PLTO: Path Loss-Aware Task Offloading for Vehicular Cooperative Perception

Amr M. Zaki[1], Sara A. Elsayed[1], Khalid Elgazzar[2], Hossam S. Hassanein[1]
[1]School of Computing, Queen's University, Kingston, ON, Canada
[2]IoT Research Laboratory, Ontario Tech University, Oshawa, ON, Canada
amr.zaki@queensu.ca, selsayed@cs.queensu.ca, khalid.elgazzar@ontariotechu.ca, hossam@cs.queensu.ca

***Abstract*—Leveraging task offloading in Vehicular Edge Computing (VEC) via V2X can present unique and robust solutions to the challenges associated with cooperative perception in Autonomous Vehicles (AVs). However, making task offloading decisions that account for the risk of communication failure due to path loss, while adhering to the stringent QoS requirements of cooperative perception has been mostly overlooked. In this paper, we propose PLTO, a Path Loss-Aware Task Offloading scheme that accounts for path loss for Line-of-Sight (LOS), Obstructed LoS (OLoS), and Non-LoS (NLoS) propagation in vehicular communications. We formulate the task offloading problem as a 0-1 Integer Linear Program (0-1 ILP) that aims to minimize the path loss and response delay, while sustaining a certain satisfactory level of improved perception and situational awareness demanded by users. We also propose PLTO-Heuristic (PLTO-H), a scheme to solve the task offloading problem using the MTHG heuristic. Extensive simulations show that PLTO yields significant improvements of up to 17%, 10%, and 23% in terms of packet delivery ratio, Received Signal Strength Indicator (RSSI), and average response delay, respectively, compared to a baseline task offloading scheme that does not consider communication efficiency. In addition, PLTO-H achieves a near optimal solution, with a small gap of up to 6%, 5% and 1.2% in terms of packet delivery ratio, RSSI, and satisfaction ratio, respectively.**

***Index Terms*—Autonomous Vehicles, Vehicular Edge Computing, Cooperative perception, Task Offloading, Path Loss.**

I. Introduction

The wide proliferation of Autonomous Vehicles (AVs) is expected to achieve over 800 billion annual social benefits by 2025, since AVs can help improve road safety, reduce congestion, and decrease energy consumption [1]. A core function for the perception system of an AV is the accurate representation of the driving environment [2], which necessitates the reliance on multiple on-board sensors (LiDAR, radar, and GPS, etc.) [2] [3]. However, inaccurate sensing of the surrounding environment can result in tragic accidents, as have been reported in recent incidents by Tesla [4] and Uber [5]. The vehicle's traffic perception can be improved by aggregating the sensed data from other vehicles [6]. This can be achieved by sharing sensor data among vehicles and exploiting extraneous data from other vehicles to ameliorate the detection capabilities of a single vehicle and enhance traffic situational awareness [3]. This data aggregation benefits a multitude of cooperative vehicular applications, such as collaborative perception [7], collaborative maneuvering [8], and traffic monitoring and route planning [9].

In cooperative vehicular applications, the surging need for collaborative and real-time decision making imposes severe demands on the on-board computing resources to meet the stringent real-time requirements of such applications. Leveraging task offloading in Vehicular edge computing (VEC) via Vehicle-to-everything (V2X) can bring the computing service closer to vehicles and end-users, which can drastically curtail the delay [10], and significantly reduce ineffective transmissions [9].

The advantageous prospect of task offloading in VEC can be hindered by the risk of propagation loss in vehicular communications, which is triggered due to the impact of other vehicles and buildings that act as physical obstructions. Ignoring communication uncertainties (due to path loss, fading and shadowing, packet collision, etc.) can lead to unrealistic assumptions about the performance of cooperative systems [11]. In [12], [13], it has been reported that shadow fading due to Obstructed-Line-of-Sight (OLoS) communication of about 10-20 dB can be induced from a single obstructing vehicle, depending on its shape, size, and location. Moreover, buildings acting as obstacles can induce additional path loss due to Non-Line-of-Sight (NLoS) communication [14]. Mangel et al. [14] show that when two vehicles are approaching an intersection with a speed of 50 km/h, only 30% of all intersection corners provide LoS, while the rest incur NLoS communication. However, accounting for the impact of the different obstacles contributing to path loss, while satisfying the strict requirements of cooperative perception in terms of time and situational awareness, has been mostly overlooked in task offloading in VEC, particularly in condensed urban environments [15].

In this paper, we propose PLTO, a Path Loss-Aware Task Offloading scheme that accounts for path loss due to LoS, OLoS, and NLoS, while adhering to the stringent requirements of the cooperative perception task. The main contributions of this article can be summarized as follows: 1) We propose PLTO which makes task offloading decisions that balance between the time criticality of cooperative perception, the underlying requirement of improved perception awareness, as well as the risk of path loss in vehicular communications. Towards that end, PLTO formulates the task offloading problem as a multi-objective 0-1 Integer Linear Program (0-1 ILP) that jointly minimizes the response delay and the risk of communication failure due to path loss, while ensuring that the benefit gained by users (i.e., the level of additional and non-redundant perception acquired due to perception aggregation) is no less than a certain satisfactory level demanded by users. 2) We propose PLTO-Heuristic (PLTO-H), a scheme that uses the Martello and Toth's heuristic (MTHG) heuristic to solve the task offloading problem in a time-efficient manner to overcome the known complexity of ILP problems. We use PLTO as a baseline for the upper bound on the reachable potential that PLTO-H can achieve. 3) Finally, we conduct extensive evaluations and analysis to demonstrate the inherent limitations and leverages of PLTO. We compare PLTO to a baseline that does not account for communication loss, and we highlight the importance of accounting for the various communication losses when designing a cooperative perception system in VEC.

The remainder of the paper is organized as follows. Section II highlights important related work. Section III describes the proposed schemes (PLTO and PLTO-H). Section IV discusses the performance evaluation and simulation results. Section V presents our conclusions and future directions.

II. Related Work

Task offloading via V2X can be used to foster a broad range of applications. Extending situational awareness through cooperative perception, with the focus of offloading computational tasks to Roadside Units (RSUs) or moving vehicles that act as mobile edge nodes through V2X has recently been studied in multiple works [16]–[19].

Zhu et al. [16] study the task allocation of video applications (with different qualities) to vehicles or RSUs for object detection. They propose a multiobjective scheme that focuses on minimizing latency and quality loss. Krijestorac et al. [17] propose a task offloading scheme that strives to maximize the data transmission rate of vehicles to maximize the number of tasks that can run on vehicles or RSUs. Kazmi et al. [18] present a task offloading scheme that considers the task's deadline requirement, as well as the energy restriction of vehicles when acting as edge nodes. Zaki et al. [19] introduce Cooperative Perception-based Task Offloading (CPTO) as a scheme for cooperative perception using VEC. CPTO aims to both minimize the latency of the task of perception aggregation at the worker while maximizing the level of cooperation between users to enhance the situational awareness of vehicles. It should be noted that maximizing the number of users sharing a worker, would increase the latency of the task of perception aggregation running at this worker. However, their work only considers enhancing the situational awareness through maximizing the number of collaborating users without accounting for perception redundancy that may occur. In addition, optimizing the communication efficiency is not considered in their study.

Lin et al. [20] employ a contextual Multi-armed Bandit approach to minimize the total expected offloading energy consumption, while satisfying stringent delay requirements. They model the channel gain between users and workers due to small-scale fading and path loss to consider the transmission rate of both the uplink and downlink channels. In addition, they utilize a simple metric reflecting the channel state dynamics to train their Multi-armed Bandit scheme. However, making path loss-aware task offloading decisions and optimizing communication efficiency is overlooked in these type of schemes. Moreover, the simulated vehicles in their analysis move in a two lane freeway, thus NLoS path loss is not considered. Zeng et al. [21] propose a QoS-aware task offloading with the aim of enhancing the network resource utilization, while adhering to the delay and energy requirements of users. Their work decomposes the task offloading combinatorial problem into two subproblems. First, a resource block group allocation scheme is proposed to get the effective data transmission rate, then a task offloading scheme based on game theory is employed to reach the best scheduling among multiple edge servers. The channel gain of the wireless channel between users and workers is utilized to model the upper limit of effective data transmission rate. However, the risk of communication failure due to path loss is not considered in the task offloading decision.

In addition to overlooking path loss in the task offloading decision, most schemes also fail to reduce redundancy and disregard the quality of perception. In this paper, we make task offloading decisions that jointly minimize the response delay of cooperative perception and the path loss triggered by various obstacles in urban environments, while abiding by a certain satisfactory level of perception demanded by users to ensure that the gained level of additional and non-redundant perception exceeds a certain customized threshold.

III. Path Loss-Aware Task Offloading (PLTO)

In this section, we provide a detailed description of the the system model, the ILP optimization problem presented in PLTO, and the heuristic utilized in PLTO-H.

A. System Model

Consider a set of users $U = \{u_1, u_2,..., u_n\}$ that are identified as vehicles moving in an urban intersection, and a set of workers $W = \{w_1, w_2, ...w_m\}$ that can execute the cooperative perception tasks offloaded by users. Note that a worker can either be a static RSU or a moving vehicle. Each user $u_i \in U$ subscribes to the service by periodically sending its current location and perception area to the orchestrator. The latter is a centralized entity that is responsible for making the task offloading decision. The orchestrator solicits the computational resources of moving vehicles acting as workers in exchange for some incentives. Once the task offloading decision is made, the orchestrator sends the resulting user-worker association to all users. Each user $u_i \in U$ then sends its cooperative perception task to the designated worker w_j.

The perception task of u_i is denoted $\Psi_i = \{q_i, \lambda_i\}$, where q_i is the computation workload or intensity (in CPU cycles/bit), and λ_i constitutes the size of the perception frame (in bits). Note that cooperative perception, and thus task offloading, is performed periodically over a certain number

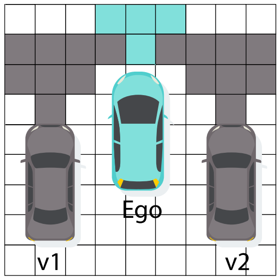


Fig. 1: Triangular rasterization of the ego-vehicle's perception to the projection matrix. Detected awareness is the percentage of the matrix covered by the aggregation of the ego's perception (green) and the perception of the collaborating vehicles (v1 and v2 in gray).

of time steps τ (i.e., rounds of operation). In order to improve the task offloading decision in a given round, we determine the total non-redundant area of perception, denoted a_{ij}, gained by each user u_i due to perception aggregation performed by worker w_j. Towards that end, the perception of each user u_i is rasterized, as shown in Figure 1, to a perception projection matrix denoted M, which represents the covered area of a specific perception range around each ego vehicle.

The non-redundant perception area of u_i, a_{ij}, is defined as the percentage of the covered area by the sensors of user u_i (ego vehicle) and the aggregated result of its association with worker w_j (i.e., the percentage of the area that the green and gray cells constitute in Figure 1). In order to determine a_{ij}, we project the triangular coverage area of a vehicle's sensor to the projection matrix M, and then apply triangular rasterization. Each intersecting cell of the matrix with the triangular coverage area is assigned 1, while the remaining cells are assigned 0. More specifically, the associated worker w_j receives different perceptions from multiple users, where it performs the aggregation task. It then sends the output back to the associated users, hence extending their perception.

Each worker $w_j \in W$ has a maximum CPU frequency, denoted C_j (in CPU cycles/sec). The CPU frequency of worker w_j is divided equally among all the perception tasks offloaded to it. The number of offloaded tasks to worker w_j is the number of vehicles currently using this worker (i.e., cooperating together), and is denoted as perception intensity η_j. The distance between user u_i and worker w_j is denoted d_{ij}, and the propagation speed is denoted v. The data rate of the transmission link is denoted R_{ij}.

An association range, denoted ξ is used to limit the selection of workers to only the ones found in proximity to the ego-vehicle. This association range is selected to be lower than the communication range of DSRC [22], which is the communication technology utilized throughout our experiments. If a user moves outside the association range, it needs to offload to a closer worker.

The total response delay of the perception aggregation task of user u_i on worker w_j is denoted t_{ij}, and is given by Eq. 1. It is composed of the computation latency α_{ij}, propagation latency β_{ij}, and the transmission latency γ_{ij}.

$$t_{ij} = \alpha_{ij} + \beta_{ij} + \gamma_{ij} \tag{1}$$

The computation latency α_{ij}, as given by Eq. 2, is the time it takes worker w_j to run the perception aggregation task of user u_i. Note that the computation latency is affected by the η_j (i.e., the number of users currently sharing the same worker). As the number of users increases, the worker's computational capability decreases, and thus the computation latency increases.

$$\alpha_{ij} = \frac{q_i \lambda_i}{\frac{C_j}{\eta_j}} \tag{2}$$

The propagation latency β_{ij}, given by Eq. 3, is the time it takes the perception task to propagate from user u_i to worker w_j.

$$\beta_{ij} = \frac{d_{ij}}{v} \tag{3}$$

The transmission latency γ_{ij}, given by Eq. 4, is the time it takes to push the entire perception frame on the transmission link between user u_i to worker w_j. The associated delay of the result of task offloading is so small that it can be neglected [17].

$$\gamma_{ij} = \frac{\lambda_i}{R_{ij}} \tag{4}$$

PLTO minimizes the response delay t_{ij} by maximizing the latency difference g_{ij}, which acts as the utility gain from offloading the task of user u_i to worker w_b instead of the current worker w_a, as given by Eq. 5.

$$g_{ij} = t_{ia} - t_{ib} \tag{5}$$

B. Communication Loss Model

The communication loss in vehicular communications between a sender u_i and a receiver w_j can be defined as given by Eq. 6 [15], where s_{ij} is the received power at worker w_j, Pt_i is the transmission power of user u_i , and $L_{(LoS/OLoS/NLoS)_{ij}}$ is the path loss component resulting from various fading and shadowing effects due to obstacles in an urban environment.

$$s_{ij} = Pt_i - L_{(LoS/OLoS/NLoS)_{ij}} \tag{6}$$

In this work, we consider the path loss component to be categorized into three categorizes; LoS when there is a direct uninterrupted line of sight between the user and the worker, OLoS due to signals getting obstructed by vehicles, and NLoS due to scattering and reflection from buildings at intersections. If s_{ij} drops beyond a specific threshold (receiver antenna sensitivity), the packet is assumed to be dropped.

For both LoS and OLoS, we consider empirical shadow fading models built for vehicular scenarios [13]. Such models take a dual-slope form, where the slope of the path loss changes after a break-point distance d_b, which is defined by the height of both the antennas of the receiver and the transmitter [13]. The path loss for LoS and OLoS is shown in Eq. 7, where L_0 is the path loss at a reference distance d_0, both n_1

and n_2 are path loss exponents estimated by linear regression, d_b denotes the break-point distance, and X_σ is the zero mean Gaussian distributed random variable, with standard deviation σ.

$$L_{(LoS/OLoS)_{ij}} = \begin{cases} L_0 + 10 n_1 log_{10} \dfrac{d_{ij}}{d_0} + X_\sigma & \text{if } d_0 \leq d_{ij} \leq d_b \\ L_0 + 10 n_1 log_{10} \dfrac{d_b}{d_0} + 10 n_2 log_{10} \dfrac{d_{ij}}{d_b} + X_\sigma & \text{if } d_b \leq d_{ij} \end{cases} \tag{7}$$

On the other hand, NLoS path loss occurs when the line of communication between user u_i and worker w_j is blocked by a building at an intersection. The adopted path loss is shown in Eq. 8 [14], [15], where d_i is the distance between the sender u_i to the center of the intersection, d_j is the distance between the receiver w_j to the center of the intersection, w_r is the width of the street, dw_i is the distance between a user and a wall, and λ is the wavelength of DSRC signals. n_3 is path loss exponents estimated by linear regression and L_{0-NLoS} is the path loss at a reference distance.

$$L_{NLoSij} = \begin{cases} L_{0-NLoS} + 10 n_3 log_{10} \dfrac{{d_i}^{0.957}}{(dw_i \omega_r)^{0.81}} \dfrac{4\pi d_j}{\lambda} + X_\sigma & \text{if } d_j \leq d_b \\ L_{0-NLoS} + 10 n_3 log_{10} \dfrac{{d_i}^{0.957}}{(dw_i \omega_r)^{0.81}} \dfrac{4\pi {d_j}^2}{\lambda d_b} + X_\sigma & \text{if } d_b \leq d_j \end{cases} \tag{8}$$

The values of the aforementioned model parameters are adopted from the experimental studies conducted for urban/-suburban environments for LoS/OLoS [13] and the studies done for NLoS for urban intersections [14].

C. Problem Formulation

PLTO aims to jointly minimize the path loss and response delay, thus enhancing the communication efficiency of the system. It minimizes the response delay t_{ij} by maximizing the latency difference g_{ij} (given by Eq. 5). To minimize the path loss, PLTO defines a matrix, called the signal strength matrix S. Each matrix entry $\hat{s}_{ij} \in S$ represents the score of the average RSSI $\overline{s}_{ij}$, where $\overline{s}_{ij}$ is the average RSSI transmitted from user u_i to worker w_j, over a defined number of time steps τ_s. A definition of the scoring function is provided later.

The matrix S is first initialized by a default value (i.e., cold start). Then s_{ij} gets updated with the RSSI experienced by the associated worker w_j after each perception aggregation. As given by Eq. 9, s_{ij} is averaged over a defined number of time steps τ_s to better capture the traffic dynamics. A scoring function is then applied to penalize high path loss, and to reward low ones, as given by Eq. 10, where Pr_H and Pr_L denote both the highest and the lowest power that can be received, while ψ_L and ψ_H denote the values utilized by the scoring function for rewarding and penalizing the different RSSI values.

Finally, after τ_s time steps, the S matrix is reset again to its default value. If the value of $\hat{s}_{ij}$ rendered by the associated worker w_j is less than a certain threshold $\mathbb{S}$, offloading to worker w_j is considered inefficient.

$$\overline{s}_{ij} = \frac{s_{ij}}{\tau_s} \tag{9}$$

$$\hat{s}_{ij} = f_{score}(\overline{s}_{ij}) = \frac{((\overline{s}_{ij} - Pr_L) \times (\psi_H - \psi_L))}{(Pr_H - Pr_L)} + \psi_L \tag{10}$$

In addition to minimizing the communication loss and response delay, PLTO strives to keep the benefit gained by each user u_i due to perception aggregation by worker w_j above a certain satisfactory level $\mathbb{A}_i$. Note that this benefit represents the total perception area gained by user u_i due to aggregation. Towards that end, PLTO defines matrix A, in which $\overline{a}_{ij} \in A$ represents the average perception area of user u_i due to its association with worker w_j. A perception constraint denoted as $\mathbb{A}_i$ is defined by each user u_i which defines the minimum acceptable covered area received due to the aggregation of perceptions at the associated worker w_j. If the received average perception $\overline{a}_{ij}$ falls below the user's satisfactory level $\mathbb{A}_i$, offloading is considered ineffective. At the beginning, the A matrix is first initialized by a default value (i.e., cold start) to allow for the exploration phase. The value of a_{ij} of the ego-vehicle u_i then gets updated in each time step to reflect the total perception area gained due to perception aggregation by the associated worker w_j. As given by Eq. 11, such value is averaged over a defined number of time steps τ_a to better capture traffic dynamics. Finally, after τ_a time steps, the A matrix is reset again to its default value.

$$\overline{a}_{ij} = \frac{a_{ij}}{\tau_a} \tag{11}$$

We formulate the task offloading problem as a 0-1 Integer Linear program (0-1 ILP), where the binary decision variable x_{ij} is set to 1 if user u_i is assigned to worker w_j, and 0 otherwise. The problem formulation is given by Eq. 12, where the objectives given by Eq. (12a) and Eq. (12b) strive to maximize the sum of the response delay difference g_{ij} and the sum of $\hat{s}_{ij}$ (i.e., RSSI), respectively.

$$\begin{aligned}
\max_x \quad & \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} g_{ij} && && \text{(12a)} \\
\max_x \quad & \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} \hat{s}_{ij} && && \text{(12b)} \\
\text{s.t.} \quad & (x_{ij} \overline{a}_{ij}) + ((1 - x_{ij}) \times 2\mathbb{A}_i) \geq \mathbb{A}_i && \forall j \in W \;\; \forall i \in U && \text{(12c)} \\
& \sum_{i=0}^{n} x_{ij} d_{ij} \leq \xi_j && \forall j \in W && \text{(12d)} \\
& \sum_{j=0}^{m} x_{ij} \leq 1 && \forall i \in U && \text{(12e)} \\
& x_{ij} \in \{0, 1\} && \forall i \in U \;\; \forall j \in W && \text{(12f)}
\end{aligned}$$

The objectives in Eq. (12a) and Eq. (12b) are subject to the constraints (12c)-(12e). Constraint (12c) ensures that the perception area gained by each user u_i due to perception aggregation by the associated worker w_j is above the user's demanded satisfactory level $\mathbb{A}_i$. Constraint (12d) ensures that the distance between user u_i and its associated worker w_j

Algorithm 1 : PLTO-H

Input: Workers (W), Users (U), Normalized Averaged Signal Strength (S), Distance (D), Distance threshold (ξ)

Output: WorkerAssignments

1: $WorkerAssignments = \{\}$
2: $cost = \{\}\{\}$
3: $resourceConsumption = \{\}$
▷ Enforce constraints using masks
4: **for** $u \in \mathcal{U}sers$ **do**
5:     **if**
6:         user exceeded association threshold OR
7:         ineffective awarness OR
8:         ineffective rssi offloading **then**
9:         **for** $w \in \mathcal{W}orkers$ **do**
10:             **if** $(d_{uw} < \xi)$ **then**
11:                 $cost[u][w] \leftarrow (g_{uw}/\phi) * \hat{\bar{s}}_{uw} * -1$
12:             **else**
13:                 $cost[u][w] \leftarrow \infty$
14:             **end if**
15:         **end for**
16:         $resourceConsumption[u] \leftarrow 1$
17:     **end if**
18: **end for**
▷ MTHG algorithm
▷ Step1: Constructive search
19: Compute the desirability f_{uw} metric of all the assignments
20: **while** node not assigned **do**
21:     Find the worker w $(chosenWorker)$ with the maximum difference between the largest and second largest assignment desirability f_{uw} (greatest regret) that can be feasibly assigned
22:     $WorkerAssignments[u] \leftarrow chosenWorker$
23: **end while**
▷ Step2: Local search
24: **for** $w \in \mathcal{W}orkers$ **do**
25:     $C_{uw} \leftarrow min(f_{uw})$
26:     **if** ($C_{uw} <$ regret of assigning this worker) **then**
27:         $WorkerAssignments[u] \leftarrow w$
28:     **end if**
29: **end for**

does not exceed the association range ξ_j. Constraint (12e) guarantees that each user is served by at most one worker. Constraint (12f) ensures that the binary decision variable x_{ij} representing the association between each user u_i and worker w_j is either set to 1 or 0.

PLTO can be viewed as a multi-objective 0-1 generalized assignment problem, which has been shown to be NP-hard [23]. Using optimization solvers to solve this type of problems can be impractical, due to time-inefficiency. Prompted by that, we propose the PLTO-H scheme to solve the problem in a time-efficient manner using the MTHG heuristic. We present PLTO-H in the next subsection.

D. *PLTO-Heuristic (PLTO-H)*

To solve the aforementioned problem, we employ the MTHG heuristic [24] [25]. MTHG is an effective polynomial-time heuristic capable of solving the Generalized Assignment Problem (GAP) [24]. It is a combination of both constructive and local search heuristics. It is a combination of both constructive and local search heuristics, and it has been shown to render near optimal solutions [26].

As detailed in Algorithm 1, PLTO-H first considers users who have either exceeded their association distances, or have experienced ineffective offloading due to low RSSI or low perception (Lines 5-8). PLTO-H then enforces the association constraint, which stops the algorithm from choosing workers that exceed the specified association distance ξ. This is done by masking the assignment cost values that lie outside this constraint by setting their values to infinity to prevent them from being considered by the MTHG heuristic (Lines 9–14). The assignment cost is set to the negative value of the product of the delay gain and the score of the average RSSI. A scaling down factor, denoted ϕ is applied on the delay parameter in order for it to match the same scale of the score of the average RSSI (line 11). Additionally, in order to ensure that each user is allocated to only one worker, we set the resource node utilization of each user to 1 so that each device can use at most one node (Line 16).

The MTHG heuristic is applied after setting the constraints of the optimization problem. In its first phase (Lines 19–23), MTHG calculates f_{dn}, which measures the desirability of assigning the user u to a certain worker w (benefit) using the previously defined cost matrix. By iteratively considering all unassigned users, the scheme assigns each user u_i to the worker that has the maximum difference between the largest and second largest f_{dn} (regret). In the second phase (Lines 24–29), MTHG improves on the solution found in the first phase through local search, by iteratively analyzing a subset of the search space close to the found solution (within its neighborhood).

IV. Performance Evaluation

In this section, we evaluate the performance of PLTO compared to the Cooperative Perception-based Task Offloading (CPTO) scheme [19]. CPTO is a representative of task offloading decisions that focus on minimizing the perception delay without accounting for the risk of communication loss or perception redundancy and without adhering to a certain satisfactory level of improved perception awareness demanded by users. We also compare PLTO-H to PLTO to assess its performance compared to the optimal solution.

We use the following performance metrics. 1) The average response delay starting from the time a perception frame request is sent until a response is received. 2) The average perception intensity, which is calculated as the average number of collaborating vehicles associated with workers. 3) The average RSSI experienced by the associated workers, which is calculated as given by Eq. 6. 4) The packet delivery ratio, which is the ratio of data packets successfully received by workers to the total number of generated data packets. 5) The satisfaction ratio, which is defined as the ratio of users that achieve their specified satisfactory perception level $\mathbb{A}_i$ to the total number of users.

A. Simulation Setup

CPTO, PLTO, and PLTO-H are implemented using python. The IBM CPLEX optimization solver [27] is used to generate the optimal solution in PLTO and CPTO. The Simulation of Urban MObility (SUMO) traffic simulator [28] is used to generate realistic vehicular mobility traces and traffic environments. Simulations are performed over a 2 x 2 road grid topography that consists of four intersections, each of which is composed of 3 two-way lanes. Two types of vehicles are generated using the *randomTraffic.py* script provided in SUMO. The first type is the passenger vehicles, where 120 vehicles are generated with an arrival rate of 0.5 seconds. The second type are buses, where a total of 20 buses were generated with a rate of arrival of 3 seconds. The width and the length of the passenger vehicle are set to 1.8 m and 5 m respectively, while for the buses they are set to 2.5 m and 12 m respectively. The number of moving workers is set to 10 throughout the different simulations. A total of 18 RSUs (i.e., static workers) are uniformly distributed along the different lanes of the intersection.

The computation frequency of workers ranges from 5 to 7 GHz [17]. The uplink data rate of ego vehicles ranges from 20 to 21 Mbps. Note that DSRC, which supports data rates of 4.5 to 27 Mbps [22], is the communication technology adopted. In accordance with the range specified by ETSI for short and medium distances [7], the perception range is set to 20 m around each ego-vehicle. In PLTO, we set the time-steps τ_s and τ_a to 5. The perception constraint $\mathbb{A}_i$ is set to 35%. The ineffective RSSI threshold ($\mathbb{S}$) is set to -84 db. The transmitter power is set to be 20 db [13]. Pr_H is set to be of -40 db which represents the highest RSSI, while the lowest RSSI (Pr_L) is set to -100 db. These values correspond to the lowest and highest channel gains for urban environments as per the analysis done by [13]. The value of the scoring function ranges between -5 to penalize low received power denoted ψ_L and 10 to reward high power received denoted ψ_H. The association range ξ is set to 40 m and the receiver antenna sensitivity is set to -60 db. The perception task used is similar to the one adopted in [17] [19], which is the computation of image data. This computing task can serve a broad range of endeavors, such as feature detection, and perspective transformation, which are essential tasks for traffic perception. The task's computation intensity is set to 1e9 cycles/sec and the compressed perception frame size is set to 20 KB. The resolution and field-of-view are conservative estimates of a new generation of advanced driver-assistance systems (ADAS) front cameras [29]. A camera with a field-of-view of 90 degrees and a range of 20 m is assumed to be used.

The simulation period is set to 30 seconds, and the optimization problem is solved periodically every 1 second. The number of users (i.e., ego vehicles) involved in the optimization problem changes in each run of the decision making process. Only the users who have either exceeded their association range or have experienced ineffective offloading due to low RSSI or low perception are considered in the offloading decision.

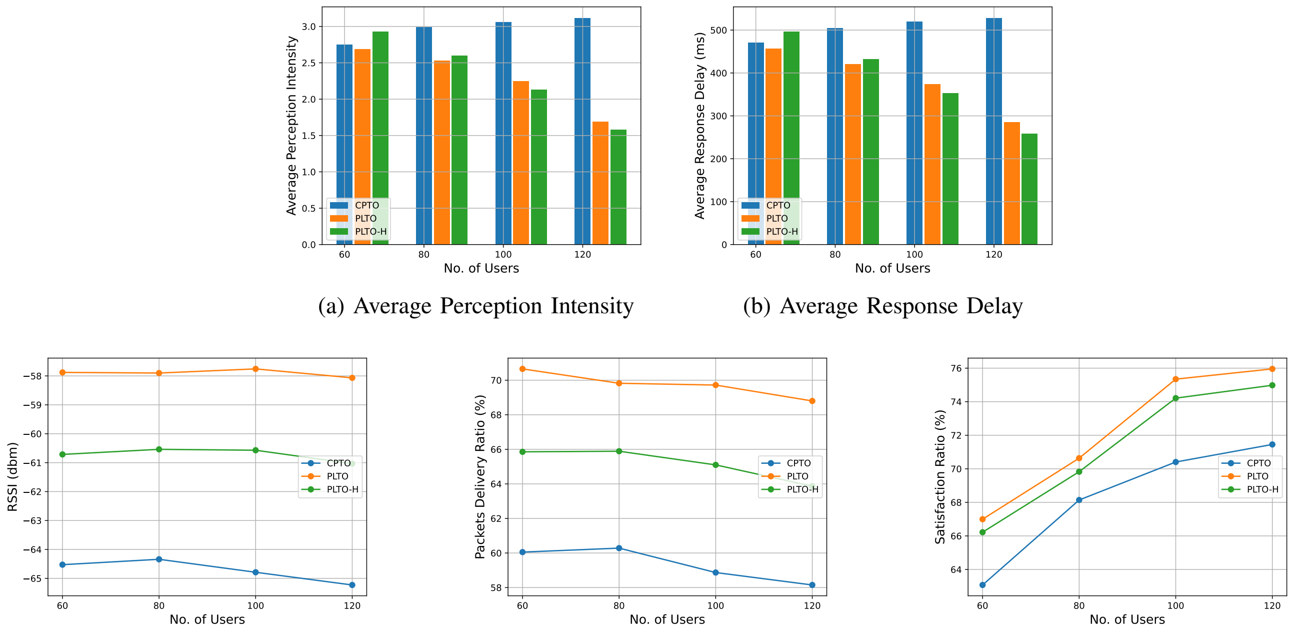
B. Simulation Results and Analysis

We evaluate the performance of CPTO, PLTO, and PLTO-H over a varying number of users, ranging from 60 to 120, to study the impact of the density of users on the task offloading procedure. Simulation results are presented at a confidence level of 95%.

We evaluate the average perception intensity of CPTO, PLTO, and PLTO-H over a varying number of users. As shown in Figure 2a, as the number of users increases in the system, the perception intensity increases in CPTO. This is since CPTO strives to maximize the number of users associated with each worker. Consequently, increasing the number of users in the system tends to increase the chance of increasing the number of collaborating users associated with each worker, thus increasing the perception intensity. In contrast, as depicted in Figure 2a, as the number of users increases, the perception intensity decreases in PLTO and PLTO-H. This is since PLTO and PLTO-H account for path loss. Thus, they tend to filter out user-worker associations that yield low received signal strength, and only select those that render high communication efficiency. This filtering capability increases as the number of users increases, since vehicles are more densely packed. In addition, PLTO and PLTO-H strive to minimize the response delay, while maintaining a certain level of perception awareness. Thus, they avoid increasing the aggregation delay by focusing on achieving a satisfactory level of non-redundant perception rather than relying on uselessly maximizing the perception intensity. Due to the aforementioned reasons, PLTO renders a lower perception intensity of up to 22.4% compared to CPTO. In addition, PLTO-H yields a performance gap of up to 6%, compared to PLTO.

Figure 2b shows the average response delay of CPTO, PLTO, and PLTO-H over a varying number of users. It can be observed that the average response delay is directly proportional to the perception intensity shown in Figure 2b. This is because as the perception intensity increases, the amount of data that needs to be aggregated at the worker increases, which increases the computation delay of the aggregation task, thus increasing the average response delay. As a result, PLTO significantly outperforms CPTO, with a reduction of up to 23.3% in delay. In addition, the yielded performance gap between PLTO-H and PLTO is 6.5%.

The communication efficiency of CPTO, PLTO, and PLTO-H is analyzed in both Figure 2c and Figure 2d. The RSSI of the workers is shown in Figure 2c, while the packet delivery ratio is shown in Figure 2d. As depicted in Figure 2c, the RSSI of CPTO, PLTO, and PLTO-H decreases as the number of users increases in the system. This is seen as a direct result of the increased traffic congestion, which in turn increases the



(a) Average Perception Intensity (b) Average Response Delay

(c) RSSI (d) Packet Delivery Ratio (e) Satisfaction Ratio

Fig. 2: Performance results of CPTO, PLTO and PLTO-H over a varying number of users.

signal obstruction, thus deteriorating the received signal. This same behavior is reflected on the packet delivery ratio depicted in Figure 2d, where the packet delivery ratio decreases with the increase in the number of users. As depicted in Figure 2c, PLTO and PLTO-H significantly outperform CPTO, with an increase of up to 10.5% and 6% in terms of RSSI, respectively. This positively impacts the packet delivery ratio shown in Figure 2d, where PLTO and PLTO-H yield an improvement of up to 17.5% and 9.8%, respectively, compared to CPTO. This is since CPTO overlooks communication efficiency in the task offloading decision and focuses on minimizing the delay and maximizing the number of collaborating users, which increases the risk of packet loss and communication failure. In contrast, PLTO and PLTO-H account for path loss and endeavor to maximize communication efficiency, thus rendering higher RSSI and packet delivery ratio. Note that the performance gap yielded between PLTO-H and PLTO is 5% in terms of RSSI and 7% in terms of packet delivery ratio, which indicates a room for improvement in the heuristic solution.

Finally, we assess the satisfaction ratio of CPTO, PLTO, and PLTO-H over a varying number of users. As depicted in Figure 2e, the three schemes exhibit an increased trend in the satisfaction ratio as more users join the system. This is since as the number of users increases, the chance of workers acquiring more perceptions to be aggregated increases, which can thus increase the level of perception rendered by users in CPTO, PLTO, and PLTO-H. Note that PLTO and PLTO-H outperform CPTO by up to 6% and 4%, respectively. This can be attributed to the fact that PLTO and PLTO-H consider the quality of perception by sustaining a certain satisfactory level of non-redundant perception demanded by users. In contrast, CPTO only focuses on increasing the number of collaborating users, which increases the risk of useless and redundant perceptions, thus reducing the satisfaction of users. It can also be seen that PLTO-H closely approaches PLTO, with a small performance gap of 1.2%, thus indicating its ability to achieve a near-optimal solution.

V. Conclusion and Future Work

This paper presents PLTO, a Path Loss-Aware Task Offloading scheme for vehicular Cooperative Perception applications. PLTO enhances the communication efficiency of cooperative perception in VEC by selectively offloading to workers that exhibit the least path loss and response delay, while adhering to a certain quality of perception required by users. PLTO formulates the task allocation problem as an Integer Linear Program (ILP) and considers path loss for LoS, OLoS, and NLoS propagation. We also propose PLTO-H, which utilizes the MTHG heuristic approach to solve the task allocation problem in a time-efficient manner. Extensive simulations show that PLTO outperforms the baseline scheme that does not consider path loss or the quality of perception, by up to 10.5%, 17.5%, 23.3%, and 6% in terms of RSSI, packet delivery ratio, perception delay, and satisfaction ratio. Performance evaluation also shows that PLTO improves perception awareness rather than perception intensity, which proves its ability to focus on the quality instead of the quantity of perception. Moreover, PLTO-H attains results close to the optimal solution in a more time-efficient manner. This work substantiates the importance of optimizing communication efficiency and considering perception quality when designing task offloading schemes for cooperative perception in VEC. In the future, we plan to apply reinforcement learning to estimate the quality of perception and quantify the level of uncertainty in such estimations.

Acknowledgment

This research is supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant number ALLRP 549919-20.

References

[1] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," *IEEE Access*, vol. 8, pp. 58 443–58 469, 2020.

[2] J. Van Brummelen, M. O'Brien, D. Gruyer, and H. Najjaran, "Autonomous vehicle perception: The technology of today and tomorrow," *Transportation Research Part C: Emerging Technologies*, vol. 89, 03 2018.

[3] S. Tang, Z. Gu, S. Fu, and Q. Yang, "Vehicular edge computing for multi-vehicle perception," in *2021 Fourth International Conference on Connected and Autonomous Driving (MetroCAD)*, 2021, pp. 9–16.

[4] Tesla, "An update on last week's accident. [online]. available:," https://www.tesla.com/blog/tragic-loss, 2018.

[5] U.S. National Transportation Safety Board, "Preliminary report, highway hwy18mh010, Nat. Transp. Saf. Board, U.S. Government, Washington, DC, USA, Tech. Rep. HWY18MH010 [online available]:," https://www.ntsb.gov/investigations/ AccidentReports/Reports/HAR1903.pdf, May 2018.

[6] Business-Wire, "Intel editorial: For self-driving cars, there's big meaning behind one big number: 4 terabytes," https://www.businesswire.com/news/home/20170414005225/en/Intel-Editorial-For-Self-Driving-Cars-There.

[7] ETSI-ITS, "Intelligent transport system (its); vehicular communications; basic set of applications; analysis of the collective - perception service (cps)," June 2019.

[8] V. Milanés, J. Alonso, L. Bouraoui, and J. Ploeg, "Cooperative maneuvering in close environments among cybercars and dual-mode cars," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 15–24, 2011.

[9] R. Yu, D. Yang, and H. Zhang, "Edge-assisted collaborative perception in autonomous driving: A reflection on communication design," in *2021 IEEE/ACM Symposium on Edge Computing (SEC)*, 2021, pp. 371–375.

[10] K. Zhang, Y. Mao, S. Leng, S. Maharjan, and Y. Zhang, "Optimal delay constrained offloading for vehicular edge computing networks," in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[11] M. G. Nilsson, C. Gustafson, T. Abbas, and F. Tufvesson, "A measurement-based multilink shadowing model for v2v network simulations of highway scenarios," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, pp. 8632–8643, 2017.

[12] J. B. P. S. M. Boban, R. Meireles and O. K. Tonguz, "Tvrtall vehicle relaying in vehicular networks," in *IEEE Trans.Mobile Comput.*, vol. 13, no. 5, 2014, p. 1118–1131.

[13] J. K. T. Abbas, Katrin Sjöberg and F. Tufvesson, "A measurement based shadow fading model for vehicle-to-vehicle network simulations," in *International Journal of Antennas and Propagation*, 2015.

[14] O. K. T. Mangel and H. Hartenstein, "5.9 ghz inter-vehicle communication at intersections: A validated non-line-of-sight path-loss and fading model," in *EURASIP J. Wireless Commun. Netw*, vol. 2011, no. 182, 2011.

[15] D. D. Yoon, B. Ayalew, and G. G. M. N. Ali, "Performance of decentralized cooperative perception in v2v connected traffic," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–14, 2021.

[16] C. Zhu, G. Pastor, Y. Xiao, Y. Li, and A. Ylae-Jaeaeski, "Fog following me: Latency and quality balanced task allocation in vehicular fog computing," in *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, 2018, pp. 1–9.

[17] E. Krijestorac, A. Memedi, T. Higuchi, S. Ucar, O. Altintas, and D. Cabric, "Hybrid vehicular and cloud distributed computing: A case for cooperative perception," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[18] S. M. A. Kazmi, S. Otoum, R. Hussain, and H. T. Mouftah, "A novel deep reinforcement learning-based approach for task-offloading in vehicular networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[19] A. M. Zaki, S. A. Elsayed, K. Elgazzar, and H. S. Hassanein, "Multi-vehicle task offloading for cooperative perception in vehicular edge computing," in *IEEE International Conference on Communications*, 2023.

[20] Y. Lin, Y. Zhang, J. Li, F. Shu, and C. Li, "Popularity-aware online task offloading for heterogeneous vehicular edge computing using contextual clustering of bandits," *IEEE Internet of Things Journal*, vol. 9, no. 7, pp. 5422–5433, 2022.

[21] L. Zeng, W. Wen, and C. Dong, "Qos-aware task offloading with noma-based resource allocation for mobile edge computing," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 1242–1247.

[22] G. G. M. N. Ali, B. Ayalew, A. Vahidi, and M. Noor-A-Rahim, "Feedbackless relaying for enhancing reliability of connected vehicles," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 4621–4634, 2020.

[23] X. Liang and M. Huang, "Modern intelligent optimization hybrid algorithm and its application." 2011.

[24] P. T. Martello, "An algorithm for the generalized assignment problem," *J.P. Brans (ed.), Operational Research'81*, pp. 589–603, 1981.

[25] "Generalized assignment solver," https://github.com/fontanf/generalized assignmentsolver, Dec. 2021.

[26] A. M. Zaki, S. A. Elsayed, K. Elgazzar, and H. S. Hassanein, "Heuristic-based proactive service migration induced by dynamic computation load in edge computing," in *IEEE Global Communications Conference*, 2022.

[27] IBM ILOG Cplex, "V12. 1: User's manual for cplex," *International Business Machines Corporation*, vol. 46, no. 53, p. 157, 2009.

[28] M. Behrisch, L. Bieker-Walz, J. Erdmann, and D. Krajzewicz, "Sumo – simulation of urban mobility: An overview," vol. 2011, 10 2011.

[29] R. Bielby, "Adas camera requirements – driving memory needs," https://www.micron.com/about/blog/2017/may/adas-camera-requirements, 2017.