Mobility-Aware Sensing Data Orchestration for Communication-Efficient Cooperative Perception

Pengjun Hou*, Yida Ding*, Mushu Li*, Yu Cheng†
*Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA 18015
†Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA 60616
Email: {peh324, yid324, mul224}@lehigh.edu, cheng@illinoistech.edu

***Abstract*—Cooperative perception enables connected autonomous vehicles (CAVs) to enhance environmental awareness by sharing multi-view sensing data, but its practical deployment is constrained by extensive communication overhead and dynamic vehicle mobility. In this paper, we address a critical problem for enabling cooperative perception in practice: how to orchestrate sensing data sharing efficiently under limited communication resources and spatio-temporally varying sensing environments. To tackle this problem, we propose a novel two-stage sensing data sharing framework. The long-timescale stage plans the data-sharing budget by balancing perception performance against communication cost, while the short-timescale stage performs fine-grained sensing data selection that adapts to rapid changes in road environments and vehicle mobility. We further develop a deep reinforcement learning (DRL)-empowered scheme that derives an index-based sensing data selection policy, enabling prompt and scalable decision making for real-time cooperative perception. Extensive experiments on real-world driving datasets and vehicle mobility traces demonstrate that the proposed framework substantially improves cooperative perception quality for CAVs operating under communication constraints.**

***Index Terms*—cooperative perception, connected autonomous vehicles, deep reinforcement learning, sensing data sharing.**

I. Introduction

Connected autonomous driving is transforming intelligent transportation systems by integrating advanced sensing and communication technologies to enhance safety, efficiency, and driving comfort [1]–[3]. A key enabler of this paradigm is *cooperative perception*, which refers to the collaborative process where connected autonomous vehicles (CAVs) share and fuse their locally sensed data through vehicle-to-everything (V2X) communications. By combining multi-view information, cooperative perception overcomes individual sensing limitations such as occlusions and restricted ranges, thereby improving overall environmental awareness [4].

While cooperative perception is promising for enabling comprehensive environmental understanding in vehicles, it faces practical challenges arising from excessive communication overhead. Safety-critical driving applications impose strict latency and reliability requirements, yet the exchange of large volumes of sensing data places tremendous pressure on limited communication bandwidth [5]. Recent advances in sensing data segmentation and multi-view fusion techniques [6] have enabled segmented sensing data sharing, opening new opportunities to further reduce communication redundancy. Building on these capabilities, prior research has explored selective data-sharing strategies to balance cooperative perception performance against communication costs. However, many existing approaches rely on instantaneous sensing status [7] or perform frame-level selection without considering the semantic importance of sensing content [4], [8]. Several challenges still remain that hinder communication-efficient cooperative perception. First, the optimal amount of data to share, which balances perception accuracy and communication delay, varies dynamically with the overlap between vehicles' sensing ranges, and computing this quantity in real time is computationally expensive. Second, identifying which sensing data segments are most valuable to neighboring vehicles is nontrivial [9], as their importance evolves with vehicle mobility and changing road environments. A segment that seems unimportant now may become critical moments later.

In this paper, we investigate a communication-efficient sensing data sharing method for cooperative perception that adapts to vehicle mobility and dynamic sensing environments. The objective is to maximize cooperative perception performance, i.e., the perception quality of vehicles, by jointly optimizing the amount of sensing data to share and identifying the segments to be transmitted. We develop a two-stage sensing data sharing framework that operates over different timescales: the long-timescale stage plans the overall data-sharing budget by balancing perception performance against communication overhead, and the short-timescale stage performs fine-grained sensing data selection while adapting to spatio-temporal dynamics. The contributions of this paper are twofold. (1) The proposed sensing data sharing framework effectively enhances perception quality in cooperative perception while balancing sensing performance and communication overhead. (2) We propose a novel deep reinforcement learning (DRL) scheme to obtain an index-based sensing data selection policy that efficiently prioritizes high-value sensing segments while rapidly adapting to dynamic sensing environments.

II. System Model

The cooperative perception scenario for CAVs is illustrated in Fig. 1, where CAVs travel alongside human-driven vehicles and continuously sense their surroundings with embedded RGB cameras for visual perception in an urban road environment. Due to the limited sensing range and potential

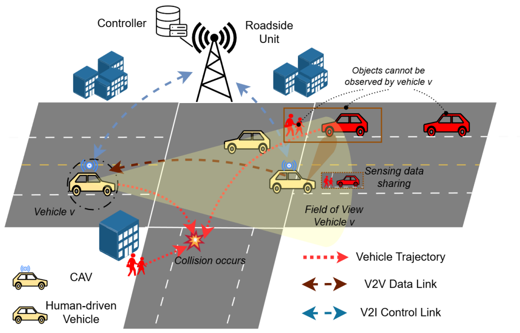


Fig. 1: The network scenario with cooperative perception

occlusions caused by surrounding objects, the field of view (FoV) of each camera is constrained. We refer to CAVs as vehicles for the rest of the paper. To extend the sensing range and enhance perception reliability, vehicles within a certain proximity form a vehicle cluster and broadcast partial sensing data (e.g., segmented frames) using shared communication resources through vehicle-to-vehicle (V2V) data links. Concurrently, each vehicle fuses the received sensing data from nearby vehicles with its locally captured data to achieve cooperative perception for object detection. As shown in Fig. 1, target vehicle v cannot sense the red cars and pedestrians since they lie outside its FoV or are occluded. Receiving sensing data from other vehicles will avoid potential collisions. A roadside unit (RSU) covering vehicle cluster coordinates the inter-vehicle communications to ensure efficient sensing data exchange through vehicle-to-infrastructure (V2I) control links. Denote $\mathcal{V} = \{1, \ldots, V\}$ as the set of vehicles in a cluster. We study the sensing data scheduling for a single vehicle cluster.

A. Perception Quality Model

We consider a time-slotted system in which time slots are indexed by t. In each time slot, vehicles capture their surroundings and share sensing data in a single round, where the duration of a time slot is denoted by ϕ. The FoV of vehicle $v \in \mathcal{V}$ at time slot t is segmented into a set of subregions $\mathcal{I}^F_{v,t}$, where each subregion $i \in \mathcal{I}^F_{v,t}$ represents a spatial segment within the vehicle's camera FoV. We use *perception confidence* to evaluate the perception quality of each subregion. The confidence value can be obtained from common object detection algorithms (e.g., YOLO [10]), quantifying the detector's belief that a given subregion contains a real object and that its predicted location is accurate. Denote by $\hat{r}_{v,i,t}$ the perception confidence of vehicle v for subregion i at time slot t. The collection of perception confidence values across all subregions captured by a vehicle in a time slot is referred to as the vehicle's *confidence map* (CM). In dynamic road environments, a vehicle's effective FoV may be limited, e.g., due to occlusions, road geometry, or high driving speeds. To capture the areas most relevant for safe navigation, we define the region of interest (RoI) for each vehicle as $\mathcal{I}^R_{v,t}$, representing the set of subregions crucial to the driving status of vehicle v at time slot t. By definition, $\mathcal{I}^F_{v,t} \subseteq \mathcal{I}^R_{v,t}$.

The perception quality of a vehicle for a given subregion is determined by three factors: 1) the local perception of vehicle v from its currently captured sensing data; 2) the sensing data received from other vehicles; and 3) historical sensing data, whose usefulness decays as it becomes outdated over time. Accordingly, after fusing these three sources of information, the perception quality of vehicle v for subregion i at time slot t is defined as

$$r_{v,i,t} = \max\{\hat{r}_{v,i,t}, e^{-\beta} r_{v,i,t-1}, \max_{v' \in \{\mathcal{V} \setminus v\}} U_{v',i,t} e^{-\beta \Delta_t} \hat{r}_{v',i,t}\}, \quad (1)$$

where β is a discount factor representing the degree of information staleness over a time slot, Δ_t denotes the time interval between when vehicles broadcast their sensing data and when the fused data is received, and $U_{v',i,t}$ is an indicator variable denoting whether vehicle v' uploads subregion i at time slot t, where $U_{v',i,t} = 1$ if the subregion is uploaded and $U_{v',i,t} = 0$ otherwise. To ensure effective cooperative sensing, Δ_t should be shorter than the duration of a time slot.

B. RSU Coordination Workflow

To avoid redundant sensing data sharing, the RSU selects which subregions should be broadcast and allocates the corresponding communication resources. This process requires vehicles to first update their status for coordination. Each vehicle updates its current local CM and location. The RSU maintains the historical CMs and trajectories of each vehicle over the preceding W time slots, as well as each vehicle's RoI. Using this information, the RSU evaluates the importance of all subregions covered by the vehicles' FoVs through a *score map evaluation function*. Subregions whose sensing data yield higher improvements in overall perception quality across vehicles are assigned higher importance scores. The resulting importance distribution over all subregions forms a *score map*, which indicates their priority in the data-sharing process. Based on the score map, the RSU selects the top-K_t subregions and instructs the corresponding vehicles, which capture each selected subregion with the highest confidence in their updated CMs, to broadcast the associated sensing data to others. Thus, the RSU determines $U_{v,i,t}$ in Eq. (1) for all vehicles at time slot t. Furthermore, the RSU also executes a *cooperative perception planning function* that determines the number of subregions whose sensing data should be shared in each time slot, denoted by K_t. This function balances the communication overhead of cooperative perception against the improvement in perception quality. Note that K_t is updated once in several time slots to accommodate significant changes in the sensing environment.

C. Communication Model

As shown in Eq. (1), the fused perception quality is directly affected by the delay associated with receiving and processing sensing data, denoted by Δ_t. A fixed V2I band is used for uploading local CMs from vehicles and distributing coordination decisions from the RSU. The corresponding delay, together with the time for generating coordination decisions, is denoted by τ^{CM}_t and is treated as a constant. In addition, to obtain the delay for broadcasting the selected subregions' sensing data within the vehicle cluster, denoted by τ_t^{SD}, the transmission rate for vehicle v broadcasting its sensing data is:

$$R_v = B\log_2(1+\frac{P_v}{\sigma^2}), \tag{2}$$

where P_v is the received transmission power from vehicle v. We assume identical received power for all vehicles due to the short transmission distance within a cluster. In (2), σ^2 denotes the noise power, and B represents the available frequency bandwidth for data transmission. We adopt a time-division multiplexing scheme for sharing sensing data among vehicles in a cluster, where transmission time slots are allocated to vehicles in proportion to the number of subregions selected for sharing. Accordingly, the transmission delays required to receive all shared data within one time slot can be derived as:

$$\tau_t^{SD} = \sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}_{v,t}^R} SU_{v,i,t}/R_v, \tag{3}$$

where S is the size of the sensing data for each subregion. The total number of selected subregions shared among vehicles in slot t, i.e., $\sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}_{v,t}^R} U_{v,i,t}$, must be no greater than K_t.

III. Problem Formulation

We aim to determine an optimal strategy for selecting subregions to broadcast over time to maximize the long-term average perception quality across all vehicles. The optimization problem is formulated as follows:

$$\begin{aligned}\mathbf{P}_0: \max_{\{K_t,U_{v,i,t},\forall v,i,t\}} & \frac{1}{T}\frac{1}{|\mathcal{V}|}\sum_{t=1}^{T}\sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}_{v,t}^R} r_{v,i,t} \qquad (4)\\ \text{s.t.} & \sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}_{v,t}^R} U_{v,i,t}\le K_t, \forall t;\\ & \Delta_t = \tau_t^{CM}+\tau_t^{SD}\le\phi, \forall t;\\ & \sum_{v\in\mathcal{V}} U_{v,i,t}\le 1, \forall i,t;\\ & U_{v,i,t}=\{0,1\}, \forall v,i,t;\\ & (1)-(3).\end{aligned}$$

The optimization problem is challenging due to several intrinsic complexities. The underlying environment is highly dynamic, with future vehicle states being time-varying and unknown a priori. In addition, decisions made in preceding time slots have long-term effects on future system performance, introducing strong temporal dependencies that complicate long-horizon optimization. Moreover, the decision space of $U_{v,i,t}$ expands combinatorially with the number of vehicles and subregions, rendering traditional RL models computationally infeasible for exhaustive exploration.

IV. Two-Stage Sensing Data Sharing Framework

To address problem $\mathbf{P}_0$, we develop a two-stage optimization method. The first stage, referred to as cooperative perception planning, determines the number of subregions to be shared in each time slot, i.e., K_t, on a larger timescale. Given the reduced decision space, the second stage derives a subregion selection strategy that adapts to spatio-temporal mobility and environmental dynamics on a smaller timescale.

A. Stage 1: Cooperative Perception Planning

Increasing the number of subregions for which sensing data should be shared, i.e., K_t, can substantially improve overall perception quality by leveraging richer multi-view information. However, transmitting excessive sensing data increases communication latency under limited bandwidth, ultimately degrading the effectiveness of cooperative perception.

Lemma 1. *A unique global maximizer of K_t exists in Problem $\mathbf{P}_0$ under stationary environmental dynamics, assuming that all subregions have equal average selection priority over time.*

Proof. Given stationary environmental dynamics, the value of K_t at each time slot is independent and unaffected by previous decisions. Therefore, we can determine K_t with one decision window and drop the index t for this optimization. Let λ_R and λ_F denote the probabilities that a subregion i lies within a vehicle's RoI and FoV, respectively. Correspondingly, let $P_{i,n}^R$ and $P_{i,n}^F$ denote the probabilities that subregion i is covered by the RoIs and the FoVs of n vehicles, respectively, which follow the binomial distributions $\mathcal{B}(n,\lambda_R)$ and $\mathcal{B}(n,\lambda_F)$.

Given that subregion i is covered by the RoIs of n vehicles and that the number of selected subregions is K, the average perception quality of subregion i over the decision window associated with K can be estimated, based on Eq. (1), as

$$\begin{aligned}\mathbb{E}_t[r_i|K,n] =& p_{i|K,n}(1-P_{i,0}^F)\mathbb{E}_t[\max\{\kappa_{v,i},\epsilon_{v,i}\}]+\\ & [1-p_{i|K,n}(1-P_{i,0}^F)]\mathbb{E}_t[\epsilon_{v,i}],\end{aligned} \tag{5}$$

In Eq. (5), $p_{i|K,n}$ represents the probability that subregion i is selected for sensing data sharing given that n vehicles interest the subregion. The probability can be modeled as $p_{i|K,n}=\sum_{k=1}^{K} q_{i,k}$, where $q_{i,k}$ denotes the probability that, when subregions are ranked by their selection priority, subregion i appears in the k-th position. Under our assumption, $q_{i,k}$ is identical for all k. Furthermore, if a subregion is selected and at least one vehicle's FoV covers it, the sensing data with the highest local perception quality among those vehicles will be broadcast, yielding a perception quality $\kappa_{v,i}=\max\{\hat{r}_{v,i}, e^{-\beta\Delta}\max_{\{v'\in\mathcal{V},v'\neq v\}}\hat{r}_{v',i}\}$. Otherwise, each vehicle relies on its own local perception and compares it with its historical perception information, resulting in a quality $\epsilon_{v,i}=\max\{\hat{r}_{v,i}, e^{-\beta}r_{v,i,t-1}\}$. Therefore, the objective function in Eq. (4) can be rewritten as

$$\mathbf{P}_1: \max_K \frac{1}{|\mathcal{V}|}\sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}_{v,t}^R}\sum_{n=0}^{|\mathcal{V}|} P_{i,n}^R\mathbb{E}_t[r_i|K,n]. \tag{6}$$

It can be further simplified as follows:

$$\max_K F(K)=\frac{1}{|\mathcal{V}|}\sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}_{v,t}^R}\alpha_i(\sum_{k=1}^{K} q_{i,k})\mathbb{E}_t[\max\{\kappa_{v,i}-\epsilon_{v,i},0\}],$$

where $\alpha_i = P_{i,n}^R(1-P_{i,0}^F)$ and it remains constant for different values of K. Furthermore, Δ can be simplified to $\tau^{CM}+(S/R_v)K$, and $\max\{\kappa_{v,i}-\epsilon_{v,i},0\} = \max\{0,(e^{-\beta\Delta}\max_{v'\in\{\mathcal{V}\setminus v\}}\hat{r}_{v',i})-\epsilon_{v,i}\}$.

By relaxing K to a continuous variable to make $F(K)$ differentiable, the derivative of $F(K)$ is given in Eq. (7a), where $\mathbf{1}_{\kappa_{v,i}>\epsilon_{v,i}}$ indicates $\mathbb{E}_t[r_i|K,n] \neq 0$. When $\nabla_K F(K) = 0$, Eq. (7a) can be simplified to Eq. (7b). The left-hand side (LHS) of this equation is monotonically decreasing in K because $q_{i,K}$ is identical for all K under the stated assumption. In contrast, the right-hand side (RHS) is monotonically increasing, since the probability that a subregion is selected grows as K increases. Therefore, if the optimal solution lies within the range $K^* \in (0, \max_{v,i} \frac{R_v}{S}(1 - \beta^{-1} \ln \frac{r_{v,i,t-1}}{\max_{v' \in \{\mathcal{V}\setminus v\}} \hat{r}_{v',i}}))$ (which follows from the condition $\kappa_{v,i} > \epsilon_{v,i}, \exists (v,i)$), then there exists a unique stationary point K^* such that $\nabla_K F(K^*) = 0$. Furthermore, when the LHS exceeds the RHS, i.e., LHS $>$ RHS, we have $\nabla_K F(K) > 0$ for $K < K^*$. Conversely, when LHS $<$ RHS, we have $\nabla_K F(K) < 0$ for $K > K^*$. Thus, there exists only one maximizer within the range. Otherwise, $K^* = 0$, which occurs when 1) the local perception quality is sufficiently high such that cooperation yields negligible improvement, or 2) the delay introduced by cooperative sensing is so large that it provides no benefit to any vehicle in any subregion. □

Over a sufficiently long horizon between updates of K_t, vehicle mobility and dynamic obstacle positions cause the selection likelihood of all subregions to homogenize, thereby justifying the assumption in Lemma 1. The closed-form solution for K_t can be obtained from (7b) using techniques such as fixed-point analysis. To reduce the complexity of deriving parameters (e.g., λ_R^n and λ_F^n), a simple heuristic tuning procedure can be adopted: periodically increase K_t until the perception performance begins to decline.

B. Stage 2: Dynamic Subregion Selection

Selecting the K_t most important subregions within each time slot to maximize the average perception quality for all vehicles can be naturally formulated as a Restless Multi-Armed Bandit (RMAB) problem. Specifically, RMAB generalizes the classical Multi-Armed Bandit (MAB) model by allowing the state of each arm (i.e., subregion in our case) to evolve over time even when the arm is not selected. Additionally, unlike the traditional MAB setting where only a single arm can be activated at each decision step, RMAB permits multiple arms to be scheduled simultaneously. The corresponding RMAB problem can then be reformulated from Problem $\mathbf{P}_0$ as follows:

$$\mathbf{P}_2: \max_{\{\mu_{i,t}, \forall i,t\}} \frac{1}{T}\frac{1}{|\mathcal{V}|}\sum_{t=1}^{T}\sum_{i\in\mathcal{I}_t} R(s_{i,t},\mu_{i,t}) \tag{8a}$$

$$\text{s.t.} \sum_{i\in\mathcal{I}_t} \mu_{i,t} \le K_t, \forall t; \mu_{i,t} = \{0,1\}, \forall i,t; \tag{8b}$$

$$R(s_{i,t},1) = |\mathcal{V}_i|(r_{i,t}^* - \bar{r}_{i,t}); \tag{8c}$$

$$R(s_{i,t},0) = \max\{|\mathcal{V}_i|(e^{-\beta} r_{i,t-1}^* - \bar{r}_{i,t}), 0\}; \tag{8d}$$

$$(1)-(3),$$

where the set $\mathcal{I}_t$ represents all subregions within the RoIs of vehicles in the cluster, i.e., $\mathcal{I}_t = \cup_v \mathcal{I}_{v,t}^R$. The optimization variable in Problem $\mathbf{P}_2$, denoted by $\mu_{i,t}$, indicates whether subregion (arm) i is selected at time slot t. If subregion i is chosen for broadcasting, then $\mu_{i,t} = 1$, and the RSU selects the vehicle v^* with the highest confidence value for that subregion to upload its sensing data. In this case, $U_{i,v^*,t} = 1$. Otherwise, if the arm is not activated, then $\mu_{i,t} = 0$ and $U_{i,v,t} = 0$ for all v. In Problem $\mathbf{P}_2$, $s_{i,t}$ denotes the state of subregion i at time slot t, which consists of three components: 1) $\mathbf{V}_i$: The set of vehicles whose RoI includes subregion i in past time W slots, where $\mathbf{V}_i = \{\mathcal{V}_{i,t-W+1}, \cdots, \mathcal{V}_{i,t}\}$ and $\mathcal{V}_{i,t} = \{v | i \in \mathcal{I}_{v,t}^R\}$; 2) $\mathbf{r}_{i,t}^*$: The maximum perception confidence value of subregion i in past time W slots, where $\mathbf{r}_{i,t}^* = \{r_{i,t-W+1}^*, \cdots, r_{i,t}^*\}$ and $r_{i,t}^* = \max_{\{v \in \mathcal{V}_i\}} r_{i,v,t}$; and 3) $\bar{\mathbf{r}}_{i,t}$: The average perception confidence value of subregion i in past time W slots, where $\bar{\mathbf{r}}_{i,t} = \{\bar{r}_{i,t-W+1}, \cdots, \bar{r}_{i,t}\}$ and $\bar{r}_{i,t} = \frac{1}{|\mathcal{V}_i|}\sum_{v\in\mathcal{V}_i} \hat{r}_{i,v,t}$. The function $R(s_{i,t},\mu_{i,t})$ represents the total perception quality improvement of subregion i at time slot t, given the current state and the selection decision.

To solve Problem $\mathbf{P}_2$, we propose an index-based scheduling scheme derived from Whittle's index policy, which assigns an index to each arm to quantify the value of activating it in the current state. The Whittle index policy is widely used for addressing complex sequential RMAB problems, as it circumvents the intractability of jointly considering the combined state space of all arms. Accordingly, Constraint (8b) in Problem $\mathbf{P}_2$ is relaxed into a stochastic constraint:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i\in\mathcal{I}_t} \mu_{i,t} = \frac{1}{T}\sum_{t=1}^{T} K_t = \bar{K}_t, \quad \mu_{i,t} \in \{0,1\}. \tag{9}$$

We then introduce a service charge that is incurred whenever an arm takes an active action. Let $C_{i,t}$ denote this service charge, i.e., the Whittle index, for subregion i, which can be obtained from the Lagrangian dual objective function:

$$\max_{\{\mu_{i,t}, C_{i,t}\}} \frac{1}{T|\mathcal{V}|}\sum_{t=1}^{T}\sum_{i\in\mathcal{I}_t}\Big[R(s_{i,t},\mu_{i,t}) - C_{i,t}\mu_{i,t}\Big] + C_{i,t}K_t. \tag{10}$$

After neglecting the constant term and decoupling the problem across subregions, the problem can be simplified as:

$$\max_{\{\mu_{i,t}, C_{i,t}\}} \frac{1}{T}\frac{1}{|\mathcal{V}|}\Big[R(s_{i,t},\mu_{i,t}) - C_{i,t}\mu_{i,t}\Big]. \tag{11}$$

The optimal service charge for each subregion is obtained by ensuring that the long-term costs of taking the active and passive actions (i.e., selecting or not selecting a subregion) are equal. By incorporating the long-term cost difference through a Bellman equation (details omitted due to space limitations), the resulting closed-form Whittle index is:

$$C_{i,t} = \underbrace{(r_{i,t}^* - e^{-\beta} r_{i,t-1}^*)\Big[|\mathcal{V}_{i,t}|}_{\text{Term 1}} + \underbrace{\sum_{l=1}^{T-t}\Big(e^{-\beta l}|\mathcal{V}_{i,t+l}|\prod_{j=1}^{l} P_{i,t+j}\Big)\Big]}_{\text{Term 2}},$$

where $P_{i,t}$ denotes the probability that subregion i is not selected for cooperative-perception uploading at time slot t, following the Whittle index policy. Term 1 in the Whittle index captures the immediate perception gain of subregion i at the current time slot relative to the previous one; a

$$\nabla_K F(K) = \frac{1}{|\mathcal{V}|}\sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}^R_{v,t}}\sum_{n=0}^{|\mathcal{V}|}\alpha_i \mathbf{1}_{\kappa_{v,i}>\epsilon_{v,i}}\{q_{i,K}[(e^{-\beta\Delta}\max_{\{v'\in\mathcal{V},v'\neq v\}}\hat{r}_{v',i})-\epsilon_{v,i}]-(\sum_{k=1}^{K} q_{i,k})(\frac{S}{R_v}e^{-\beta\Delta}\max_{\{v'\in\mathcal{V},v'\neq v\}}\hat{r}_{v',i})\}; \quad (7a)$$

$$\sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}'}\sum_{n=0}^{|\mathcal{V}|} q_{i,K}[1-\epsilon_{v,i}e^{\beta\Delta}(\max_{\{v'\in\mathcal{V},v'\neq v\}}\hat{r}_{v',i})^{-1}] = \sum_{v\in\mathcal{V}}\sum_{i\in\mathcal{I}'}\sum_{n=0}^{|\mathcal{V}|}\frac{S}{R_v}\int_{k=0}^{K} q_{i,k}, \text{ where } \mathcal{I}'=\{i|i\in\mathcal{I}^R_{v,t},\kappa_{v,i}>\epsilon_{v,i}\}. \quad (7b)$$

larger gain indicates a stronger necessity to broadcast this subregion at the current time. Term 2 represents the discounted cumulative expected number of vehicles that will be interested in this subregion in future time slots. When more vehicles are expected to rely on this subregion (i.e., larger $|\mathcal{V}_{i,t+l}|$) and the probability of the subregion not being uploaded in the future is high (i.e., larger $P_{i,t+j}$), uploading the subregion at the current time becomes increasingly valuable. The Whittle indexes for all subregions in a time slot form a score map in the RSU.

C. ST-GAT-Empowered Sensing Data Selection

Although a closed-form expression for the Whittle index is derived, the future vehicle mobility and the state-action transition probabilities remain unknown. To address this challenge, we develop a DRL-empowered scheme to estimate the Whittle index $C_{i,t}$. We denote the term related to future information in the Whittle index by $f(s_{i,t})$, where

$$f(s_{i,t}) = |\mathcal{V}_{i,t}| + \sum_{l=1}^{T-t}\left(e^{-l\beta}|\mathcal{V}_{i,t+l}|\prod_{j=1}^{l}P_{i,t+j}\right). \quad (12)$$

By expanding the above mathematical formulation, we can obtain its Bellman equation form as follows:

$$f(s_{i,t}) = |\mathcal{V}_{i,t}| - (1-P_{i,t+1}e^{-\beta})|\mathcal{V}_{i,t+1}| + f(s_{i,t+1}). \quad (13)$$

Therefore, $C_{i,t}$ can be reformulated as follows:

$$C_{i,t} = (r^*_{i,t} - e^{-\beta}r^*_{i,t-1})f(s_{i,t}). \quad (14)$$

We then approximate $f(s_{i,t})$, referred to as the value function via a DRL-empowered scheme, thereby obtaining real-time Whittle indexes for all subregions covered by the vehicles' RoIs. Under the new form of the Bellman equation, the reward of subregion i at each decision step is $R_{i,t} = |\mathcal{V}_{i,t}| - (1 - P_{i,t+1}e^{-\beta})|\mathcal{V}_{i,t+1}|$. The action is the selection of subregions based on the resulting Whittle indexes.

To handle the large state space, we adopt a neural network for approximating the value functions for all subregions from aggregated state $\mathbf{s}_t = \{s_{i,t}, \forall i \in \mathcal{I}_t\}$ due to their identical structure. Inspired by DQN, we use two neural networks, i.e., an evaluation network and a target network, to learn the value function. The loss function can be defined as follows:

$$L(\theta) = \frac{1}{|\mathcal{I}_t|}\sum_{i\in\mathcal{I}_t}\left\|\hat{f}(s_{i,t};\theta) - \left(R_{i,t} + \hat{f}(s_{i,t};\theta')\right)\right\|_2^2, \quad (15)$$

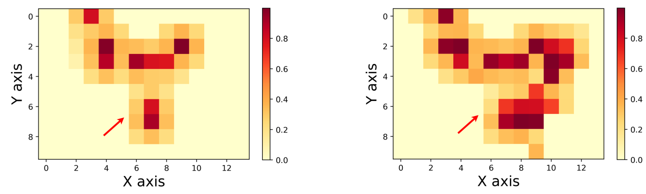
where $\hat{f}(s_{i,t},\theta)$ represents the estimated value function obtained using weights θ in the evaluation network, and $\hat{f}(s_{i,t};\theta')$ represents the estimated value function obtained using weights θ' in the target network. Furthermore, since each subregion is associated with vehicle positions and vehicles mutually influence one another over time, the system naturally forms a temporal graph structure. Therefore, we adopt a spatio-temporal graph attention network (ST-GAT) [11] as the backbone of our DQN architecture to jointly capture spatial and temporal dependencies and effectively learn the value function $f(s_{i,t})$. This design also enables the neural network to differentiate among subregions and learn the value function for each subregion individually. The overall DRL-empowered scheme is summarized in Alg. 1. To reduce the large state space, we employ a variational autoencoder (VAE), denoted by model $\mathcal{E}$, to extract compact latent features from the CMs and the distribution of sensing interests (i.e., $\mathcal{V}_{i,t}, \forall i$) in state $s_{i,t}$. In Steps 4-6, these latent features from the past W time slots are used to construct a latent feature graph, where edges are defined based on pairwise inter-vehicle distances. In Steps 7-8, the ST-GAT-empowered DRL module outputs the Whittle indexes for sensing data selection. Meanwhile, in Steps 9-14, the ST-GAT evaluation and target networks are trained and updated to adapt to sensing environment dynamics.

Algorithm 1 ST-GAT-Empowered Sensing Data Selection

1: **Input:** K_t and pretrained VAE model $\mathcal{E}$.
2: Initialize the evaluation and target ST-GAT networks (θ, θ'), and the experience replay buffer.
3: **for** each time slot t **do**
4: Obtain $\mathbf{s}_t$ for past W time slots in the RSU data set.
5: Use the VAE $\mathcal{E}$ to construct W latent feature graphs.
6: Feed the latent feature graphs into the evaluation ST-GAT network to approximate value function $\hat{f}(s_{i,t};\theta), \forall i$.
7: Compute Whittle index $C_{i,t}$ for each subregion according to Eq. (14), i.e., score map at slot t.
8: Select top-K_t subregions with the highest $C_{i,t}$ values for cooperative perception.
9: Observe the reward $R_{i,t}$ and next state $\mathbf{s}'_t$.
10: Store $(\mathbf{s}_t, \mathbf{s}'_t, \{R_{i,t}, \forall i\})$ into the replay buffer.
11: Sample a mini-batch transitions from the buffer.
12: Update evaluation network θ by minimizing Eq. (15).
13: Periodically update target network $\theta' \leftarrow \tau(1-\tau)\theta + \tau\theta'$.
14: **end for**

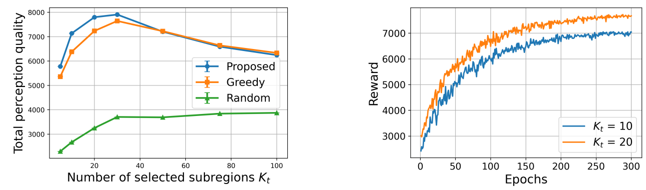
V. Simulation Results

We conducted simulation experiments using the V2X-Sim-2.0-mini dataset [12]. The scenario consists of a cluster of 5 vehicles equipped with multiple sensors, operating in an urban environment that includes pedestrians, cyclists, and non-autonomous vehicles. The simulation lasts for 20 seconds, with each time slot set to 200 ms. To simplify the setup, only the front-view RGB camera data are utilized. Object detection is performed using Faster R-CNN [13], which generates the CMs



(a) Single perception. (b) Cooperative perception.

Fig. 2: The comparison of perception quality for a vehicle.



(a) Perception quality vs. K_t. (b) Model training convergence.

Fig. 3: Simulation results of the proposed DRL approach.

for each vehicle and serves as the basis for our evaluation. Each vehicle's RoI is aligned with its FoV and defined as a square area of 64 m×64 m centered on the vehicle. Each RoI is divided into subregions of size 4 m×4 m. The total available radio spectrum bandwidth is $B = 50\text{MHz}$.

The perception quality of a vehicle at a given time slot under single-vehicle perception and cooperative perception is illustrated in Fig. 2. Each block represents a subregion in a bird's-eye view, with the center block indicating the vehicle's position. The color intensity of each block reflects the perception quality of the corresponding subregion, ranging from low (yellow) to high (red). In the subregion pointed to by the arrow, cooperative perception (Fig. 2b) yields noticeably higher perception quality than single-vehicle (Fig. 2a). This indicates the vehicle alone is unable to accurately or fully detect the objects in the subregion due to limited sensing range or occlusions caused by surrounding obstacles. Cooperative perception effectively compensates for these limitations. The corresponding frames in the original dataset also confirm that a vehicle ahead is indeed blocking the targeted vehicle's view.

The proposed sensing data selection scheme is further compared with two benchmarks: *Greedy*: each vehicle broadcasts the best K_t subregions based solely on the instantaneous distribution of sensing interests and current perception quality; *Random*: each vehicle randomly selects K_t subregions to broadcast. The performance comparison under different values of K_t is shown in Fig. 3a. Our proposed scheme consistently achieves the highest perception quality when K_t is smaller than 50. This is because the proposed scheme selects subregions based on the Whittle index, which captures not only the immediate importance of each sensing region but also its long-term contribution to future perception quality. As K_t increases, the performance of all methods becomes closer to that of the greedy approach due to sufficient broadcast opportunities. However, increasing K_t also leads to higher transmission latency, which in turn degrades the perception quality. The results also demonstrate the existence of a unique optimal K_t that maximizes cooperative perception quality, validating the feasibility of our two-stage framework. In addition, the convergence behavior is illustrated in Fig. 3b. The proposed DRL-empowered scheme converges within fewer than 200 training epochs (100 time slots per epoch), demonstrating the efficiency of the Whittle index-based selection policy.

VI. Conclusion

This paper presented a communication-efficient sensing data sharing framework designed to enhance cooperative perception for CAVs operating under limited wireless bandwidth and dynamic mobility conditions. We developed a two-stage sensing data sharing strategy that jointly determines the optimal amount of data to share and selects the most informative sensing segments, effectively balancing perception performance and communication overhead. The proposed solution demonstrates the strong potential of index-based decision-making for enabling real-time cooperative perception, especially as CAV penetration continues to increase in intelligent transportation systems. For future work, we will explore semantic-based sensing data sharing to further reduce communication overhead while improving cooperative perception efficiency.

References

[1] Y. Han, H. Zhang, H. Li, Y. Jin, C. Lang, and Y. Li, "Collaborative perception in autonomous driving: Methods, datasets, and challenges," *IEEE Intell. Transp. Syst. Mag.*, vol. 15, no. 6, pp. 131–151, 2023.

[2] L. Wang, Y. Zhang, H. Shan, C. Chen, F. Hou, H. Ghafoor, and Y. Cheng, "Performance analysis and optimization of ISAC vehicular networks with 360° radar detection," in *Proc. of 2024 IEEE/CIC Int. Conf. Commun. China (ICCC)*, 2024, pp. 580–585.

[3] P. Wu, X. Yang, P. Wei, and J. Chen, "Safety assured online guidance with airborne separation for urban air mobility operations in uncertain environments," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 19 413–19 427, 2022.

[4] K. Qu, Z. Qin, and W. Zhuang, "Meta learning based adaptive cooperative perception in nonstationary vehicular networks," in *Proc. 2024 Int. Conf. Meta Comput. (ICMC)*, 2024, pp. 84–90.

[5] T. Huang, J. Liu, X. Zhou, D. C. Nguyen, M. Rahimi Azghadi, Y. Xia, Q.-L. Han, and S. Sun, "Vehicle-to-everything cooperative perception for autonomous driving," *Proc. IEEE*, vol. 113, no. 5, pp. 443–477, 2025.

[6] G. Krispel, M. Opitz, G. Waltner, H. Possegger, and H. Bischof, "Fuseseg: Lidar point cloud segmentation fusing multi-modal data," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2020, pp. 1874–1883.

[7] Y. Hu, S. Fang, Z. Lei, Y. Zhong, and S. Chen, "Where2comm: Communication-efficient collaborative perception via spatial confidence maps," in *Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 4874–4886.

[8] M. Li, J. Gao, L. Zhao, and X. Shen, "Adaptive computing scheduling for edge-assisted autonomous driving," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 5318–5331, 2021.

[9] J. Ying, X. Dong, B. Li, and Z. Tian, "Auto-regressive model with exogenous input (ARX) based traffic flow prediction," in *CICTP 2021*, pp. 295–304.

[10] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 7263–7271.

[11] C. Park, C. Lee, H. Bahng, Y. Tae, S. Jin, K. Kim, S. Ko, and J. Choo, "ST-GRAT: A novel spatio-temporal graph attention networks for accurately forecasting dynamically changing road speed," in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manag.*, 2020, pp. 1215–1224.

[12] Y. Li, D. Ma, Z. An, Z. Wang, Y. Zhong, S. Chen, and C. Feng, "V2X-Sim: Multi-agent collaborative perception dataset and benchmark for autonomous driving," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 10 914–10 921, 2022.

[13] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2016.