semantic data,

Algorithm 1 Training Process of MAHAS-DRL Framework

Initialize: Actor ρ, critics θ_1, θ_2, target networks $\theta_1' \leftarrow \theta_1$, $\theta_2' \leftarrow \theta_2$, experience buffer $\mathcal{B}$, vehicle positions, and encoders.
for each episode $k = 1, \ldots, K_{\max}$ **do**
  Reset environment, observe $s_q(0)$, and set $\mathcal{R}_{\text{curr}} = 0$.
  for each time step $t = 0, \ldots, T_{\max}$ **do**
    Encode $S_i[w]$ into $X_i[w]$ (19) and $X_i'[w]$ (17).
    Select $a_q(t)$ via $\pi_\theta(a_q|s_q)$ (24).
    Execute $a_q(t)$, observe $r_q(t)$, and state $s_q(t+1)$.
    Store $(s_q(t), a_q(t), r_q(t), s_q(t+1))$ in $\mathcal{B}$.
    Update $\mathcal{R}_{\text{curr}} += r_q(t)$.
    if $\mathcal{B}$ is full **then**
      Sample mini-batch from $\mathcal{B}$
      Compute $A_{q,q'}$ (20), and target $y(t)$ (23).
      Update critics by minimizing (21) and (22).
      Update actor network via gradient descent (24).
      Soft update target networks: $\theta_{\text{target}} \leftarrow \tau\theta + (1-\tau)\theta_{\text{target}}$.
    end if
  end for
  Stopping Criterion:
  if $|\mathcal{R}_{\text{curr}} - \mathcal{R}_{\text{prev}}| < \epsilon_{\text{stop}}$ for several episodes **then**
    Terminate training.
  else
    $\mathcal{R}_{\text{prev}} \leftarrow \mathcal{R}_{\text{curr}}$
  end if
  Decay exploration factor $\tau_q(t)$.
end for

$T_q(t)$ is the time budget, and $\tau_q(t)$ is an exploration factor. Each agent selects an action $a_q(t) = (b_q(t), P_{\mathrm{V2V}}(t), u_q(t))$ where $b_q(t) \in \{0,1\}$ indicates channel sharing with V2I, $P_{\mathrm{V2V}}(t) \in [0, P_{\max}]$ is the V2V transmission power, and $u_q(t) \in [u_{\min}, u_{\max}]$ is the number of semantic symbols for V2V transmission. The reward function balances agent and system performance, calculated as:

$$
r_q(t) = \lambda \cdot \mathrm{SSEE}(t) + (1-\lambda) \cdot \mathrm{SR}(t)\,, \tag{16}
$$

where $\lambda \in [0,1]$ controls the trade-off between SSEE and SR.

A. MAHAS-DRL Architecture

We implement MAHAS-DRL using a centralized training, decentralized execution (CTDE) paradigm. During training, agents share information with a central edge server at the BS, which coordinates policy updates. After training, each agent independently executes its learned policy without central communication. Each agent operates under an actor-critic framework, where the *actor* generates actions based on the current state, and dual *critics* estimate value functions. The hierarchical policy enables each agent to make decisions at both channel and semantic levels, optimizing spectrum efficiency and communication success.

The hierarchical policy has two levels: at the *channel level*, each agent selects power allocation $P_q(t)$ and sub-band sharing $b_q(t)$ to optimize SINR for V2V and V2I links, balancing SINR and interference reduction as:

$$
b_q(t) = \arg\max_{b_q} \left(\mathrm{SINR}_{\mathrm{V2V}}[w,t] - \lambda \cdot I_{\mathrm{V2V}}[w] \right)\,, \tag{17}
$$

where λ is the SINR-interference trade-off. Power allocation maximizes V2V SINR within $P_{\max}$:

$$
P_q(t) = \arg\max_{P_q} \mathrm{SINR}_{\mathrm{V2V}}[w,t] \quad \text{s.t.}\, P_q \leq P_{\max}\,. \tag{18}
$$

Algorithm 2 Testing Process of MAHAS-DRL Framework

```
Initialize: ρ*, vehicles, semantic & channel encoders.
Set number of testing episodes K_test.
for each test episode k = 1, ..., K_test do
    Reset environment, observe initial state s_q(0).
    for each time step t = 0, ..., T_max do
        Encode S_i[w] to X_i[w] (19) and X'_i[w] (17).
        Select a_q(t) via π_θ(a_q|s_q) (24).
        Execute a_q(t), observe r_q(t), and state s_q(t+1).
        Store (s_q(t), a_q(t), r_q(t), s_q(t+1)) for evaluation.
    end for
    Evaluation Metrics: Compute SSEE (11) and SR. (13).
end for
Output: Testing results (SSEE, SR).
```

At the *semantic level*, the agent decides $u_q(t)$, the number of semantic symbols to transmit, optimizing similarity ξ and transmission delay $T_{\text{trans}}(t)$:

$$u_q(t) = \arg\max_{u_q} \left(\xi \cdot \text{SSEE}(t) - \gamma \cdot T_{\text{trans}}(t)\right) , \tag{19}$$

where γ balances semantic similarity and delay. An attention mechanism in the critic network improves decision-making by focusing on critical links, and calculating attention scores as:

$$A_{q,q'} = \text{softmax}\left(W_a^T [s_q \oplus s_{q'}]\right) , \tag{20}$$

where W_a^T is the attention weight matrix. Each agent has two critic networks, Critic_1 and Critic_2, which minimize the Bellman error for Q-value estimation:

$$\mathcal{L}_{\text{Critic}_1}(t) = \mathbb{E}_{(s_q,a_q)\sim\mathcal{B}}\left[\left(Q_{\text{Critic}_1}(s_q,a_q) - y(t)\right)^2\right] , \tag{21}$$

$$\mathcal{L}_{\text{Critic}_2}(t) = \mathbb{E}_{(s_q,a_q)\sim\mathcal{B}}\left[\left(Q_{\text{Critic}_2}(s_q,a_q) - y(t)\right)^2\right] , \tag{22}$$

where the target $y(t)$ is calculated by the Bellman equation:

$$y(t) = r_q(t) + \gamma \cdot \min\left(Q_{\text{Critic}_1}(s'_q, a'_q), Q_{\text{Critic}_2}(s'_q, a'_q)\right) , \tag{23}$$

with reward $r_q(t)$, discount factor γ, and next state-action (s'_q, a'_q). The actor-network maximizes cumulative reward by minimizing:

$$\mathcal{L}_{\text{Actor}}(t) = \mathbb{E}_{s_q\sim\mathcal{B}}\Big[\log \pi_\theta(a_q|s_q) \cdot \min\Big(Q_{\text{Critic}_1}(s_q,a_q), Q_{\text{Critic}_2}(s_q,a_q)\Big)\Big] - \alpha H(\pi_\theta) , \tag{24}$$

where $H(\pi_\theta)$ is the policy entropy and α controls exploration. During training, the edge server optimizes actor and critic parameters ρ, θ_1, and θ_2 to establish channel and semantic policies under dynamic urban conditions. Target critic networks undergo soft updates at a rate τ.

The detailed steps of the training process are provided in **Algorithm 1**. After training, learned policies are deployed for independent agent execution, with testing evaluating SSEE and SR in real-time vehicular scenarios to assess MAHAS-DRL performance as described in **Algorithm 2**.

V. Simulation Results

We simulate a cellular-based vehicle communication network in an urban area of 1299×750 meters, using 3GPP TR 36.885 standards for V2V and V2I communications [16]. Vehicle movement follows a spatial Poisson process with updates every 100 ms. At intersections, vehicles change direction with probabilities of 50% straight, 25% left, and 25% right. Roads have four lanes (two per direction) with 3.5-meter width, and intersections are spaced $d_{\text{adj}} = 500$ meters apart. Vehicles travel at a constant speed of 35 km/h, maintaining uniform density. The semantic data size D_{sem} is set to $K \times 1060$ bits per u_q suts. The remaining parameters are given in Table I.. We use the European Parliament dataset [17] of 2.0 million sentences, preprocessed to sentence lengths of 1 to 20 words, with 90% for training and 10% for testing.

TABLE I: Simulation Parameters

Parameter	Value	Parameter	Value	Parameter	Value
N	30	M	5	v_i	36 km/h
p_s	50%	p_l	25%	p_r	25%
w_{lane}	3.5 m	d_{adj}	433 m	B	20 MHz
b_w	1 MHz	$P_{\text{max,V2V}}$	23 dBm	$P_{\text{max,V2I}}$	30 dBm
f_{carrier}	2 GHz	σ^2	-114 dBm	γ_{URLLC}	15 dB
γ_{V2V}	10 dB	$\mathcal{S}_{\text{V2V}}$	3 dB (std. dev.)	$\mathcal{S}_{\text{V2I}}$	8 dB (std. dev.)
d_{dec}	50 m	u_q	20 suts	l_i	10 words
I_i	100 bits	N_s	50 symbols	ξ_{min}	0.85
τ_{max}	1 ms	ϵ_{URLLC}	10^{-5}	K_{max}	10000
T_{max}	1000	ϵ_{stop}	10^{-3}	τ_{explore}	0.1
γ	0.99	λ	0.5	α	0.01
τ	0.005	ρ	3×10^{-4}	θ	3×10^{-4}
$H(\pi_\theta)$	0.01	$\mathcal{B}$	1×10^6	Batch size	256

A. Benchmark Schemes

We compare MAHAS-DRL with the following benchmarks:

1) **MAHAS-DRL$_{\text{NC}}$:** MAHAS-DRL with traditional, non-semantic communication.
2) **SAC$_{\text{SC}}$:** Multi-agent soft actor-critic (SAC) optimizing resource allocation with semantic encoding.
3) **SAC$_{\text{NC}}$:** Multi-agent SAC using traditional (non-semantic) communication.
4) **FDRL$_{\text{SC}}$:** Federated DDQN for decentralized semantic communication and resource allocation.

Note: MAHAS-DRL with semantic communication (**MAHAS-DRL$_{\text{SC}}$**) incorporates semantic encoding. Traditional communication refers to raw bit-level data transmission without semantic encoding.

B. Results

1) Impact of SSEE on different parameters: Fig. 2 illustrates the impact of vehicle count on SSEE for various DRL methods. In the figure, as vehicle numbers rise, MAHAS-DRLSC shows consistent SSEE gains due to its hierarchical attention mechanism, which prioritizes critical links, minimizes redundant transmissions and optimizes spectrum in dense urban networks. SACNC performs well initially but lacks scalability, while FDRLSC improves under high density through federated updates. Non-semantic methods (MAHAS-DRLNC, SAC$_{\text{NC}}$) achieve lower SSEE due to limited resource optimization in dense environments.

Fig. 3 illustrates the impact of V2I power on SSEE where MAHAS-DRLSC achieves high SSEE across V2I power levels by dynamically adjusting power to minimize interference, maintaining spectral efficiency. FDRLSC and SAC$_{\text{NC}}$ achieve

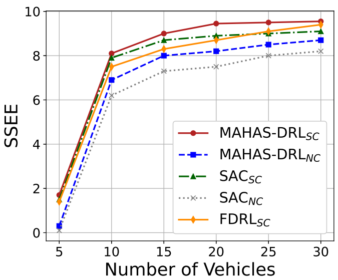

Fig. 2: SSEE vs. Number of vehicles.

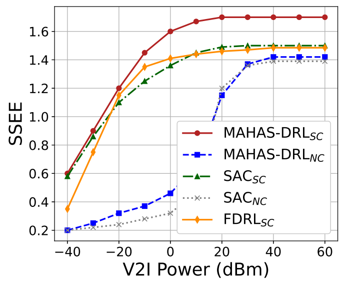

Fig. 3: SSEE vs. V2I power.

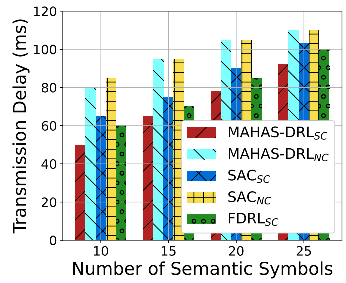

Fig. 4: Delay vs Semantics

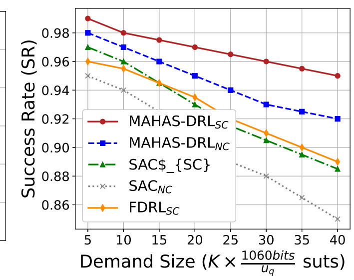

Fig. 5: Success rate vs. Demand size.

moderate efficiency but lack adaptive power control, while non-semantic variants yield lower SSEE, underscoring the role of semantic communication in efficient power and spectrum management.

2) Impact of Semantic Symbol Volume on Transmission Delay: Fig. 4 shows the effect of semantic symbol count on transmission delay for different methods, highlighting the MAHAS-DRL$_{\text{SC}}$ framework. MAHAS-DRL$_{\text{SC}}$ achieves consistently lower delay across various symbol counts, demonstrating its ability to handle larger data loads efficiently by integrating semantic communication with deep reinforcement learning. In contrast, non-semantic methods (MAHAS-DRL$_{\text{NC}}$ and SAC$_{\text{NC}}$) experience significantly higher delays, particularly as symbol counts increase, due to the lack of semantic encoding, which limits their scalability for symbol-rich data. FDRL$_{\text{SC}}$ also incurs higher delays among semantic methods due to federated learning's iterative updates, impacting convergence time.

3) Impact of demand size on success rate: Fig. 5 shows SR trends as a function of semantic demand size (K). MAHAS-DRL$_{\text{SC}}$ maintains high SR with increasing data demand, highlighting the effectiveness of semantic communication for managing larger data volumes and complex tasks. In contrast, MAHAS-DRL$_{\text{NC}}$ and SAC$_{\text{NC}}$ experience a sharper decline in SR as demand grows, due to the limitations of traditional communication methods in handling the increased load. FDRL$_{\text{SC}}$, while outperforming non-semantic approaches, lags behind MAHAS-DRL$_{\text{SC}}$ due to federated averaging overhead.

VI. Conclusions and Future Work

In this paper, we proposed a semantic communication framework tailored for 6G vehicular networks in URLLC scenarios, combining semantic encoding with multi-agent DRL and hierarchical attention for enhanced spectrum efficiency and reliability. Simulation results show that MAHAS-DRL significantly outperforms traditional methods in SSEE, SR, and transmission delay by optimizing spectrum sharing and resource allocation, especially in dense environments. The SCF6 and MAHAS-DRL frameworks provide a foundation for intelligent spectrum management in future 6G networks. Future work will extend this approach with advanced semantic models and adaptive risk-aware optimization for scalable, real-time vehicular networks.

References

[1] J. Li, S. Dang, M. Wen, Q. Li, Y. Chen, Y. Huang, and W. Shang, "Index modulation multiple access for 6G communications: Principles, applications, and challenges," *IEEE Netw.*, vol. 37, no. 1, pp. 52–60, 2023.

[2] Q. Wu, W. Wang, P. Fan, Q. Fan, J. Wang, and K. B. Letaief, "URLLC-awared resource allocation for heterogeneous vehicular edge computing," *IEEE Trans. Veh. Technol.*, vol. 73, no. 8, pp. 11 789–11 805, 2024.

[3] Y. Hui, J. Hu, N. Cheng, G. Zhao, R. Chen, T. H. Luan, and K. Aldubaikhy, "RCFL: Redundancy-aware collaborative federated learning in vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 6, pp. 5539–5553, 2024.

[4] B. Hazarika and K. Singh, "AFL-DMAAC: Integrated resource management and cooperative caching for URLLC-IoV networks," *IEEE Trans. Intell. Veh.*, vol. 9, no. 6, pp. 5101–5117, 2024.

[5] H. Xie, Z. Qin, and G. Y. Li, "Semantic communication with memory," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2658–2669, 2023.

[6] H. Zhou, Y. Deng, X. Liu, N. Pappas, and A. Nallanathan, "Goal-oriented semantic communications for 6G networks," *IEEE Internet Things Mag.*, vol. 7, no. 5, pp. 104–110, 2024.

[7] W. Xu, Y. Zhang, F. Wang, Z. Qin, C. Liu, and P. Zhang, "Semantic communication for the internet of vehicles: A multiuser cooperative approach," *IEEE Veh. Technol. Mag.*, vol. 18, no. 1, pp. 100–109, 2023.

[8] L. Xia, Y. Sun, D. Niyato, D. Feng, L. Feng, and M. A. Imran, "xURLLC-aware service provisioning in vehicular networks: A semantic communication perspective," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 5, pp. 4475–4488, 2024.

[9] P. Singh, B. Hazarika, K. Singh, W.-J. Huang, and C.-P. Li, "Augmented multi-agent DRL for multi-incentive task prioritization in vehicular crowdsensing," *IEEE Internet Things J.*, pp. 1–1, 2024.

[10] H. Zhang, H. Wang, Y. Li, K. Long, and V. C. M. Leung, "Toward intelligent resource allocation on task-oriented semantic communication," *IEEE Wirel. Commun.*, vol. 30, no. 3, pp. 70–77, 2023.

[11] 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.

[12] 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.*, pp. 1–1, 2024.

[13] D. Cao, J. Zhao, W. Hu, F. Ding, Q. Huang, and Z. Chen, "Attention enabled multi-agent DRL for decentralized volt-VAR control of active distribution system using PV inverters and SVCs," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1582–1592, 2021.

[14] 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.

[15] B. Wang and C.-C. J. Kuo, "Sbert-wk: A sentence embedding method by dissecting bert-based word models," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 28, pp. 2146–2157, 2020.

[16] Technical Specification Group Radio Access Network, "Study on LTE-based V2X services (release 14)," 3rd Generation Partnership Project (3GPP), 3GPP TR 36.885 V14.0.0, jun 2016, available: https://www.3gpp.org.

[17] P. Koehn, "Europarl: A parallel corpus for statistical machine translation," in *Proceedings of machine translation summit x: papers*, 2005, pp. 79–86.