Spectral–Topology-Aware KG-MARL for 5G V2V Sidelink

Vehicular-to-Vehicular (V2V) communication is a cornerstone for enabling cooperative safety, real-time traffic management, and autonomous driving. A key enabler in the 5G NR standard is sidelink mode-2, where vehicles autonomously select transmission resources without centralized scheduling. While this approach ensures scalability, its baseline mechanism—Semi-Persistent Scheduling (SPS)—underperforms in dense or highly mobile environments. SPS suffers from slow reselection, hidden-terminal collisions, and lacks adaptability to diverse QoS requirements such as latency, reliability, and throughput. These shortcomings compromise safety-critical applications where packet reception ratio (PRR), low delay, and high reliability are crucial.

To address these limitations, we propose a **Koopman-augmented Graph Multi-Agent Reinforcement Learning (KG-MARL)** framework for decentralized V2V sidelink resource allocation. Unlike SPS, KG-MARL empowers each vehicular link to dynamically select both its subchannel and transmit power using a richer representation of the environment. The framework combines: (i) *spectrogram-based spectral maps* via short-time Fourier transform (STFT), capturing temporal and frequency-domain interference dynamics; (ii) *Graph Attention Network (GAT) embeddings*, modeling the interference topology among neighboring links; and (iii) *Koopman operator-based prediction*, which linearizes nonlinear state dynamics to enable stable and sample-efficient prediction of short-horizon interference evolution. Each agent optimizes a reward shaped as a potential game, aligning local and global objectives. The per-link reward (utility) is The per-link utility is $R_i = U_i = \alpha\,\mathrm{PRR}_i + \beta\log(1+\mathrm{SINR}_i) - \gamma\,\mathrm{Int}_i - \lambda P_i, \quad \mathrm{SINR}_i = \frac{P_i g_{ii}}{N_0+\sum_{j\neq i,\, r_j=r_i} P_j g_{ji}}$, where PRR_i is the packet reception ratio, Int_i the measured interference, P_i the transmit power, g_{ii} and g_{ji} the desired and interfering channel gains, N_0 the noise power, and $\alpha, \beta, \gamma, \lambda$ weighting factors for reliability, spectral efficiency, interference mitigation, and power cost.

The framework follows a Soft Actor–Critic (SAC)-style actor–critic architecture with centralized training and decentralized execution. Koopman operators accelerate value updates by approximating state transitions linearly, while GAT embeddings enhance coordination via graph-structured observations. Once trained, vehicles execute decisions autonomously with minimal overhead, ensuring practicality for real deployment.

Algorithm 1: KG-MARL Training

1: Initialize actor π_θ, critic Q_ϕ, GAT f_ψ, Koopman $\mathcal{K}$, replay buffer $\mathcal{D}$
2: **for** each episode and frame **do**
3: Agents sense spectrum $\rightarrow$ STFT heatmap
4: Build interference graph $\rightarrow$ GAT embedding z_i
5: Form state s_i, sample action $a_i = (r_i, P_i) \sim \pi_\theta$
6: Execute, observe reward R_i, next state s'_i, store in $\mathcal{D}$
7: Update step: Koopman prediction $\Phi(s') \approx \mathcal{K}\Phi(s)$
8: Update critic with target; update actor (SAC objective); update GAT and $\mathcal{K}$

Table 1: Results

Scheme	PRR	Collisions	Avg. Power
KG-MARL	**93–95%**	**15%**	**0.22 W**
GAT-A2C	90%	18%	0.25 W
DIRAL	88%	22%	0.30 W
DQN	85%	25%	0.30 W
SPS	75–85%	30%	0.30 W

Simulations show that KG-MARL consistently outperforms all considered baselines, as summarized in Table I. The comparisons include Semi-Persistent Scheduling (SPS), which is the standard 3GPP mode-2 mechanism; Deep Q-Network (DQN), a single-agent reinforcement learning method for resource allocation; Distributed Resource Allocation using Multi-Agent Reinforcement Learning (DIRAL), a decentralized MARL-based approach; and Graph Attention Network–based Advantage Actor–Critic (GAT-A2C), which leverages graph neural representations for policy learning. Against these benchmarks, KG-MARL achieves notable gains: it improves median SINR by 5-7 dB, halves collision probability in dense scenarios, sustains packet reception ratio (PRR) above 90–95% for safety-critical messages, and reduces average transmit power by 20–25%.

References

[1] Ye, Hao, Geoffrey Ye Li, and Biing-Hwang Fred Juang. "Deep reinforcement learning based resource allocation for V2V communications." IEEE Transactions on Vehicular Technology 68.4 (2019): 3163-3173.

[2] Liang, Le, Hao Ye, and Geoffrey Ye Li. "Spectrum sharing in vehicular networks based on multi-agent reinforcement learning." IEEE Journal on Selected Areas in Communications 37.10 (2019): 2282-2292.