IEEE Open Journal of
Vehicular Technology

Received 11 April 2022; accepted 18 May 2022. Date of publication 6 June 2022;
date of current version 14 June 2022. The review of this article was coordinated by Editor Hongsheng Lu.

Digital Object Identifier 10.1109/OJVT.2022.3177437

Safe Intersection Management With Cooperative Perception for Mixed Traffic of Human-Driven and Autonomous Vehicles

SHUNSUKE AOKI [1] (Member, IEEE), AND RAGUNATHAN RAJKUMAR[2] (Fellow, IEEE)

[1]National Institute of Informatics, Tokyo 101-8430, Japan
[2]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA

CORRESPONDING AUTHOR: SHUNSUKE AOKI (e-mail: aoki@nii.ac.jp)

This work was supported by JST, PRESTO under Grant JPMJPR2131, Japan.

ABSTRACT Autonomous driving systems are highly expected to be used on public roads to improve traffic throughput and road safety, but it will likely take a long transition period before all human-driven vehicles can be replaces with autonomous vehicles. Hence, CAVs have to safely cooperate with the surrounding human drivers, in order to gain the benefits of autonomous driving technologies during such transition periods. In this paper, we present a Distributed Synchronous Intersection Protocol (DSIP) and a Cooperative Perception-based High-Definition Map (CP-HD Map) for mixed traffic environments, in which autonomous vehicles co-exist with human-driven vehicles. First, in DSIP, each CAV utilizes dynamic decision-making mechanisms to adaptively change the vehicle behaviors based on the surrounding environments by using *vehicle states* and *vehicle mode*. In addition, in CP-HD Map, each CAV uses Vehicle-to-Vehicle (V2V) communications to share the information of detected objects to improve the road safety in the mixed traffic environments. Under these protocols, human-driven vehicles simply follow the traffic lights just like they do today. Finally, we show that DSIP and CP-HD Map increase the traffic throughput around the road intersections when we compared to existing signalized intersections and other V2V communications-based protocols.

INDEX TERMS Autonomous vehicles, connected vehicles, intersection management, intelligent transportation systems, vehicular networks.

I. INTRODUCTION

Cyber-Physical Systems (CPS) technologies [1] are widely studied and developed, and a variety of CPS applications are becoming increasingly prevalent. In particular, Connected Autonomous Vehicles (CAVs) [2] are rapidly developed and deployed. These CAVs are widely expected to improve traffic throughput, road safety, energy efficiency, and user comfort. For example, the US Government [3] estimates that more than 35,000 people die in motor vehicle-related crashes every year in the US. In particular, driving at road intersections can be safer, because more than 44% of all automotive crashes happen around intersections [4]. Since more than 90% of serious traffic accidents occur due to human error [3], autonomous driving technologies have potential to reduce that number.

However, we may need a long transition period to replace all human-driven vehicles with CAVs [5]. For example, as shown in Fig. 1, we have to consider two different types of traffic environments: **Homogeneous Traffic** and **Mixed Traffic**. Even after the automation technologies satisfy the SAE Level 5 [6], such CAVs have to cooperate with human-driven

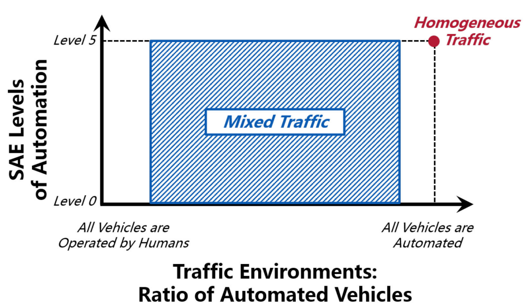


FIGURE 1. Transitions for Traffic Environments.

vehicles safely in the mixed traffic environment. Hence, to leverage the advantages of autonomous driving technologies during such transition periods, CAVs have to safely and effectively cooperate with the surrounding human-driven vehicles.

In this paper, we present a DSIP (Distributed Synchronous Intersection Protocol) and a CP-HD Map (Cooperative Perception-based High-Definition Map) that can be used in a mixed traffic environment, where human-driven and autonomous vehicles cooperate with one another to avoid vehicle collisions and possible deadlocks. First, in DSIP protocol, all CAVs enter the road intersection without stopping before and inside the intersection when all the vehicles are automated. To guarantee road safety, these CAVs use both vehicular communications and perception systems. Also, to develop safe and affordable intersection management system, we use Vehicle-to-Vehicle (V2V) communications for DSIP. Unlike the existing intersection protocols [7]–[9], DSIP is decentralized and no additional road infrastructure is required. Under the protocol, each CAV utilizes dynamic decision-making mechanisms to adaptively change the vehicle behaviors based on the surrounding environments by using *vehicle states* and *vehicle mode*. More specifically, in DSIP, each CAV keeps modifying the decision-making policy around intersections based on the presence of human-driven vehicles. When CAVs confirms that there are no human-driven vehicles at the road intersection, the CAVs negotiate with each other to determine the priority. Then, these CAVs traverse the road intersection efficiently without coming to a stop around the intersection. In addition, these CAVs follow the traffic lights like the existing human driver, when the CAVs detect human-operated vehicles, in order to cooperate and co-exist with each other. Also, under the protocol, human drivers simply follow the traffic lights just like they do today. To guarantee road safety with DSIP, each CAV has to know whether human-driven vehicles are around the intersection or not.

In addition, in order to detect human-operated vehicles in a more reliable manner, DSIP uses *Cooperative Perception* to develop the CP-HD Map, in which each CAV stores the information of detected moving/static objects to improve road safety. In cooperative perception [10]–[12], each CAV exchanges the sensor data with its neighboring vehicles, and fuses the data from its local perception system and from the neighboring vehicles. To develop the CP-HD Map, we can use Vehicle-to-Vehicle (V2V) and/or Vehicle-to-Infrastructure (V2I) communications. Since the network resources are limited, the CP-HD Map framework shares only the presence of human-driven vehicles around the intersection. Hence, each vehicle processes the sensor data locally, and shares the processed data via vehicular communications.

Here, the primary contributions of this paper are presented.

1) We present a Distributed Synchronous Intersection Protocol (DSIP) to improve traffic throughput by using V2V communications while guaranteeing road safety for mixed traffic environments.
2) We present a Cooperative Perception-based High-Definition Map (CP-HD Map) to share the static/dynamic objects by using V2V and V2I communications.
3) We introduce a dynamic decision-making mechanism to adaptively change the vehicle behaviors and policy.
4) We evaluate DSIP and CP-HD Map by simulations and demonstrate superior performances.

II. RELATED WORK

Safe cooperation for CAVs has been widely studied for intelligent intersection management, platooning, merging [13], and lane-changing maneuvers [14], [15]. In this section, we review and study the related works for such cooperation and cooperative perception. Many researchers have worked on multi-vehicle coordination for the homogeneous traffic environment where all vehicles are fully automated and connected, and few research projects have focused on the cooperation for human drivers and automated vehicles [16].

First, many works have focused on agent-based centralized control for intersection management [7], [9], [17]–[20]. Dresner and Stone [7], [9] studied a multi-agent approach for autonomous intersection management. In the autonomous intersection management, all CAVs call ahead to a reservation manager agent at the intersection to reserve conflict-free trajectories and time slots. They also studied the system to semi-autonomous vehicles [21] to enable safe intersection management. In addition, Khayatian *et al.* [18], [22] added the robustness to the system against external disturbances and model mismatches. These mechanisms heavily rely on V2I communications and on the central manager, and the infrastructure might become a single point of failure. In addition, the protocols require dedicated hardware devices for road infrastructure, which will be costly. On the other hand, our decentralized protocol does not need additional investments for road infrastructure to manage traffic at intersections.

Secondly, there are multiple studies for decentralized intersection protocols. For example, Azimi *et al.* [23] proposed several spatio-temporal intersection protocols, which use the priority-based negotiation. In these works, vehicles are assigned priorities based on their arrival times at the intersection, with vehicles reaching the intersection earlier assigned higher priorities than vehicles coming later. A vehicle with trajectory conflicts will come to a complete stop before entering the intersection, and waits until all vehicles having higher priority have crossed the conflicting areas. These priority-based intersection protocols were also extended to merge points and other driving context [24], [25]. In these works, CAVs cooperate with other CAVs and human-driven vehicles by using vehicular communications and on-board sensors while accounting for the traffic contexts around the merge points.

In addition, there are a variety of situations to lead to vehicle collisions and/or deadlocks on public roads [25]–[27]. In [25], the authors defined "dynamic intersection" and classified them to avoid vehicle deadlocks and collisions. The paper presented a cooperative dynamic intersection protocol, in which CAVs account for the dynamic situations and lane-based priority while using the vehicular communications. A longitudinal motion planning technique was also considered for lane-changing maneuvers in [27]–[29], where automated vehicles select an appropriate inter-vehicle traffic gap and time instants to perform the lane-changing maneuver. Motion-planning techniques were widely studied for cooperative driving [30]–[32]. For example, Kim *et al.* [30], [31] studied a motion-planning framework for autonomous driving and they focused on Model-Predictive Control (MPC) to understand the future behaviors for the surrounding vehicles.

Cooperative perception and/or collective perception [11], [12], [33]–[40] are highly expected to improve road safety by having connected vehicles exchange their raw or processed data with the neighboring vehicles. For example, Kim *et al.* [34] presented a cooperative perception framework for CAVs and for human-driven vehicles. They proposed an on-road sensing system to enable a see-through/lifted-seat/satellite view for the human drivers. In addition, Tsukada *et al.* [39] proposed an open-source software named AutoC2X to enable cooperative perception with CAVs. They showed the feasibility of cooperative perception by conducting field experiments with autonomous driving systems.

To enable the cooperative perception mechanisms on public roads, resource management [11], [12], [37] is essential, because the network resources are limited and network congestion may happen. In [11], Higuchi *et al.* proposed the value-anticipating networks-based approach to save the network resources. In this approach, each CAV anticipates the value of information for the potential receivers, by using the message histories and the perception systems. In addition, in [12], the deep reinforcement learning-based approach has been proposed to include a variety of factors to determine the message transmissions. By using the convolutional neural networks, the approach uses a variety of factors to determine which pieces of information are transmitted by each connected vehicle. In addition, in [37], the authors reviewed the ETSI proposals and proposed the data redundancy techniques to save the network resources. These works, however, have no specific objectives and applications. Compared to these works, our CP-HD Map are dedicated to safe cooperation and to vehicle collision avoidance at road intersections, and hence, our framework is designed as resource-efficient and resource-effective.

III. OUR ASSUMPTIONS

In this section, we present our assumptions for vehicle systems, traffic environments, and road intersections & controllers. In addition, we show existing designs and discussions for the synchronous intersection protocols [8], [41], [42] and cooperative perception [5], [12].

A. VEHICLE AUTOMATION AND TRAFFIC ENVIRONMENTS

According to the Society of Automotive Engineers International (SAE International) [6], vehicle automation are classified into 6 levels, that are ranging from fully manual to fully automated. For example, traditional vehicles driven by human operators are classified into SAE Level 0. Also, vehicles in SAE Level 5 are fully automated and might have communication devices.

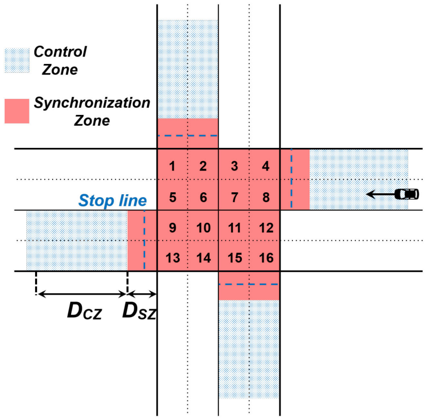


FIGURE 2. Illustration of Intersection Grid and Cells.

In this paper, we assume that each CAV includes a map database, a navigation system, a perception system, an autonomous vehicle controller, a wireless communication interface, and a localization system. In particular, the perception system collects the information of the surrounding environment by using vision cameras, radar, and/or LiDAR. Also, the wireless interface enables the system to communicate with surrounding vehicles and infrastructural sensors wirelessly. We assume that each CAV has an On-Board Unit (OBU) that supports Dedicated Short-Range Communications (DSRC), the Wireless Access in a Vehicular Environment (WAVE) protocol stack [43], [44], and/or Cellular-V2X (C-V2X).

Our protocols assume that the traffic has different automation levels. We present the focus of this paper in Fig. 1, in which the ratio of automated vehicles on public roads may be gradually changed. Hence, there might take a long transition period to replace all human-driven vehicles with CAVs [5], [12]. To guarantee road safety in the mixed traffic environment, there are 2 principles to design a vehicle cooperation protocol: (i) *CAVs cannot rely on the vehicular communications* and (ii) *CAV's behaviors in front of human drivers should be straightforward and reasonable, and never mislead the surrounding human drivers*. First, human-driven vehicles may or may not have a communication device. Therefore, the sensor-based perception systems have to guarantee road safety. Secondly, CAVs are required to obey the existing traffic rules in front of human-driven vehicles for safely cooperating with human-driven vehicles. Specifically, when there are no human-driven vehicles around the intersection, the CAVs can follow their own protocols to traverse the intersection. On the other hand, when there are human-driven vehicles around the intersection, the automated vehicles must follow the operation of the traffic light and the existing traffic rules to safely cooperate with the surrounding human-driven vehicles.

B. ROAD INTERSECTION

Road intersections are one of the main bottlenecks for road safety and throughput. In this paper, we focus on the four-way perfect-cross intersection, as shown in Fig. 2. As shown in Fig. 2, a road intersection is considered as a large grid and it has multiple small cells. When we consider the intersection shown in Fig. 2, the intersection has four lanes for each way and the area is divided into 16 (4 × 4) cells. Each cell has a unique identifier, from 1 to 16. The length of each side of a cell is equal to the width of each lane.

For our protocol, we define two dedicated zones, as shown in Fig. 2: Control Zone and Synchronization Zone.

Control Zone: Within the Control Zone, all CAVs communicate and negotiate with each other to determine the priority, and they adjust their own speeds and accelerations to enter the Synchronization Zone synchronously.

Synchronization Zone: Within the zone, all CAVs drive synchronously with a pre-determined speed. They maintain their speed from the entrance to the exit of the area.

In our vehicle systems, these physical zones are explicitly defined in the map database. To represent these two physical zones in this paper, we use two parameters: D_{CZ} and D_{SZ}. First, D_{CZ} is the length of the Control Zone. Also, D_{SZ} is the physical distance from the entrance to the exit of the Synchronization Zone. To satisfy the safety requirements for the intersection protocol, the stop line for traffic lights are within the Synchronization Zone. Further discussion to determine these two distances are shown in Section IV-E.

In addition, to enable the safe and reliable negotiation, our protocol uses the **Trajectory Cell List** (TCL) [23] that is the ordered list to include the cell numbers to be traversed at the intersection. Each CAV calculates its TCL before entering the intersection and exchanges its TCL to avoid vehicle collisions and/or deadlocks.

C. SYNCHRONOUS INTERSECTION PROTOCOLS

To improve road safety and traffic throughput at road intersections, *synchronous intersection protocols* [8], [41] were studied for the homogeneous traffic environment, where all the vehicles are fully automated and connected. Under the protocols, all CAVs enter and traverse the intersection without stopping by following a strict but well-defined spatio-temporal pattern.[1] Synchronous intersection protocols also can configure the inter-vehicle distance to improve passenger comfort and road safety [41]. Such intersection protocols are originally designed for CAVs but might be beneficial for mobile robots operated in automated warehouses.

Intersection management has quality and quantity measures, and the synchronous intersection protocols have superior performance for these two perspectives. For example, the US Federal Highway Administration (FHWA) shows different perspectives for traffic throughput, and they claim that Intersection Stops [45] is one of the most significant measurements to determine traffic throughput and efficiency. Here, Intersection Stops represents the number of stops around an intersection to traverse the area. A smaller value is better in terms of the traffic throughput and energy efficiency, because the vehicles can maintain the speed around the intersection having the small value. When we use the synchronous intersection protocols, such as Ballroom Intersection Protocol (BRIP) [8] and Configurable Synchronous Intersection Protocol (CSIP) [41], the Intersection Stops become zero and the traffic throughput becomes significantly better than that of an intersection operated by the existing traffic lights and/or controllers.

Although the protocols are promising, there are two significant drawbacks: i) possible single point of failure and ii) inability to accommodate human-operated vehicles. First, the protocols require the infrastructure and manager agent to manage traffic around the intersections. Hence, the intersection manager can become a single point of failure. Secondly, to join the protocols, each vehicle must have an appropriate wireless interface, and has to control itself very accurately. Due to its strict requirements, the existing protocols cannot accomodate human-operated vehicles.

To solve these two challenges, we propose DSIP in this paper to enable the safe cooperation of human-driven vehicles and CAVs. The protocol has the decentralized mechanisms and uses Vehicle-to-Vehicle (V2V) communications for reliable and feasible negotiations.

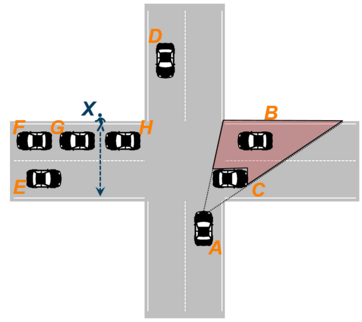


FIGURE 3. Cooperative Perception and Occlusion.

D. COOPERATIVE PERCEPTION

To improve detection accuracy and reliability of the perception systems, multiple projects are working for cooperative perception [11], [12], [33], [34], in which multiple vehicles and roadside sensors exchange the sensor data by using vehicular communications. For example, a CAV may not detect human pedestrians from its blind spot, as shown in Fig. 3. In this example, vehicle **A** cannot see vehicle **B** because of the presence of vehicle **C**. Also, vehicle **E** might not detect pedestrian **X** because the pedestrian is not within the Field-of-View (FoV) of vehicle **E**. Such occlusion might easily lead to vehicle accidents. Cooperative perception and/or collective perception are one of the most prospective solutions for

[1]The pattern is tailored to each intersection, which has various layout and geographical design.

such occlusion problems. ETSI (European Telecommunications Standards Institute) has already launched the proposals for the standardization of cooperative and/or collective perception. In Fig. 3, cooperative perception tells the presence of vehicle **B** to vehicle **A**. In addition, vehicles **G** and **H** can tell the presence of pedestrian **X** to vehicle **E** to avoid vehicle accidents. In CP-HD Map, each CAV exchanges the information of detected human-driven vehicles to improve road safety when we use DSIP on public roads.

IV. DSIP FOR MIXED TRAFFIC

We now present the DSIP protocol for mixed traffic environments to enable the cooperation of human-driven vehicles and CAVs at road intersections. The decentralized synchronous intersection protocol is designed to avoid vehicle collisions and deadlocks while keeping traffic throughput high.

One of the main features in DSIP is a *dynamic decision-making mechanism*, in which each CAV adaptively changes the decision-making policy based on the surrounding environment by using 2 elements: (i) *vehicle states* and (ii) *vehicle mode*. First, each CAV holds a vehicle state Φ_{state} that dynamically transitions, and is used to determine the vehicle decisions around the intersection while considering the surrounding situation, the presence of human-driven vehicles, and traffic congestion.

Secondly, to enable the safe cooperation with human-driven vehicles, under the protocol, CAVs modify their policies to determine the future maneuvers, which includes **Traffic-light mode** or **Synchronous mode**. Under the Traffic-light mode, automated vehicles follow the existing traffic rules and the operation of the traffic light. Under the Synchronous mode, CAVs negotiate with the surrounding vehicles and traverse the intersection synchronously. The vehicles control their speeds and arrival times within the Control Zone, and cross the road intersection efficiently without stopping before and/or inside the intersection. To guarantee road safety, all vehicles around the intersection have to be in the same mode by keeping exchanging the information.

Once the CAVs detect human-driven vehicles around the intersection, they change their mode to **Traffic-light mode** and follow the existing traffic rules for human drivers, to prevent human drivers from making mistakes and/or to avoid vehicle accidents. To ensure road safety, DSIP uses cooperative perception to detect any surrounding human-driven vehicles cooperatively, as we discuss in Section V.

To utilize the protocol in practice, DSIP uses Advanced Safety Messages for V2V communications [44], [46], which can be implemented with the second optional part of Basic Safety Messages (BSM) [44].

A. PROTOCOL POLICY

To safely manage both human-driven vehicles and CAVs, the intersection protocol should satisfy 2 requirements: i) *Human drivers do not need to change their behaviors* and ii) *CAVs must not let human drivers' make a mistake*. First, human drivers must follow the traffic rules as they do today but do not change their behaviors based on the presence of CAVs. For safety purposes, CAVs have to adapt to human behaviors. In addition, CAVs' behaviors in front of human-driven vehicles must be very straightforward and comfortable, and never mislead the surrounding human drivers.

To satisfy these 2 requirements, we use a *mode-based decision-making* mechanism in which each CAV dynamically changes the decision-making policy. We present the vehicle mode for the mode-based decision-making mechanism below.

Traffic-light mode

Under the mode, CAVs follow the existing traffic lights; they enter the intersection when the dedicated light is green and they stop before the intersection when the dedicated light is yellow or red.

Synchronous mode

Under the mode, CAVs enter and traverse the intersection in a synchronous manner. All the vehicles adjust their own speeds, accelerations, and arrival time within the Control Zone, and traverse the Synchronization Zone without coming to a stop.

When there are human-driven vehicles around the intersection, the CAVs are in the the Traffic-light mode and they follow the existing traffic rules. On the other hand, when there are no human-driven vehicles around the intersection, the CAVs are in the Synchronous mode, and the vehicles follow a V2V-based vehicle synchronization.

Here, we list up the notation to be used in this paper.

- (r, r): Cell size. (lane width r)
- (l_α, w_α): Physical size of vehicle α (length l_α and width w_α).
- v_{sync}: Vehicle's velocity for the synchronization.
- v_{limit}: Speed limit.
- ω: Safety gap.
- t_α: Original Arrival Time of vehicle α.
- $\hat{t_\alpha}$: Assigned Arrival Time of vehicle α.
- D_{SZ}: Distance for Synchronization Zone.
- D_{CZ}: Distance for Control Zone.
- D_{SL}: Distance from the intersection entrance to the Stop Line.

Here, v_{sync} represents the speed for vehicle synchronization within the Synchronization Zone. v_{limit} is the pre-determined speed limit for the area. For safety reasons, the synchronization speed v_{sync} has to be smaller value than the limit v_{limit}

In addition, the inter-vehicle gap ω is designed as a configurable value to improve road safety. Hence, the protocol accounts for the GPS accuracy level and system reliability and provide a sufficient safety gap.

As for Original Arrival Time t and Assigned Arrival Time $\hat{t}$, each vehicle first calculates its Original Arrival Time t while approaching the intersection. Each CAV uses the value t to determine the priority at the intersection. Once the negotiation is completed, each vehicle gets its Assigned Arrival Time $\hat{t}$ and it adjusts the speed and acceleration to enter the intersection at time $\hat{t}$. To guarantee vehicle safety, the assigned time $\hat{t}$ is not earlier than the original time t and each vehicle never accelerates to satisfy the assigned arrival time.

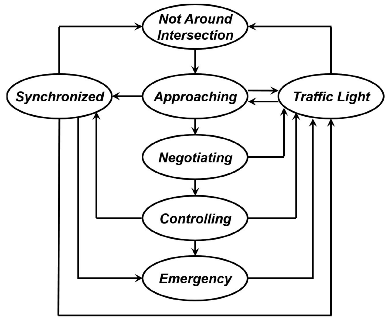


FIGURE 4. State Transition used by CAVs at Intersections.

B. MESSAGE FORMAT

In DSIP, each vehicle uses Advanced Safety Messages for negotiations. The messages are broadcast at 10 Hz, the same rate as the BSM.

The message contains 6 variables.

- TCL (Trajectory Cell List).
- Vehicle State Φ_{state}.
- Original Arrival Time t.
- Assigned Arrival Time $\hat{t}$.
- Visible HV Flag HV_{flag}.
- Timestamp of Last HV τ_{HV}

These parameters are used for the negotiation and information sharing. First, TCL, Φ_{state}, t, and $\hat{t}$, are for the negotiation to determine the priority. Vehicles hold their constant TCL and t, and also have two variables, Φ_{state} and $\hat{t}$.

For the mixed traffic environment and cooperative perception, each CAV uses HV_{flag} and τ_{HV}. The HV_{flag} is the flag to represent the presence of the surrounding human-operated vehicle(s). τ_{HV} is the timestamp that is taken when a CAV detects any surrounding human-driven vehicles. τ_{HV} will be updated whenever the neighboring CAVs detect human-operated vehicles and/or the information is shared through V2V communications. Further discussions and explanations for these values are shown in Section V.

C. VEHICLE STATE TRANSITIONS

In Fig. 4, we present 7 vehicle states and state transition diagram. Each CAV holds one of these vehicle states Φ_{state}, and it changes to either the *Traffic Light state* or the *Synchronized state* when CAV enters the Synchronization Zone to determine their modes. We show the vehicle behaviors within each state next.

1) NOT AROUND INTERSECTION STATE

When the vehicle is not near a road intersection, the vehicle is in the *Not Around Intersection state*. Also, the vehicle state Φ_{state} transitions to the *Not Around Intersection* once the CAV exits the intersection.

2) APPROACHING STATE

When the CAV approaches a road intersection, the vehicle state Φ_{state} changes to the *Approaching state* from the *Not Around Intersection state*. Under this state, the CAV calculates its TCL and Original Arrival Time t by using the map database while keep broadcasting these information to initiate the negotiation.

3) NEGOTIATING STATE

When the CAV detects the surrounding CAVs by using the perception systems and/or vehicular communications, the vehicle state Φ_{state} becomes the *Negotiating state*. In this state, each vehicle negotiates and determine the priority to use the intersection area.

4) CONTROLLING STATE

Once the CAV takes the priority and determines the Assigned Arrival Time, the vehicle state transitions to the *Controlling state*. The primary maneuver for this state is controlling their speeds, accelerations, and arrival times to enter the intersection to fit its target ones.

5) SYNCHRONIZED STATE

There are two ways to be Synchronized state. First, when the CAV in the *Approaching state* confirms that there are no vehicles around the intersection, the vehicle state Φ_{state} transitions to the *Synchronized state*. Secondly, once the vehicle in the *Controlling state* completes the required control maneuvers, the state transitions to the *Synchronized state*. In this vehicle state, the vehicle maintains the assigned speed and arrival time to traverse the intersection area synchronously. For emergency situations and for safety, the vehicles can change their speeds and/or arrival times.

6) TRAFFIC LIGHT STATE

Once the CAV detects human-operated vehicles by using their own sensors and/or CP-HD Map, the vehicle state Φ_{state} transitions to the *Traffic Light state* to follow the traffic light. In this state, the vehicles simply obey the existing traffic rules to collaborate with the surrounding human-driven vehicles.

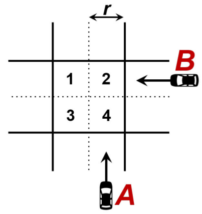
7) EMERGENCY STATE

When the CAV meets emerge situations, such as a traffic accident and/or a traffic rule violations, the vehicle state Φ_{state} becomes the *Emergency state*. Under this state, all of the CAVs stop to guarantee their own safety. The protocol never compromise vehicle safety.

Whenever the CAV detects the human-driven vehicle(s) around the intersection, the vehicle state transitions to the *Traffic Light state*. Hence, when the CAV in the *Synchronized state* detects the human-driven vehicle(s) near the intersection and when the traffic light is red, the CAV has to decelerate and stop before the intersection. Therefore, the distance D_{SZ} is designed for the required distance for the safety buffer to stop.

Algorithm 1: Priority Assignment for V2V-based Vehicle Synchronization.

```
foreach k ∈ Neighboring CAVs do
    if Φ^k_state == Controlling or Synchronized then
        if potential collision then
            if t'_{ego,β} < t̂_{k,β} − ω or
              t'_{ego,β} > t̂_{k,β} + r/v_sync + ω then
                Priority = -1;
            else
                Priority = 0;
            end
        else
            Priority=-1;
        end
    else if Φ^k_state == Approaching or Negotiating then
        if t_k > t_ego then
            if Priority ≠ 0 then
                Priority =1;
            else
                Priority = 0;
            end
        end
    end
end
```



(a) Two Vehicles Arrival.

Vehicle	Estimated Arrival Times at cell1	cell2	cell3	cell4
A		$t_A + \frac{r}{v_{sync}}$		t_A
B	$t_B + \frac{r}{v_{sync}}$	t_B		

(b) Estimated Arrival Times.

FIGURE 5. Scenario under *Synchronous mode.*

D. DECENTRALIZED VEHICLE SYNCHRONIZATION

We now present the decentralized vehicle synchronization mechanisms. In the protocol, when there are no surrounding human-driven vehicles at the intersection, all CAVs have the same constant speed and are not required to stop around and/or inside the intersection.

In this section, we present the priority assignment, deadlock-free features, and configurable safety gap.

1) PRIORITY ASSIGNMENT

In the protocol, each vehicle uses the original arrival time t, TCL, and the vehicle state Φ_{state} to determine the priority. After the priority is assigned, each vehicle gets the assigned arrival time $\hat{t}$ to safely traverse the intersection area. We show the procedures to determine the priority.

First, the CAVs in the *Controlling state* and in the *Synchronized state* have completed the priority assignment and already had their own Assigned Arrival Times. Hence, these CAVs will not change their Assigned Arrival Times and other surrounding CAVs have to select the appropriate Assigned Arrival Times for themselves. In other words, when the CAV has at least one neighboring vehicle in *Controlling state* or in *Synchronized state*, the CAV has to determine the arrival time and control its own speed to avoid vehicle collisions and/or deadlocks.

The vehicles not in *Controlling state* and in *Synchronized state* have to participate in the negotiation to determine the priority and get the assigned arrival time. In DSIP, in principle, the vehicle with the smallest original arrival time gets precedence. Any tie of the original arrival times is broken by using pre-appointed identifiers, such as a Lane ID that is explicitly defined in the map database.

After getting the priority and the assigned arrival time, the vehicle state transitions to *Controlling state* to control the vehicle speed and arrival time. Knowing the Assigned Arrival Time of the neighboring vehicles with higher priority, the CAV calculates its Assigned Arrival Time at Cell β, $\hat{t}_{ego,\beta}$, as shown in (1). Here, $\hat{t}_{neigh,\beta}$ represents the Assigned Arrival Time at Cell β of the neighboring CAV. β is the potential collision cell between the ego vehicle and the neighboring vehicle.

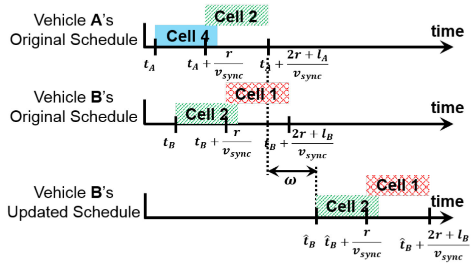
$$\hat{t}_{ego,\beta} = max\left\{t_{ego,\beta}, \quad \hat{t}_{neigh,\beta} + \frac{r}{v_{sync}} + \omega\right\} \tag{1}$$

We show the pseudo-code for such priority assignment in Algorithm 1. Here, $t'_{ego,\beta}$ is the arrival time at Cell β of the ego vehicle. Also, in Algorithm 1, when the *Priority* becomes 1, the ego vehicle takes the priority to determine the assigned arrival time. Here, for safety purposes, each CAV calculates its priority and determine the maneuvers. No CAV commands other CAVs' behaviors and future plans.

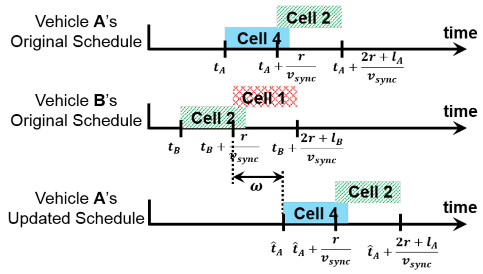
As presented in Fig. 5(a), consider the case when there are two vehicles and vehicles **A** and **B** are approaching from different directions. Here, since vehicle **A** traverses cell 4 and cell 2 at the intersection, vehicle **A** transmits its TCL $\{4, 2\}$ and the Arrival Time t_A. In addition, vehicle **B** broadcast its TCL $\{2, 1\}$ and its own arrival time t_B. The priority assignment follows (2). After the message exchanges and negotiations, they have to determine the priority and control themselves to avoid vehicle accidents.

$$t_A - t_B \begin{cases} \leq 0 \ (\mathbf{A} \text{ has priority}) \\ > 0 \ (\mathbf{B} \text{ has priority}) \end{cases} \tag{2}$$

Let us consider two cases. First, when vehicle **A** has the priority (i.e. $t_A \leq t_B$), vehicle **A** does not need to change its arrival time ($t_A = \hat{t}_A$). However, Vehicle **B** has to arrive at the intersection after $t_A + \frac{2r}{v_{sync}} + \omega$, as shown in Fig. 6(a). To satisfy this condition, vehicle **B** has to slow down within the Control Zone and delay its arrival time. Secondly, when vehicle **B** has the priority (i.e. $t_A > t_B$), it does not change the arrival time. In the protocol, vehicle **A** has to arrive at the intersection after $\hat{t}_B + \omega$, as shown in Fig. 6(b). Here, the

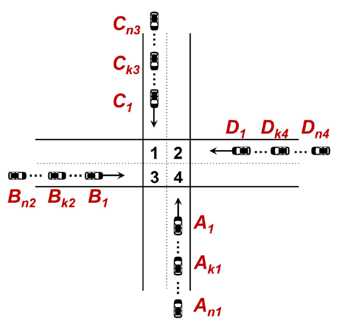


(a) Case I: Vehicle **A** has the priority.

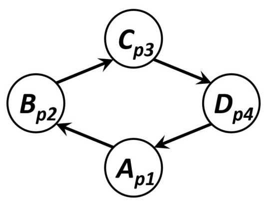


(b) Case Ⅱ: Vehicle **B** has the priority.

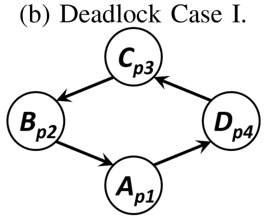
FIGURE 6. **Timelines and Schedules for Vehicles A and B.**



(a) Consecutive Vehicle Arrivals.



(b) Deadlock Case I.



(c) Deadlock Case II.

FIGURE 7. **Negotiation and Deadlocks.**

protocol provide the safety gap ω for inter-vehicle distance to be resilient against GPS inaccuracies and control system failure.

2) VALIDATION FOR DEADLOCK FREEDOM

Next, we validate the deadlock-free features of DSIP. A deadlock represents a situation in which two or more vehicles are waiting for each other to take action, and subsequently, nobody is able to progress.

We show an example scenario in Fig. 7(a), and suppose vehicles $A_1, A_2, \cdots A_{n1}$, $B_1, B_2, \cdots B_{n2}$, $C_1, C_2, \cdots C_{n3}$, and $D_1, D_2, \cdots D_{n4}$ arrive from different directions. $(n1, n2, n3, n4 \in \mathbb{N}.)$ Also, t_{A_1}, t_{B_1}, t_{C_1}, and t_{D_1} are the original arrival times for vehicles A_1, B_1, C_1, and D_1, respectively. Also, we show two possible wait-for graphs in Fig. 7(b) and -(c), in which the deadlock happens among vehicles A_{p1}, B_{p2}, C_{p3}, and D_{p4}. $(p1, p2, p3, p4 \in \mathbb{N}.)$ We use these notation to show the theorem that our negotiation-based protocol is a deadlock-free protocol.

Theorem DSIP is a deadlock-free protocol.

Proof: Under the protocol, only the vehicles in *Approaching state* and *Negotiating state* negotiate with each other to determine which CAV has the priority to get the Assigned Arrival Time. Therefore, the vehicles in *Synchronized state* and *Controlling state* do not join the priority assignment procedure and do not lead the deadlock situation. Suppose $A_1, \cdots A_{k1}$, $B_1, \cdots B_{k2}$, $C_1, \cdots C_{k3}$, and $D_1, \cdots D_{k4}$ are in either *Synchronized state* or *Controlling state*, and we only need to consider the deadlock among $A_{k1+1} \cdots A_{n1}$, $B_{k2+1} \cdots B_{n2}$, $C_{k3+1} \cdots C_{n3}$ and $D_{k4+1} \cdots D_{n4}$. $(1 \leq k1 \leq n1,\ 1 \leq k2 \leq n2,\ 1 \leq k3 \leq n3,\ 1 \leq k4 \leq n4.)$

Next, suppose the leading vehicles for each direction in the *Approaching state* or *Negotiating state*, A_{k1+1}, B_{k2+1}, C_{k3+1} and D_{k4+1}, are deadlocks like Case I, as shown in Fig. 7(b). Here, the possible conditions to lead to the deadlock are captured in (3).

$$\begin{cases} t_{A_{k1+1}} < t_{D_{k4+1}} & \text{(A has higher priority than D)} \\ t_{B_{k2+1}} < t_{A_{k1+1}} & \text{(B has higher priority than A)} \\ t_{C_{k3+1}} < t_{B_{k2+1}} & \text{(C has higher priority than B)} \\ t_{D_{k4+1}} < t_{C_{k3+1}} & \text{(D has higher priority than C)} \end{cases} \tag{3}$$

Here, the relationships are cyclic and no values of $t_{A_{k1+1}}$, $t_{B_{k2+1}}$, $t_{C_{k3+1}}$, and $t_{D_{k4+1}}$ can satisfy (3), we can conclude that there are no such deadlocks among these four vehicles like Case I in Fig. 7(b). Hence, one of them takes the priority to determine the arrival time and can transition to the *Controlling state*. Likewise, the peer-to-peer negotiation among four vehicles happens, continuously. From the processes, we can conclude that there are no deadlock situations like Case I in Fig. 7(b). Similarly, there are no possible deadlocks like Case II shown in Fig. 7(c). ■

3) CONFIGURABLE SAFETY GAP

The protocol provides a configurable inter-vehicle distance ω to satisfy safety requirements and to keep human passengers comfortable. In particular, we can change the vehicle distance based on the GPS accuracy levels and system reliability. Since the lateral error is easily calibrated by using lane markers [47], the protocol focuses on the longitudinal error.

More specifically, by increasing the inter-vehicle distance ω, we can improve vehicle safety. The longer safety gap provides the resiliency against the localization failures and control failures. On the other hand, the price to be paid is the lower traffic throughput. Overall, we have to configure the distance ω by accounting for the system reliability and occasional conditions.

We show the simulation video for the negotiation of DSIP here: https://youtu.be/W0GSDhHHXi0

E. CONFIGURATION FOR MAP DATABASE

Here, we study the design and configuration for the distances of Synchronization Zone D_{SZ} and Control Zone D_{CZ}. These values should be determined by accounting for the decelerating distances and general vehicle abilities. We show the required conditions for D_{SZ} and D_{CZ} in (4).

$$\begin{cases} D_{SZ} \geq \ D_{SL} + \theta(v_{sync}, 0) \\ D_{CZ} \geq \ \theta(v_{limit}, v_{sync}) \end{cases}, \tag{4}$$

where $\theta(v_w, v_z)$ represents the distance to decelerate from the initial velocity v_w to the target velocity v_z. For human drivers, the deceleration distance is determined by the sum of the reaction distance and the braking distance [48], [49]. For automated vehicles, since they have short reaction time, the protocol only accounts for the braking distance. Here, the deceleration distance $\theta(v_w, v_z)$ is generally determined from the kinetic energy equation as shown in (5).

$$E = \frac{1}{2}m \cdot v_w^2 = \frac{1}{2}m \cdot v_z^2 + \mu \cdot m \cdot g \cdot \theta(v_w, v_z), \tag{5}$$

in which m is the vehicle's mass and μ is the coefficient of friction between the road surface and the tires. Based on these equations, we can determine the required distances for D_{SZ} and D_{CZ} to satisfy the safety requirements.

$$\theta(v_w, v_z) = \frac{v_w^2 - v_z^2}{2\mu g} \tag{6}$$

$$\begin{cases} D_{SZ} \geq \ D_{SL} + \frac{v_{sync}^2}{2\mu g} \\ D_{CZ} \geq \ \frac{v_{limit}^2 - v_{sync}^2}{2\mu g} \end{cases} \tag{7}$$

Here, such designs should be on top of the existing road designs. For example, the distance from the intersection entrance to the stop line D_{SL} is typically pre-determined by the existing designs for human-driven vehicles. In addition, the friction μ has widely studied in the transportation engineering and has been widely considered when the road networks are designed [45], [49].

V. COOPERATIVE PERCEPTION-BASED HIGH-DEFINITION MAP

In this section, we present the Cooperative Perception-based High-Definition Map (CP-HD Map) to improve detection accuracy and coverage. To enable the safe cooperation between human-driven vehicles and CAVs, CAVs have to reliably detect human-driven vehicles. Cooperative perception is a very prospective technique to use DSIP on public roads. As shown in Fig. 8, CP-HD Map considers two cooperative ways: i) V2I communications-based data sharing and ii) V2V communications-based data sharing. First, in order to enhance the detection accuracy, CAVs might use the sensor data from the roadside infrastructure. In fact, many IoT (Internet-of-Things) devices are becoming prevalent on public roads and they are available for CAVs. Secondly, by using peer-to-peer communications, each CAV can exchange the sensor data to help each other. Hence, each CAV gets the data from its own sensors and V2X interfaces, and it processes these data in the Data Fusion Component and Data Selection Component. In the data selection component, each CAV processes the sensor data to select which data are beneficial for the potential receivers. Also, the fused data is used to determine vehicle's behaviors and navigation.

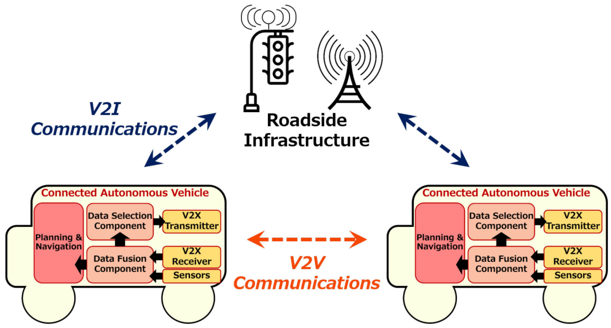


FIGURE 8. Cooperative Perception Framework with V2I Communications and V2V Communications.

Since cooperative perception is heavily relying on the vehicular communications, efficient resource management is essential to use CP-HD Map. The excessive network congestion might lead to the risk that important data packets are delayed or even lost, potentially leading to serious safety concerns. In our protocol, to save the network resources, each CAV only transmits a flag HV_{flag} and a timestamp τ_{HV}. HV_{flag} is to represent the presence of human-driven vehicle(s) around the vehicle. τ_{HV} is used to share the time when a CAV detects any surrounding human-driven vehicles. Our protocol never transmits raw sensor data, in order to keep the communication reliability high.

Our protocol has two different components: (i) Local Data Processing and (ii) Cooperative Data Sharing. Each vehicle independently senses its surrounding environment, and exchanges the information of the surrounding human-driven vehicle(s) cooperatively. By using the cooperative perception mechanism, CAVs are able to detect and track the human-driven vehicles that might be occluded and/or be in the blind spots.

A. LOCAL DATA PROCESSING

We now present the local perception and local data processing for CP-HD Map. To develop reliable CP-HD Map in real time, each CAV firstly collects the data from its Field-of-View (FoV). Here, each CAV holds two tables presented as below, in order to detect human-driven vehicles effectively.

- Ψ_{comm}: Neighborhood table built from the safety beacon (BSM) of connected vehicles. Each vehicle in this table has its location and operation (human-driven or CAV).
- $\Psi_{perception}$: Neighborhood table built from the local perception. Each vehicle in this table is detected by the table holder. Each vehicle in this table has its approximate location.

Algorithm 2: Human-Driven Vehicle Detection with *Cooperative Perception.*

foreach $i \in \Psi_{perception}$ **do**
  foreach $j \in \Psi_{comm}$ **do**
    if $\Psi_{perception}$*(i).location* $==$ Ψ_{comm}*(j).location* **then**
      Ψ_{CAV}.add($\Psi_{perception}$(i));
    end
  end
end
if *size*($\Psi_{perception}$) $\neq$ *size*(Ψ_{CAV}) **then**
  $HV_{flag} = 1$;
  τ_{HV} = Current Time;
end
if $HV_{flag} == 0$ **then**
  if $HV_{flag,neigh} == 1$ **then**
    $\tau_{HV} = \tau_{HV,neigh}$;
  end
end
if $HV_{flag} == 1$ **then**
  if τ_{HV} − *Current Time* $> \delta$ **then**
    $HV_{flag} = 0$;
  end
end

Each CAV compares the location of the vehicles in Ψ_{comm} and in $\Psi_{perception}$, and knows which vehicle is human-operated or autonomous driving. The table Ψ_{comm} is developed by using the BSMs, and hence, the CAVs hold other CAVs' locations, speeds, and operation in the table Ψ_{comm}. The surrounding and visible CAVs might be stored both in Ψ_{comm} and in $\Psi_{perception}$, but the surrounding and visible human-driven vehicles might be stored only in $\Psi_{perception}$. Hence, each CAV knows the operation of each vehicle. In addition, in practice, semi-automated vehicles, such like SAE Level 3 and 4 vehicles, have human-operated mode and autonomous driving mode, and our framework can accommodate these semi-automated vehicles. Each vehicle can transmit the operation explicitly. Also, some human-driven vehicles, such like SAE Level 0 or 1 vehicles, may have the wireless communication interface, and these connected vehicles may express their operations via vehicular communications.

To be used in DSIP, we show the pseudo-code for the decision-making algorithm in lines 1 through 11 in Algorithm 2. To store all visible CAVs, DSIP builds an additional table named Ψ_{CAV}. By using the information stored in the table Ψ_{CAV} and local perception system, each CAV ensure whether there are surrounding human-driven vehicles or not. In the CP-HD Map, each CAV updates a timestamp τ_{HV} when it detects surrounding human-driven vehicles. The timestamp τ_{HV} is used in the cooperative sharing step, in order to design resource-efficient cooperative perception.

B. COOPERATIVE DATA SHARING

Next, we present the cooperative data sharing for developing CP-HD Map. To keep the data integrity and consistency for detected human-driven vehicles, the protocol uses both the flag HV_{flag} and the timestamp τ_{HV}, which represents the last time when the CAVs detect the human-driven vehicles at the intersection. Note that CP-HD Map is designed to be used for safe cooperation of human-driven and autonomous vehicles at road intersections. Since CP-HD Map only transmits two values, the framework uses 8 Bytes for each message, as shown in Table 1. The message is used only for sharing the presence of human-driven vehicles. On the other hand, when each CAV greedily transmits the basic information for detected moving/static objects, the message size might become $d \times 52$ Bytes, where the vehicle detects d objects ($d \in N$) [11]. The basic information might include the estimated orientation, speed, and object types. Hence, when there are two vehicles around the transmitter vehicle, the message size for cooperative perception becomes 104 Bytes. Overall, CP-HD Map keeps transmitting the fixed-size messages to save the network resources.

TABLE 1. Parameters for Message Size

	Message Size	
CP-HD Map	8	(Bytes)
Greedy Transmission	$d \times 52$	(Bytes)

To calculate these two values, we use multiple values, which are acquired from the neighboring CAVs by using the vehicular communications.

- $HV_{flag,neigh}$: HV_{flag} from the neighboring CAV.
- $\tau_{HV,neigh}$: τ_{HV} from the neighboring CAV.

By using these two values, each CAV cooperatively shares the detected human-driven vehicles at the intersection. To be used for safe cooperation of human-driven vehicles and CAVs, we present the cooperative detection mechanisms in Algorithm 2. As shown in line 12 to line 16, when the neighboring CAVs detect human-driven vehicles at the intersection, the receiver vehicle updates the timestamp τ_{HV} and it goes to Traffic-light mode in DSIP.

Also, we present the logic to keep updating the detection information, in Algorithm 2 from line 17 to line 21. We use a time threshold δ, and when the CAV cannot detect any human-driven vehicles for the period δ locally, the flag HV_{flag} becomes 0 and we can consider there are no human-driven vehicles in its FoV. Note that DSIP considers both HV_{flag} and $HV_{flag,neigh}$ to determine the driving mode, Synchronous mode or Traffic-light mode. By using such timeout function, each CAV can revert from the Traffic-light mode to the Synchronous mode.

To guarantee road safety in DSIP, autonomous vehicles can be classified as human-driven vehicles, but not the other way around. Hence, when the classification for the driving operation is difficult, these vehicles have to be classified as human-driven vehicles in the CP-HD Map framework. To wrap up, the CP-HD Map framework is essential to enable DSIP on public roads.

A simulation video of the cooperation between human-driven vehicles and CAVs by using CP-HD Map can be seen at: https://youtu.be/LtAGiDEfWlE

C. COMPONENT FAILURES AND RESILIENCY

Our framework, including DSIP and CP-HD Map, guarantee road safety by using both V2X communications and the on-board sensor systems cooperatively, and does not heavily rely on either of them for system resiliency. In fact, in our framework, a CAV without regular transmission of the safety beacon (BSM) will be considered as human-operated vehicles. Hence, under the framework, CAVs can simply switch the decision-making policy to Traffic-light mode even when one of them has problems for communication interfaces and/or devices. Also, for packet delay, since our framework uses the Basic Safety Messages (BSM) and Advanced Safety Messages, it inherently accommodates the loss of transmitted packets. Such safety beacon is broadcast at 10 Hz and occasional packet losses are easily compensated by subsequent successful transmission [50]. Therefore, the framework is sufficiently resilient against communication failures and mishaps.

Also, when a vehicle has anomalies, such a breakdown at the road intersection, V2V communications contribute to preventing vehicle accidents by spreading the information to the surrounding vehicles. Since the intersection protocol is decentralized one, each vehicle can determine its maneuver. Since CAVs are safety-critical and life-critical computer systems, DSIP and CP-HD Map cannot compromise road safety and should be resilient against some degree of external disturbance.

VI. IMPLEMENTATION AND EVALUATION

We now present the implementation and evaluations for DSIP and cooperative perception framework for the mixed traffic environments in the *AutoSim*, a simulator-emulator for vehicular ad-hoc networks [2].

AutoSim [2] is a model-based simulator with 3-D graphics, and consists of various individual models, including mobility models, communication models, control models, and localization models for each vehicle. By configuring each model, we can study the positive and negative effects among multiple models. In the simulator-emulator, each simulated connected vehicle can transmit and receive a Basic Safety Message (BSM) and an Advanced Safety Message. The message formats are along with the SAE J2735 standard [50], DSRC Message Set Dictionary.

To show the feasibility of our frameworks, we first evaluate DSIP from the viewpoints of traffic throughput, by comparing against a existing signalized intersection and the previous V2V-based intersection protocols: *Traffic Light-based Protocol* and *Spatio-Temporal Intersection Protocol* (STIP) [23]. We secondly evaluate the cooperative perception technique in terms of the network load, compared to a baseline protocol named *Greedy Transmission for Cooperative Perception*. We describe these baseline protocols below.

Traffic Light-based Protocol
: Traffic signals control traffic movement at the intersection. In our baseline measurements, the time length for a green light is 15 seconds, and the time length for a yellow light is 3 seconds. No vehicle begins entering the intersection from the lanes whose light is yellow or red.

TABLE 2 Environmental Settings for the Experiments

Communication range	400 (m)
Communication frequency	10 (Hz)
Synchronization speed v_{sync}	25 (km/h)
Speed limit v_{limit}	40 (km/h)
Average vehicle length l	2.6 (m)
Average vehicle width w	1.6 (m)

Spatio-Temporal Intersection Protocol (STIP)
: Under the protocol, each CAV negotiates the priority by using vehicular communications and the vehicle having the earlier arrival time has the priority. The higher-priority vehicle traverses the intersection without stopping before and/or inside the intersection area. The protocol originally designed only for autonomous vehicles and cannot accommodate human-operated vehicles.

Greedy Transmission for Cooperative Perception
: The connected vehicles always broadcast the messages for cooperative perception at 10 Hz whenever the local sensor systems detect moving/static objects on public roads. The protocol simply keeps transmitting the data for detected objects.

We present the performance metric and experimental scenarios next.

A. PERFORMANCE METRIC

To evaluate the protocols, we use the *Trip Time*, *Trip Delay*, and *Intersection Stops* [45]. First, the trip time is the time taken by a vehicle to traverse the intersection. We calculate the trip time with each protocol and each one is compared against the trip time taken by the vehicle assuming that it stays at a constant speed and does not stop at the intersection. The time difference is considered as the *Trip Delay* due to the intersection. In addition, we calculate the number of intersection stops for each protocol. The fewer intersection stops means more fuel-efficient and more comfortable protocol for the intersection management.

In addition, we use the *Network Load* to represent the network congestion for vehicular communications. We calculate the network load from the message sizes and the numbers of transmitted messages for cooperative perception. The fewer network load means more resource-efficient and more reliable framework for cooperative perception.

B. SCENARIOS

In the experimental evaluations, we use a four-way perfect-cross intersection, because we can get the similar results even when we change the structures for the intersection. In Table 2, we show the defined variables for the simulation runs. Here, we set the communication range as 400 m and the data frequency as 10 Hz. The synchronization speed is set as 25 km/h. The speed limit for this segment is set as 40 km/h. These values are set by utilizing the V2X standards [43], [44] and general vehicle abilities.

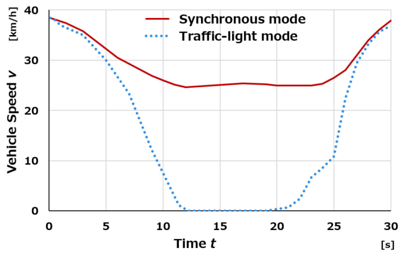
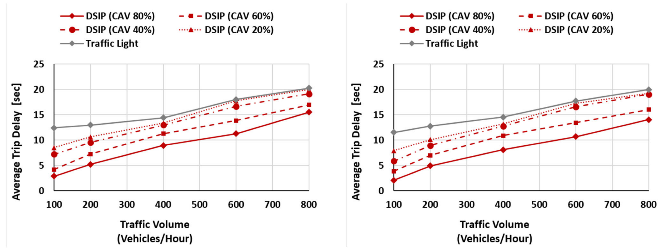


FIGURE 9. **Speed Changes for Each *Vehicle Mode*.**



(a) Arrival Pattern: Exponential Distribution.

(b) Arrival Pattern: Log-normal Distribution.

FIGURE 10. **Traffic Throughput for Mixed Traffic.**

To measure steady-state behavior, we run 30 minutes for each simulation and we evaluate the time delay and message numbers for each vehicle that are generated during the last 20 minutes.

For traffic environments, as we discuss in Section I, we focus on two traffic environments: *Homogeneous Traffic Environment* and *Mixed Traffic Environment*. In the homogeneous traffic environment, all vehicles are automated and connected. Secondly, in the mixed traffic environment, human-driven vehicles and CAVs co-exist, and they have to safely cooperate with each other to guarantee road safety. In our evaluations, we change the penetration ratio of CAVs for each simulation.

For the vehicle arrival patterns, we prepare three probability distributions: normal distribution, exponential distribution, and log-normal distribution. Since the vehicle arrival patterns and the inter-vehicle distance play a important role to analyze traffic flow, the arrival patterns have been widely studied since the 1960s [51]. One of the most prevalent distributions is the exponential distribution, and the vehicle arrival pattern for an isolated intersection follows Poisson Processes.

C. EVALUATION FOR TRAFFIC THROUGHPUT: MIXED TRAFFIC

In this section, we study the Average Trip Delay in a mixed traffic environment. We change the traffic volume, from 100 vehicles per hour to 800 vehicle per hour. Hence, we can consider both sparse rural intersections and congested urban intersections in this evaluation.

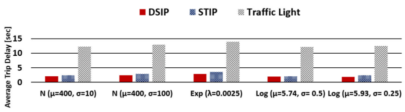
First, in Fig. 9, we show the speed changes at the intersection for DSIP vehicle modes. When the CAVs confirm that there is no human-operated vehicles, the vehicles are synchronized and keep the vehicle speed as v_{sync} around the intersection. In the synchronous mode, the CAVs never stop around the intersection. On the other hand, in the traffic-light mode, the CAVs decelerate to stop themselves to guarantee road safety, as shown in Fig. 9.

Next, we evaluate the average trip delay in mixed traffic contexts, as shown in Fig. 10(a) and (b), when we change the traffic arrival patterns. Fig. 10(a) represents the trip delay when we use the exponential distribution. Fig. 10(b) represents the trip delay when we use the log-normal distribution ($\sigma = 0.5$). For both cases, we change the penetration ratio of CAV from 20% to 80%. When the intersection traffic is sparse, DSIP demonstrates the superior performance because many CAVs do not need to stop and maintain the vehicle speed. When the traffic volume becomes high, the average trip delay is increased because more CAVs meet human-driven vehicles at the intersection. Hence, when the volume is high, more CAVs just follow the traffic light. In addition, when the ratio of the CAVs is increased, the traffic throughput becomes better performance because more CAVs can simply coordinate with CAVs and they do not need to use the traffic lights to traverse the intersection area. On the other hand, when the CAV penetration ratio is low, the average trip delay becomes closer to that in the Traffic Light-based Protocol, but it is definitely smaller. In the Traffic Light-based Protocol, since all vehicles simply follow the traffic signal, the penetration ratio of CAV does not change the throughput.

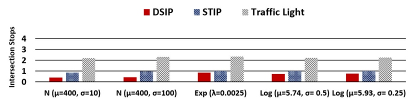
D. EVALUATION FOR TRAFFIC THROUGHPUT: HOMOGENEOUS TRAFFIC

Although DSIP is designed for the mixed traffic environment, the protocol is also useful for the fully-automated and fully-connected traffic environment. We show the performance in the homogeneous traffic environment for different arrival patterns, in Fig. 11(a), (b), (c), and (d). We study both the Average Trip Delay and Intersection Stops here. In this study, we use two baseline protocols: the Traffic Light-based Protocol and the V2V-based intersection protocol named STIP. We set as 100 vehicles/hour in Fig. 11(a) and (b), and set as 400 vehicles/hour in Fig. 11(c) and (d).

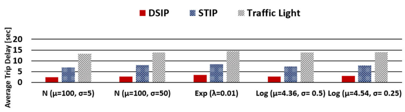
For the rural intersection with 100 vehicles/hour, we consider 5 different types of vehicle arrival patterns: i) Normal distribution ($\mu = 400$, $\sigma = 10$), ii) Normal distribution ($\mu = 400$, $\sigma = 100$), iii) Exponential distribution ($\lambda = 0.0025$) iv) Log-normal distribution ($\mu = 5.74$, $\sigma = 0.5$), and vi) Log-normal distribution ($\mu = 5.93$, $\sigma = 0.25$). From the traffic volume, the mean for the inter-vehicle distance becomes 400 m. In addition, for the urban intersection with 400 vehicles/hour, as shown in 11-(c) and -(d), we consider 5 different models: i) Normal distribution ($\mu = 100$, $\sigma = 5$), ii) Normal distribution ($\mu = 100$, $\sigma = 50$), iii) Exponential distribution ($\lambda = 0.01$) iv) Log-normal distribution ($\mu = 4.36$, $\sigma = 0.5$), and vi) Log-normal distribution ($\mu = 4.54$, $\sigma = 0.25$). The mean for the inter-vehicle distance becomes 100 m. When we see the simulation results, DSIP protocol has the shortest average trip delay and the smallest intersection stops. Overall, the decentralized protocol shows the superior performance



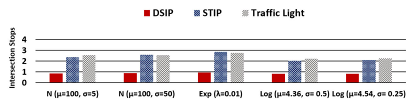
(a) Average Trip Delay (100 vehicles/hour).



(b) Intersection Stops (100 vehicles/hour).

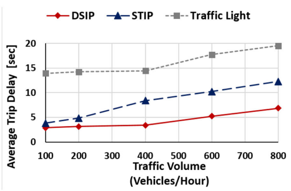


(c) Average Trip Delay (400 vehicles/hour).

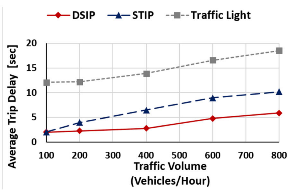


(d) Intersection Stops (400 vehicles/hour).

FIGURE 11. Traffic Throughput with Various Inter-vehicle Distance Models.



(a) Exponential Distribution.



(b) Log-normal Distribution.

FIGURE 12. Traffic Throughput for Homogeneous Traffic.

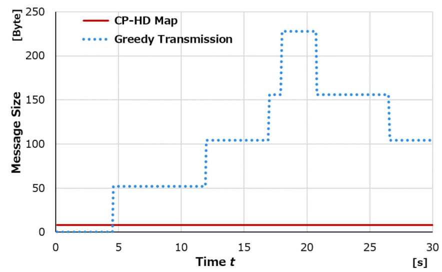


FIGURE 13. Message Size Changes of Each Transmitter.

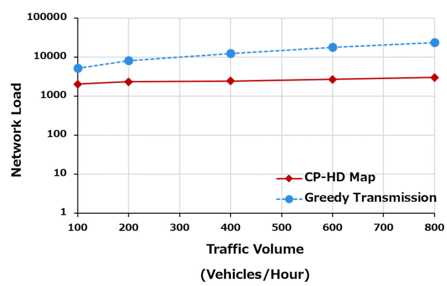


FIGURE 14. Network Load for Cooperative Perception.

against the baseline protocols. On the other hand, the Traffic Light-based Protocol always has a long average delay because approximately half of the arriving vehicles have to stop and wait for the green period to traverse the intersection. STIP protocol can have the largest value for Intersection Stops. Under the protocol, each CAV does not adjust speed and/or arrival time before reaching the entrance of the intersection. Consequently, up to 3 vehicles might need to stop before the intersection when 4 vehicles arrive at the intersection at almost the same time.

In addition, as shown in Fig. 12(a) and (b), we study the traffic throughput when we change the vehicle arrival pattern and traffic volume. We use the exponential distribution and the log-normal distribution ($\sigma = 0.5$), respectively. In all cases, DSIP protocol has the shortest average trip delay, compared to two baseline protocols. In DSIP, each CAV controls its speed and arrival time before approaching the intersection entrance. In contrast, some vehicles come to stop around the intersection in the baseline protocols. As shown in Fig. 9, the stopped vehicles delay its travel time not only for slowing down and stopping but also for accelerating later.

Overall, from multiple evaluations, we show that DSIP is efficient protocol both in the homogeneous traffic environment and in the mixed traffic environment. We also evaluate the network congestion for CP-HD Map next.

E. EVALUATION FOR NETWORK CONGESTION

We now evaluate the network loads in V2X communications for CP-HD Map. CP-HD Map is designed for a resource-efficient cooperative perception framework to support DSIP. As we present in Table 1, CP-HD Map only uses 8 Bytes for each message even when a transmitter vehicle detects d objects. On the other hand, the baseline protocol transmits all the basic information, including the orientation, speed, and object types, for each detected object. We first present the message size changes in Fig. 13, as a case study. CP-HD Map keeps transmitting the fixed-size messages, whether there are surrounding human-operated vehicles or not. On the other hand, the baseline protocol transmits the messages whose size are dynamically varied. In this case study, the transmitter vehicle detects 4 objects at the intersection, and the message size becomes over 200 Bytes.

In addition, in Fig. 14, we present the network load when we change the traffic volume (vehicles/hour). Here, the network load represents the amount of transmitted messages from each vehicle in a simulation. Since CP-HD Map framework uses the fixed-size message and fixed frequency for message transmission, the network load becomes approximately constant. On the other hand, for the baseline protocol, the network load is exponentially increased when the traffic volume is increased. When the road intersection is congested, each CAV detects more moving obstacles and increases the message sizes. From these evaluations, we show that our framework saves the network resources for cooperative perception and improves the reliability for the vehicular communications.

VII. CONCLUSION

In this paper, we presented an intersection protocol and data sharing framework named *DSIP* (Distributed Synchronous Intersection Protocol) and *CP-HD Map* (Cooperative Perception-based High-Definition Map) for the mixed traffic of human-driven vehicles and Connected and Automated Vehicles (CAVs). In the decentralized protocol, CAVs dynamically determine their decision-making policy depending on the presence of human-driven vehicles. When there are no human-driven vehicles at the intersection, the CAVs synchronously traverse the road intersection without coming to a stop before or inside the intersection. In contrast, once the human-driven vehicle appears, the surrounding CAVs follow the existing traffic rules and dedicated traffic light. In addition, to detect and track the human-driven vehicles at road intersections in a reliable manner, CP-HD Map holds the information of detected moving/static objects to improve road safety. By using cooperative perception, each CAV exchanges the sensor data with its neighboring vehicles and roadside infrastructure, and fuses the data from its local perception system and from the neighboring vehicles. The protocol and framework relies on Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications, but they are resilient against system failures.

We finally note several limitations of our work. First, we need to consider the effects and dynamics among multiple road intersections and merging points, in order to apply our framework in the real world. Secondly, connected vehicles and road infrastructure might share imprecise information with vehicular communications. In future work, we will explore the potential attacks and study more secure frameworks against false data provided by vehicular communications.

REFERENCES

[1] P. Derler, E. A. Lee, and A. S. Vincentelli, “Modeling cyber-physical systems,” *Proc. IEEE*, vol. 100, no. 1, pp. 13–28, Jan. 2012.

[2] A. Bhat, S. Aoki, and R. R. Rajkumar, “Tools and methodologies for autonomous driving systems,” *Proc. IEEE*, vol. 106, no. 9, pp. 1700–1716, Sep. 2018.

[3] NHTSA, “Automated vehicles for safety,” National Highway Traffic Safety Admin., U.S. Dept. Transp., Washington, DC, USA, 2018. [Online]. Available: https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety

[4] U.S. Department of Transportation, “Transportation statistics annual report,” U.S. Dept. of Transp., Office Sec. Transp., Bur. Transp. Stat., Washington, DC, USA, 2013. [Online]. Available: https://www.bts.gov/sites/bts.dot.gov/files/2022-01/TSAR_FULL%20BOOK-12-31-2021.pdf

[5] P. Bansal and K. M. Kockelman, “Forecasting americans’ long-term adoption of connected and autonomous vehicle technologies,” *Transp. Res. Part A: Policy Pract.*, vol. 95, pp. 49–63, 2017.

[6] SAE, “Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles,” New SAE Int. Standard J3016, SAE Int., Warrendale, PA, USA, 2021. [Online]. Available: https://www.sae.org/standards/content/j3016_202104/

[7] K. Dresner and P. Stone, “A multiagent approach to autonomous intersection management,” *J. Artif. Intell. Res.*, vol. 31, pp. 591–656, Mar. 2008.

[8] R. Azimi, G. Bhatia, R. Rajkumar, and P. Mudalige, “Ballroom intersection protocol: Synchronous autonomous driving at intersections,” in *Proc. IEEE 21st Int. Conf. Embedded Real-Time Comput. Syst. Appl.*, 2015, pp. 167–175.

[9] C.-L. Fok *et al.*, “A platform for evaluating autonomous intersection management policies,” in *Proc. IEEE/ACM 3rd Int. Conf. Cyber- Phys. Syst.*, 2012, pp. 87–96.

[10] K. Garlichs, H.-J. Günther, and L. C. Wolf, “Generation rules for the collective perception service,” in *Proc. IEEE Veh. Netw. Conf.*, 2019, pp. 1–8.

[11] T. Higuchi, M. Giordani, A. Zanella, M. Zorzi, and O. Altintas, “Value-anticipating V2V communications for cooperative perception,” in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2019, pp. 1947–1952.

[12] S. Aoki, T. Higuchi, and O. Altintas, “Cooperative perception with deep reinforcement learning for connected vehicles,” in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2020, pp. 328–334.

[13] S.-C. Lin *et al.*, “A dynamic programming approach to optimal lane merging of connected and autonomous vehicles,” in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2020, pp. 349–356.

[14] S. Y. Gelbal, S. Zhu, G. A. Anantharaman, B. A. Guvenc, and L. Guvenc, “Cooperative collision avoidance in a connected vehicle environment,” SAE Int., Warrendale, PA, USA, SAE Tech. Paper 2019-01-0488, 2019.

[15] G. Jornod, R. Alieiev, A. Kwoczek, and T. Kürner, “Environment-aware communications for cooperative collision avoidance applications,” in *Proc. IEEE 19th Int. Symp. “A World Wireless, Mobile, Multimedia Netw.” (WoWMoM)*, 2018, pp. 588–599.

[16] S. Aoki, C.-W. Lin, and R. Rajkumar, “Human-robot cooperation for autonomous vehicles and human drivers: Challenges and solutions,” *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 35–41, Aug. 2021.

[17] E. Andert, M. Khayatian, and A. Shrivastava, “Crossroads: Time-sensitive autonomous intersection management technique,” in *Proc. 54th Annu. Des. Automat. Conf.*, 2017, Art. No. 50.

[18] M. Khayatian, M. Mehrabian, and A. Shrivastava, “RIM: Robust intersection management for connected autonomous vehicles,” in *Proc. IEEE Real-Time Syst. Symp.*, 2018, pp. 35–44.

[19] B. Chalaki and A. A. Malikopoulos, “An optimal coordination framework for connected and automated vehicles in two interconnected intersections,” in *Proc. IEEE Conf. Control Technol. Appl.*, 2019, pp. 888–893.

[20] B. Zheng, C.-W. Lin, S. Shiraishi, and Q. Zhu, “Design and analysis of delay-tolerant intelligent intersection management,” *ACM Trans. Cyber- Phys. Syst.*, vol. 4, no. 1, pp. 1–27, 2019.

[21] G. Sharon and P. Stone, “A protocol for mixed autonomous and human-operated vehicles at intersections,” in *Proc. Int. Conf. Auton. Agents Multiagent Syst.*, 2017, pp. 151–167.

[22] M. Khayatian, Y. Lou, M. Mehrabian, and A. Shirvastava, “Crossroads a time-aware approach for intersection management of connected autonomous vehicles,” *ACM Trans. Cyber- Phys. Syst.*, vol. 4, no. 2, pp. 1–28, 2019.

[23] R. Azimi, G. Bhatia, R. R. Rajkumar, and P. Mudalige, “Stip: Spatio-temporal intersection protocols for autonomous vehicles,” in *Proc. ACM/IEEE Int. Conf. Cyber-Phys. Syst.*, 2014, pp. 1–12.

[24] S. Aoki and R. R. Rajkumar, “A merging protocol for self-driving vehicles,” in *Proc. ACM/IEEE 8th Int. Conf. Cyber-Phys. Syst.*, 2017, pp. 219–228.

[25] S. Aoki and R. R. Rajkumar, “Dynamic intersections and self-driving vehicles,” in *Proc. IEEE/ACM 9th Int. Conf. Cyber-Phys. Syst.*, 2018, pp. 219–228.

[26] C. Liu, C.-W. Lin, S. Shiraishi, and M. Tomizuka, “Distributed conflict resolution for connected autonomous vehicles,” *IEEE Trans. Intell. Veh.*, vol. 3, no. 1, pp. 18–29, Mar. 2018.

[27] J. Nilsson, M. Brännström, E. Coelingh, and J. Fredriksson, “Lane change maneuvers for automated vehicles,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 5, pp. 1087–1096, May 2017.

[28] T. Li *et al.*, "A cooperative lane change model for connected and automated vehicles," *IEEE Access*, vol. 8, pp. 54940–54951, 2020.

[29] Y. Chen, C. Hu, and J. Wang, "Human-centered trajectory tracking control for autonomous vehicles with driver cut-in behavior prediction," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8461–8471, Sep. 2019.

[30] K.-D. Kim and P. R. Kumar, "An MPC-based approach to provable system-wide safety and liveness of autonomous ground traffic," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3341–3356, Dec. 2014.

[31] X. Liu, K. Ma, and P. Kumar, "Towards provably safe mixed transportation systems with human-driven and automated vehicles," in *Proc. IEEE 54th Annu. Conf. Decis. Control*, 2015, pp. 4688–4694.

[32] P. Hang, C. Lv, C. Huang, J. Cai, Z. Hu, and Y. Xing, "An integrated framework of decision making and motion planning for autonomous vehicles considering social behaviors," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14458–14469, Dec. 2020.

[33] H.-J. Gunther, O. Trauer, and L. Wolf, "The potential of collective perception in vehicular ad-hoc networks," in *Proc. 14th Int. Conf. ITS Telecommun.*, 2015, pp. 1–5.

[34] S.-W. Kim, W. Liu, M. H. Ang, E. Frazzoli, and D. Rus, "The impact of cooperative perception on decision making and planning of autonomous vehicles," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 3, pp. 39–50, Mar. 2015.

[35] Q. Chen, S. Tang, Q. Yang, and S. Fu, "Cooper: Cooperative perception for connected autonomous vehicles based on 3D point clouds," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst.*, 2019, pp. 514–524.

[36] A. Miller, K. Rim, P. Chopra, P. Kelkar, and M. Likhachev, "Cooperative perception and localization for cooperative driving," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 1256–1262.

[37] G. Thandavarayan, M. Sepulcre, and J. Gozalvez, "Redundancy mitigation in cooperative perception for connected and automated vehicles," in *Proc. IEEE 91st Veh. Technol. Conf.*, 2020, pp. 1–5.

[38] S. Tang, B. H. Chen, J. Hochstetler, J. Hirsch, and S. Fu, "Cooperative mixed reality leveraging edge computing and communication," in *Proc. IEEE/ACM Symp. Edge Comput.*, 2020, pp. 425–429.

[39] M. Tsukada, T. Oi, A. Ito, M. Hirata, and H. Esaki, "AutoC2X: Open-source software to realize V2X cooperative perception among autonomous vehicles," in *Proc. IEEE 92nd Veh. Technol. Conf.*, 2020, pp. 1–6.

[40] C. Allig and G. Wanielik, "Alignment of perception information for cooperative perception," in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2019, pp. 1849–1854.

[41] S. Aoki and R. R. Rajkumar, "A configurable synchronous intersection protocol for self-driving vehicles," in *Proc. IEEE 23rd Int. Conf. Embedded Real-Time Comput. Syst. Appl.*, 2017, pp. 1–11.

[42] S. Aoki and R. R. Rajkumar, "V2V-based synchronous intersection protocols for mixed traffic of human-driven and self-driving vehicles," in *Proc. IEEE 25th Int. Conf. Embedded Real-Time Comput. Syst. Appl.*, 2019, pp. 1–11.

[43] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 GHz dedicated short range communication (DSRC) frequency band," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 8, pp. 1501–1516, Oct. 2007.

[44] J. B. Kenney, "Dedicated short-range communications (DSRC) standards in the United States," *Proc. IEEE*, vol. 99, no. 7, pp. 1162–1182, Jul. 2011.

[45] FHA, "Detection technology for IVHS," Fed. Highway Admin., U.S. Dept. Transp., Washington, DC, USA, 1996. [Online]. Available: https://www.fhwa.dot.gov/publications/research/operations/ivhs/chapter1.cfm

[46] L.-W. Chen, P. Sharma, and Y.-C. Tseng, "Dynamic traffic control with fairness and throughput optimization using vehicular communications," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 504–512, Sep. 2013.

[47] Y.-W. Seo and R. Rajkumar, "Tracking and estimation of ego-vehicle's state for lateral localization," in *Proc. IEEE 17th Int. Conf. Intell. Transp. Syst.*, 2014, pp. 1251–1257.

[48] R. Akçelik and M. Besley, "Acceleration and deceleration models," in *Proc. 23rd Conf. Australian Inst. Transport Res.*, 2001, vol. 10, Paper 12.

[49] A. K. Maurya and P. S. Bokare, "Study of deceleration behaviour of different vehicle types," *Int. J. Traffic Transport Eng.*, vol. 2, no. 3, pp. 253–270, 2012.

[50] SAE, "Dedicated short range communications (DSRC) message set dictionary," SAE Int. Standard J2735, SAE, Warrendale, PA, USA. [Online]. Available: https://www.sae.org/standards/content/j2735_201603/

[51] G. Yan and S. Olariu, "A probabilistic analysis of link duration in vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1227–1236, Dec. 2011.

SHUNSUKE AOKI (Member, IEEE) received the B.Eng. degree from Waseda University, Tokyo, Japan, the M.S. degree from The University of Tokyo, Tokyo, Japan, and the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA. After working as a Researcher with Carnegie Mellon University, and with Nagoya University, Nagoya, Japan, he is currently an Assistant Professor with the National Institute of Informatics, Tokyo, Japan. His research interests include cyber-physical systems, vehicular communications, and multirobot coordination.

RAGUNATHAN RAJKUMAR (Fellow, IEEE) received the B.E. (Hons.) degree from the University of Madras, Chennai, India, and the M.S. and Ph.D. degrees from Carnegie Mellon University, Pittsburgh, PA, USA. He is currently the George Westinghouse Professor of electrical and computer engineering with Carnegie Mellon University. His research interests include the all aspects of cyber-physical systems, with a particular emphasis on connected and automated vehicles.