Computational Intelligence Inspired Data Delivery for Vehicle-to-roadside Communications

Celimuge Wu, Member, IEEE, Tsutomu Yoshinaga, Member, IEEE, Yusheng Ji, Member, IEEE, and Yan Zhang, Senior Member, IEEE

Abstract—We propose a vehicle-to-roadside communication protocol based on distributed clustering where a coalitional game approach is used to stimulate the vehicles to join a cluster, and a fuzzy logic algorithm is employed to generate stable clusters by considering multiple metrics of vehicle velocity, moving pattern, and signal qualities between vehicles. A reinforcement learning algorithm with game theory-based reward allocation is employed to guide each vehicle to select the route that can maximize the whole network performance. The protocol is integrated with a multi-hop data delivery virtualization scheme that works on the top of the transport layer and provides high performance for multi-hop end-to-end data transmissions. We conduct realistic computer simulations to show the performance advantage of the protocol over other approaches.

Index Terms—VANETs, vehicle-to-roadside communications, routing protocol, game theory, fuzzy logic, reinforcement learning, multi-hop data delivery.

I. INTRODUCTION

With the increasing demand of collision avoidance systems, intelligent transport systems and autonomous driving systems, vehicular ad hoc networks (VANETs) are expected to attract more interests in terms of research and development. The emergence of vehicular Internet-of-Things (IoT) applications, such as vehicular data collection, police car camera video uploading, traffic information monitoring, driving habit monitoring and parking space management, motivates us to use vehicle-to-roadside (V2R) communications to deliver data from the data sources to the cloud. In this paper, we discuss a multi-hop data transmission protocol for V2R communications with specific emphasis on the routing (network) layer and upper layers.

There are two main challenges for route selection in multi-hop V2R transmission scenarios. First, the routing protocol should consider the vehicle mobility, route length, and inter-vehicle wireless link quality for the route selection, as the performance of a communication route is affected by these multiple factors. Second, since vehicles could be deployed in a high-density manner for some hours or some road segments, the number of concurrent sender nodes is expected to be large. IEEE 802.11p, the standard for wireless access in vehicular environments, has a performance degradation problem when the number of sender nodes increases due to the MAC layer contention scheme based on the exponential backoff [1], [2]. Many protocols have been proposed to solve the first problem [3]–[12] using different techniques such as predictive routing protocol [6], quality-of-service (QoS)-aware routing [7], opportunistic routing [8], street-centric approach [9], RSU controlled approach [10], intersection-centric approach [11] or geographical routing [12]. However, the second problem is still under-explored.

In recent years, some clustering approaches have been employed to mitigate the hidden node problem and improve the wireless resource utilization efficiency [13]–[19]. The design of the clustering algorithm for dynamic VANETs is particularly difficult, and most existing studies ignore some important factors specifically, vehicle mobility, vehicle distribution and channel condition which could have significant effects on the performance of V2R communications. Another issue of the existing clustering approaches is the cluster maintenance overhead which could degrade the network performance. In this paper, we use a totally distributed clustering approach that does not require explicit cluster join/leave messages. In the clustering algorithm, the abovementioned factors are jointly considered using a fuzzy logic approach. We employ a coalitional game theory-based approach to stimulate the vehicles to join clusters to reduce the number of sender nodes in the network and thus improve the channel contention efficiency at the MAC layer. Game theory has been widely applied in ad hoc network protocols especially for solving selfish behaviors in packet forwarding [20]–[26]. In contrast to these works, we employ game theory as a part of a routing metric that could directly affect the route selection of the routing protocol.

Definition of routing metrics for a multi-hop network is a difficult problem especially in a highly dynamic vehicular environment. To design an intelligent protocol in a complex and uncertain environment, we evaluate each one-hop link status by employing a fuzzy logic approach, and use a multi-agent Q-learning algorithm (a model-free reinforcement learning approach) [27] with game theory-inspired reward calculation to conduct the evaluation of multi-hop performance. The fuzzy logic can manage easy and efficient evaluation of link status by considering multiple metrics where conflicts between different metrics exist.

TCP (transmission control protocol) [28] is a widely used
transport layer protocol that provides reliable end-to-end data delivery. In a multi-hop transmission in VANETs, the end-to-end packet loss probability increases linearly with the packet hopping, which results in a low congestion window size at the TCP sender (low throughput), and a large end-to-end delay. In order to solve this problem, we propose a multi-hop data delivery virtualization approach that uses hop-by-hop acknowledgement instead of end-to-end feedback. The main contributions of this paper are as follows.

- We propose a coalitional game theory-based clustering and packet forwarding algorithm to reduce the number of sender nodes and improve the MAC layer contention efficiency for V2R communications. We jointly consider vehicle mobility, channel condition, and vehicle distribution using a fuzzy logic-based approach for the cluster head selection algorithm. A coalition game is used to stimulate vehicles to join the clusters of vehicles.
- Based on the coalitional game theory-based clustering, we employ a reinforcement learning approach to evaluate the performance of multi-hop routes. Game theory is fully integrated into the reinforcement learning as a key factor for determining the reward of the learning.
- We also integrate a multi-hop data delivery virtualization approach with the reinforcement learning-based route selection to solve the performance degradation problem of multi-hop TCP transmissions.

This paper is an extension of our previous conference paper [29]. This paper improves [29] by using a multi-hop data virtualization approach and presents new simulation results in a street scenario to show the performance advantage of the proposed approach over existing studies in a more comprehensive manner. The remainder of the paper is organized as follows: Section II gives a brief survey of related work. We give a detailed description of the proposed protocol in section III. Simulation results are presented in section IV. Finally, we draw our conclusions and future work in section V.

II. RELATED WORK

A. VANET routing protocols

Many studies have discussed the routing problem of vehicular networks [3]–[5]. Yao al. [6] proposed a predictive routing protocol that estimates the future locations of vehicles based on past traces and a hidden Markov model. Eiza et al. [7] proposed a QoS-aware routing protocol for VANETs. An ant colony optimization (ACO) technique was used to choose communication routes under multiple QoS constraints. However, the problem of selecting optimization factors associated with QoS constraints was not discussed in [7], as it focuses on the security aspect of the routing. Different from [7], this paper uses a fuzzy logic approach to jointly consider multiple metrics to achieve a practical solution for route evaluation under multiple constraints. Zhang et al. [8] proposed an opportunistic routing protocol that aims to minimize the expected transmission costs of multi-hop paths (ETCoP) and achieve high forwarding reliability by considering the link correlation. However, the drawback of the opportunistic routing is the difficulty in selecting a high data rate (modulation and coding scheme) for transmissions as the intended receiver (the next forwarder) is unknown in advance.

The concept of microtopology (MT) was proposed in [9]. The authors introduced a street-centric routing protocol that consists of two main functions, namely, inter-MT routing, and intra-MT routing. Although they pointed out the importance of considering the mobility of vehicles and the effect of channel contention, the relay node selection in intra-MT communications was not adequately discussed. Liu et al. [10] proposed an RSU-controlled data dissemination protocol over a hybrid vehicle-to-infrastructure and vehicle-to-vehicle communication environment. Due to its centralized approach where all transmission opportunities are assigned by the RSU, [10] is likely to incur a high overhead in dense vehicular environments, which could degrade the performance of the communication. In [11], Darwish et al. discussed the routing decision at the intersection, and proposed an approach that considers the structure of roads, the position of neighboring nodes, the link quality between neighbors, and the mobility for the next hop selection. Zhu et al. [12] discussed the dramatic degradation problem of wireless transmission range in multilevel scenarios, and proposed a greedy geographic routing protocol that calculates the connectivity probability at intersections for determining the forwarding direction. The most important concern about greedy opportunity routing protocols is the problem in finding a forwarder that could provide a high data rate as geographic routing protocols intend to choose the farthest node as the forwarder, which is typically inefficient in terms of the modulation and coding.

B. Clustering or backbone-based approaches

A recent survey on the clustering techniques can be found in [13]. Wang et al. [14] conducted a performance analysis of clustering in VANETs, and showed the effect of the cluster size, vehicle speed, and traffic demand on the system throughput. Abuashour and Kadoch [15] proposed a cluster-based routing protocol where the cluster areas are static, and the cluster heads are dynamically selected. Since the region of each cluster is predetermined and fixed, the change in the cluster heads is frequent when the vehicle mobility is high, which incurs a high overhead. Hafeez et al. [16] proposed a fuzzy logic-based cluster head selection algorithm that considers the relative velocity of vehicles and inter-vehicle distance. Once the cluster head is elected, [16] tries to avoid a new election process in order to reduce the cluster formation overhead. However, since the inter-vehicle link quality is not considered, [16] could result in inefficient clustering.

Togou et al. [17] proposed a CDS (connected dominating set) based backbone creation approach which selects backbone vehicles by considering vehicle speed and spatial distribution to ensure stability as well as low dissemination delay. Since backbones are generated one by one (early generated backbones specify the neighbor backbone nodes), the backbone formation algorithm of CDS-SVB is not totally distributed. As a result, the change in backbone vehicles occurs frequently in a highly dynamic vehicular network. In [18], the stable CDS-based routing protocol (SCRP) was proposed based on CDS-SVB. SCRP connects the backbone vehicles at intersections.
with bridge vehicles that maintain the whole network information and calculate the delay for transmitting data packets over road segments. Different weights are assigned to road segments by taking into account the link life time, delay, and hop count. SCRP requires the global network topology for the backbone creation, which is difficult to achieve in frequently changing networks. Moreover, the corresponding communication overhead was not adequately discussed. MoZo, a clustering approach based on the similarity of vehicle movement, was proposed in [19]. MoZo constructs multiple moving zones by grouping the vehicles that have similar movement patterns. Since explicit join request and response messages are required to maintain the moving zones, the zone maintenance overhead to manage up-to-date information of the zone members at each zone captain side could be large in a highly dense or mobile environment. Other well-known clustering approaches are discussed in [13]. The common problem of the existing studies is the cluster maintenance overhead (cluster join/leave messages).

C. Game theory based protocols

Game theory has been attracting great interest for modeling cooperative behaviors in wireless networks. An earlier survey of game theory approaches for cooperative communications was provided in [20]. Recently, Shivshankar and Jamalipour [21] proposed a public goods game (PGG)-based group interaction model for packet routing in VANETs and investigated the effects of networking properties on the evolution of cooperation. One study [22] modeled the cooperative service-based message sharing issue as a coalition game formation problem among nodes. Similar to most existing game theory-based approaches, [22] discussed the incentive of the packet relay problem, in which coalitions are formed among the nodes that belong to the same type in terms of service messages. Chen et al. [23] studied how to stimulate message forwarding according to the selected route. Different from existing game theoretical approaches that discuss how to let all nodes follow the routing decision made by the routing protocol, this paper studies how to use game theory to select the best route from multiple candidates. Kumar et al. [24] analyzed the behaviors of intermediate nodes using Bayesian coalition games (BCG) and learning automata. They showed that the probability of achieving Nash equilibrium points depends on the knowledge about the strategy space of the opponents. A coalition graph game-based peer-to-peer approach for content distribution in cognitive radio VANETs was proposed by Wang et al. [25]. Game theory was used to coordinate the vehicle-to-vehicle links and V2R links by allowing each vehicle to decide whether to connect to another vehicle. A beacon rate adaptation algorithm based on a non-cooperative game was proposed in [26]. Most existing game theory-based approaches do not adequately address the characteristics of vehicular networks, specifically lossy channels, varying density, and mobility. While these protocols mainly investigated the decision making problem of whether a vehicle should forward a packet or not, this paper employs a game theoretical approach to evaluate route candidates (game theory is used as a part of the routing metric).

III. PROPOSED PROTOCOL

A. Assumption

Each node knows its position information and road map information. The location information and velocity information of vehicles are exchanged by hello messages with an interval of 1 second, and each vehicle thus knows the neighbors in its vicinity. This interval can be tuned according to the system requirements. However, a 1-second hello interval is sufficient in most cases because the possible moving distance of vehicles in this period is much smaller than the transmission range. As we mainly discuss a connected network, the routing solutions for delay tolerant networks [30] are beyond the scope of this paper.

B. Problem definition and system overview

Due to the random backoff procedure, the contention efficiency of the IEEE 802.11p MAC layer drastically decreases with the number of concurrent sender nodes. Therefore, it is particularly important to reduce the number of sender nodes. This characteristic of wireless resource contention should be clearly addressed by the routing protocols. For the transport layer, TCP has the problem of performance degradation in a lossy multi-hop environment. The topology and environment dynamics in VANETs make these problems more challenging.

Here, we first propose a clustering approach where different sender nodes use the same forwarder node (cluster head node) to transmit the data (see Fig. 1). The clusters are formed by using hello messages that are periodically generated, and only exchanged among one-hop neighbors. A fuzzy logic approach is used to generate clusters in a distributed way. The cluster head selection algorithm considers the relative vehicle mobility and movement patterns, and therefore ensures that the generated cluster topology is stable even when the vehicle mobility is high. Second, we propose a coalitional game approach to stimulate the cooperation of sender nodes to use the cluster head nodes for the packet forwarding. The RSU distributes payoffs to the vehicles, and cluster head nodes forward the payoffs to each cluster member. Payoff allocation is integrated as an important part of a reinforcement learning algorithm used to evaluate a route in terms of the multihop performance. As a result, each vehicle is enforced to use the cluster heads for multi-hop data transmissions, which results in efficient wireless resource utilization. Third, we use a multi-hop data delivery virtualization scheme to solve the performance drop problem of TCP in a multi-hop lossy environment.

C. Dynamic clustering and cluster head selection

We use an approach where cluster heads are selected using a distributed algorithm. The cluster join/leave procedure is conducted with zero overhead, as we do not use any cluster join/leave messages for the maintenance of cluster member information. After cluster heads are determined, each cluster head announces the number of cluster members using the hello messages. We judge if a vehicle is suitable for working as a
cluster head by using a fuzzy logic-based approach. In the evaluation, we consider three different factors: 1) the moving speed of vehicles, 2) the density of vehicles that are moving toward the same direction as the current vehicle, and 3) the average channel condition between the current vehicle and its neighbors. The first two factors are used to ensure that the generated cluster heads are stable. The third factor is to give a higher priority to the vehicles that could provide better wireless links to cluster members (for example, such as the buses or trucks with higher antennae that can provide longer line-of-sight distances). Since the evaluation involves multiple factors, it is difficult to use a simple mathematical approach to conduct a fair calculation. Therefore, we use a fuzzy logic-based approach in the evaluation to jointly consider these three factors.

The cluster heads are selected based on the information shared with hello messages. Each node attaches the information about its velocity and the number of neighbor vehicles driving toward the same direction to the hello messages. Upon reception of a hello message, each node calculates a competency value (fitness value for being a cluster head) for itself and each one-hop neighbor. The node that has the largest competency value in its vicinity declares itself as a cluster head using hello messages. As shown in Fig. 2, by employing cluster head-based forwarding, multiple source nodes (S1, S2, S3 and S4) select the same nodes as the forwarder nodes, which results in a more efficient MAC layer contention.

**D. Fuzzy logic-based competency calculation**

The competency value calculation consists of three steps. First, the velocity factor, leadership factor, and signal quality factor are calculated for each one-hop neighbor within the range of $\frac{1}{2}R$. Next, we use predefined membership functions to convert the factors to fuzzy values, and use predefined fuzzy rules to calculate the final fuzzy value for each neighbor. Finally, the fuzzy value is converted to a numerical value (competency value) based on the fuzzy output membership function [31].

1) **First step – definition of three factors:**

- **Mobility Factor (MF):** After reception of a hello message from node $m$, node $s$ calculates $MF(s, m)$ as

$$MF(s, m) = \frac{|v(m)| - \min_{y \in N_s}|v(y)|}{\max_{y \in N_s}|v(y)|}$$

where $N_s$ is the neighbor set of node $s$. A lower $MF$ indicates a lower velocity. The update of $MF$ is conducted periodically at an interval of one second based on a weighted exponential moving average as

$$MF(s, m) \leftarrow (1-\alpha) \times MF_{i-1}(s, m) + \alpha \times MF_i(s, m),$$

where $MF_{i-1}(s, m)$ and $MF_i(s, m)$ denote the previous value and current value of $MF$, respectively. $MF$ is initialized to 1, and $\alpha$ is set to 0.7 based on our simulation results. The value of $\alpha$ controls how quickly the evaluation value can adapt to the change of network topology. If the value is too small, it is difficult to adapt quickly to network dynamics. In contrast, if the value is too large, then the algorithm cannot reflect the network tendency.

- **Leadership Factor (LF):** $LF(s, m)$ is calculated as

$$LF(s, m) = \frac{c(s)}{\max_{y \in N_s}c(y)}.$$  

$c(s)$ shows the number of vehicles moving toward the same direction as the node $s$. A higher $LF$ means that the node is more suitable for being a cluster head node. The initial value of $LF$ is 0. For every hello message reception, $LF$ is updated using a weighted exponential moving average as

$$LF(s, m) \leftarrow (1-\alpha) \times LF_{i-1}(s, m) + \alpha \times LF_i(s, m).$$

- **Signal Quality Factor (SQF):** For simplicity, we calculate the signal quality factor (SQF) using the hello packet
reception ratio. The signal quality factor differentiates different vehicles by setting a higher evaluation value on a vehicle that has better signal condition (for example a cluster head vehicle with higher antenna could provide longer line-of-sight distance to cluster members and therefore could result in a higher hello packet reception rate between the cluster head and members). The \( SQF \) is initialized to 0 and is updated as

\[
SQF(s, m) = (1 - \alpha)SQF_{i-1}(s, m) + \alpha \times SQF_i(s, m).
\]

2) Second step – fuzzification and fuzzy rules: The fuzzy membership functions are defined as shown in Fig. 4. The linguistic variables of the mobility factor are defined as \{Slow, Medium, Fast\}. Similarly, the linguistic variables for the leadership factor and signal quality factor are defined as \{Good, Fair, Poor\} and \{Good, Medium, Bad\}, respectively.

![Fig. 4. Fuzzy membership functions (left: MF, middle: LF, right: SQF).](image)

3) Last step – defuzzification: The output membership function is defined as shown in Fig. 5. The center of gravity (COG) method is used for the defuzzification. As shown in Fig. 5, the value of the \( x \) coordinate corresponding to the centroid is the final defuzzified value which indicates the competency value of the node.

![Fig. 5. Output membership function.](image)

### TABLE I

<table>
<thead>
<tr>
<th>Rule</th>
<th>Mobility</th>
<th>Leadership</th>
<th>Signal quality</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule 1</td>
<td>Slow</td>
<td>Good</td>
<td>Good</td>
<td>Perfect</td>
</tr>
<tr>
<td>Rule 2</td>
<td>Slow</td>
<td>Good</td>
<td>Medium</td>
<td>Good</td>
</tr>
<tr>
<td>Rule 3</td>
<td>Slow</td>
<td>Good</td>
<td>Bad</td>
<td>Unpreferable</td>
</tr>
<tr>
<td>Rule 4</td>
<td>Slow</td>
<td>Fair</td>
<td>Medium</td>
<td>Good</td>
</tr>
<tr>
<td>Rule 5</td>
<td>Slow</td>
<td>Fair</td>
<td>Medium</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Rule 6</td>
<td>Slow</td>
<td>Medium</td>
<td>Medium</td>
<td>Good</td>
</tr>
<tr>
<td>Rule 7</td>
<td>Slow</td>
<td>Poor</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 8</td>
<td>Slow</td>
<td>Poor</td>
<td>Bad</td>
<td>VeryBad</td>
</tr>
<tr>
<td>Rule 9</td>
<td>Slow</td>
<td>Poor</td>
<td>Bad</td>
<td>VeryBad</td>
</tr>
<tr>
<td>Rule 10</td>
<td>Medium</td>
<td>Good</td>
<td>Medium</td>
<td>Good</td>
</tr>
<tr>
<td>Rule 11</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Rule 12</td>
<td>Medium</td>
<td>Good</td>
<td>Bad</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 13</td>
<td>Medium</td>
<td>Fair</td>
<td>Good</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Rule 14</td>
<td>Medium</td>
<td>Fair</td>
<td>Medium</td>
<td>Unpreferable</td>
</tr>
<tr>
<td>Rule 15</td>
<td>Medium</td>
<td>Fair</td>
<td>Bad</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 16</td>
<td>Medium</td>
<td>Poor</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 17</td>
<td>Medium</td>
<td>Poor</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 18</td>
<td>Medium</td>
<td>Poor</td>
<td>Bad</td>
<td>VeryBad</td>
</tr>
<tr>
<td>Rule 19</td>
<td>Fast</td>
<td>Good</td>
<td>Good</td>
<td>Unpreferable</td>
</tr>
<tr>
<td>Rule 20</td>
<td>Fast</td>
<td>Good</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 21</td>
<td>Fast</td>
<td>Good</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 22</td>
<td>Fast</td>
<td>Fair</td>
<td>Good</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 23</td>
<td>Fast</td>
<td>Fair</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 24</td>
<td>Fast</td>
<td>Fair</td>
<td>Medium</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 25</td>
<td>Fast</td>
<td>Poor</td>
<td>Good</td>
<td>Bad</td>
</tr>
<tr>
<td>Rule 26</td>
<td>Fast</td>
<td>Poor</td>
<td>Medium</td>
<td>VeryBad</td>
</tr>
<tr>
<td>Rule 27</td>
<td>Fast</td>
<td>Poor</td>
<td>Bad</td>
<td>VeryBad</td>
</tr>
</tbody>
</table>

Each node calculates the rank (a competency value for being a cluster head) of each neighbor based on the IF/THEN rules as defined in Table I. The linguistic variables for the rank are defined as \{Perfect, Good, Acceptable, Unpreferable, Bad, VeryBad\}. In Table I, Rule 1 is expressed as follows:

**E. Coalitional game approach for stimulating clustering and better route selection**

1) Game theoretical model: The objective of this model is to stimulate the vehicles to use cluster head nodes for multi-hop data transmissions, resulting in better route selection. The payoff will be initialized by the RSU and will be further disseminated by cluster head nodes in order to disseminate the payoffs to a multi-hop distance. Each payoff allocator (RSU or cluster head) assigns payoffs equally to the one-hop neighbors. The set of players are the one-hop neighbors, which are denoted by \( \mathcal{N} = \{1, ..., N\} \). The coalitional game is defined as \( \langle \mathcal{N}, v \rangle \), where \( v \) is the coalition value, which indicates the worth of a coalition. Any coalition \( S \subseteq \mathcal{N} \) represents an agreement between vehicles on the use of cluster-based forwarding. All nodes can be benefited from using cluster head based forwarding which could reduce the number of sender nodes in the network and consequently improve the throughput of all nodes. Therefore, it is plausible to define \( v \) based on the average collision probability for the MAC layer backoff (since cooperation can result in lower collision probability, there would be incentives for nodes to form a coalition). According to the IEEE 802.11p standard, the collision probability of \( n \) nodes contending the same wireless medium can be calculated as

\[
P(n) = \begin{cases} 
1, & \text{if } CW + 1 \leq n \\
1 - \frac{\prod_{k=1}^{n} (CW+1-k)}{(CW+1)^n}, & \text{otherwise}
\end{cases}
\]

where \( CW \) is the collision window and \( n \) is the number of contending nodes.
where $CW$ is the current contention window size. We define $v(S) = \left(1 - \frac{P[|S|-1]}{2}\right) \cdot N$ where $|S|$ represents the cardinality of set $S$. For simplicity, we set the value of payoff to player $i \in S$ as

$$x_i = \begin{cases} \frac{v(S)}{N}, & \text{cooperate}, \\ 0, & \text{acting alone}. \end{cases} \quad (7)$$

**Theorem 1:** The proposed game is a game with transferable utility (TU).

**Proof:** In the proposed game, the value of coalition $S$ depends solely on the number of vehicles forming the coalition. The value is independent of how the players in $N \setminus S$ are structured. Therefore, the game is a TU game. □

**Theorem 2:** The game is superadditive.

**Proof:** The formation of a larger coalition means a reduction in the number of sender nodes, resulting in a better channel contention. Therefore, the cooperation is beneficial to all the involved players. That is, no one can do better by acting noncooperatively. Therefore, the game is superadditive. □

**Theorem 3:** The core of the game is not empty, and the payoff vector defined by Eq. (7) is an imputation.

**Proof:** For the proposed game, given the grand coalition $N$, the payoff vector $x \in \mathbb{R}^N \setminus \{N\}$ defined by (7) is group rational since it satisfies $\sum_{i \in N \setminus S} x_i = v(N)$. The payoff vector $x$ is also individually rational because every player can receive a payoff no less than acting noncooperatively $[x_i \geq v(i), \forall i]$. Therefore, the payoff vector defined by Eq. (7) is an imputation, and the core of the game is not empty. □

2) **Distributed implementation of payoff allocation:** Each cluster head (or RSU) allocates payoffs equally to the one-hop neighbors. This means that when a vehicle is a neighbor of the RSU, the vehicle will receive a payoff directly from the RSU. Otherwise, the vehicle has to select a cluster head located closer to the RSU compared with itself. Since only the cluster head or RSU distributes the payoff, each vehicle would like to cooperate (using the cluster head for data forwarding). This also results in a grand coalition because there is no way that cluster members could form a different coalition that would give them larger payoffs.

**F. Multi-hop route selection based on reinforcement learning and game theory-based metric**

1) **Q-learning model:** We use a Q-learning algorithm to evaluate a multi-hop route. The following Q-learning model is defined. The environment is the entire network. The network nodes are the learning agents, which learn the environment by exchanging hello messages with each other. The action at each node is to select the next hop node for the data transmission. Therefore, the set of one-hop neighbors is the possible actions allowed at each node. Each node maintains a Q-Table where each $Q$-value $[Q(RSU, m)]$ shows the value for choosing $m$ as the next hop to the RSU.

2) **Update of $Q$-values:** Each node has to maintain a $Q$-value for each pair of the destination node (RSU) and a one-hop neighbor. Upon reception of each hello message, the Q-Table is updated. $Q$-values are attached to the hello messages and broadcasted by all nodes. The initial value for each $Q$-value is 0. After reception of a hello message from node $m$, node $l$ updates the corresponding $Q$-value to the RSU as

$$Q_l(RSU, m) \leftarrow \alpha \times P_m \times SQF(l, m) \times \left[\hat{R} + \gamma \times \max_{y \in N_{BS}^m} Q_m(r, y)\right] + (1 - \alpha) \times Q_l(RSU, m). \quad (8)$$

where $SQF(l, m)$ is the link status value between node $l$ and $m$, as calculated by Eq. (5). $N_{BS}^m$ denotes the one-hop neighbor set of node $m$. The payoff $P_m$ is calculated as

$$P_m = \begin{cases} (1 - \frac{P(N_{BS}^m)}{2}), & \text{if } m \text{ is a cluster head} \\ 0, & \text{otherwise}. \end{cases} \quad (9)$$

The learning rate ($\alpha$) is 0.7, and the discount factor ($\gamma$) is 0.9. “$\max_{y \in N_{BS}^m} Q_m(RSU, y)$” is the maximal $Q$-value of $m$ to node RSU. The reward $\hat{R}$ is calculated as

$$\hat{R} = \begin{cases} 1, & \text{if } l \in N_{BS}^R \text{ and } m \text{ is a neighbor of RSU} \\ 0, & \text{otherwise}. \end{cases} \quad (10)$$

where $N_{BS}^R$ denotes the one-hop neighbor set of the RSU excluding cluster head nodes. If node $m$ is a neighbor of the RSU, the reward is 1 and otherwise 0. Each $Q$-value is an evaluation value for each pair of state and action. Upon reception of a hello message, each agent updates the corresponding $Q$-value as shown in Eq. (8).

The reward is discounted by three elements that are the number of hops from the RSU, the payoff calculated by the game theory algorithm [$P_m$ in Eq. (8)], and the link quality [$SQF(l, m)$ in Eq. (8)]. The consideration of hop count ensures that the routing protocol intends to choose a shorter route. First, the algorithm discounts the reward with the hop count. Therefore, a smaller hop count results in a larger reward and larger $Q$-value. Second, the reward is also discounted with the payoff calculated by the game theory algorithm [$P_m$ in Eq. (8)]. This ensures that our protocol can stimulate each node to use a cluster head node to forward data packets because only that can achieve a nonzero payoff. Third, the reward is also discounted with the packet loss probability of each link [$SQF(l, m)$ in Eq. (8)], which constitutes the communication route.

![Fig. 6. An example of Q-table update.](image-url)
Since the routing agent selects the node that shows the largest $Q$-value as the next forwarder node, the protocol is able to choose the route expected to achieve the best performance in terms of multi-hop transmissions.

**G. Multi-hop data delivery virtualization**

![Virtual multi-hop data delivery](image)

Fig. 7. Multi-hop data delivery virtualization (the numbers after “Data” and “ACK” show the sending order of data and acknowledgment respectively).

Most vehicular IoT applications require a reliable end-to-end data delivery. TCP is a widely used end-to-end acknowledgment based data delivery approach. In TCP, the sender node adapts the sending rate, namely, congestion window size, by using a congestion avoidance algorithm according to the acknowledgment packets successfully received. The throughput of a TCP connection is affected significantly by the end-to-end packet loss probability. As a result, TCP has a performance degradation problem in a lossy multi-hop networks where a high end-to-end packet loss ratio restricts the increase of the congestion window size at the sender node. To solve this problem, we use a multi-hop data delivery virtualization approach [32] to conduct multi-hop data transmissions through multiple one-hop TCP sessions. The virtualization approach works on the top of transport layer. As shown in Fig. 7, a multi-hop data transmission from vehicle 1 to the road side unit $N$ is conducted by multiple one-hop data transmissions where each hop is managed by a TCP session, which ensures the reliability and fairness while facilitating an efficient use of wireless resources. Each forwarder node is responsible for transmitting data segments to the next hop that is closer to the destination. Since each one-hop communication is conducted based on TCP, the congestion window size at each TCP sender node is not affected by the number of hops between the source node and the destination node.

**IV. Simulation Results**

To clearly explain the performance of the proposed protocol, we conducted extensive simulations using ns-2.34 [33] in two types of roads, namely, a freeway scenario and a street scenario (see Table II). The freeway had three lanes in each direction [34]. The width of each lane was set to 4m. The street scenario was generated by SUMO [35] and TraNS [36]. The street configuration consisted of five horizontal streets and five vertical streets and every street had one lane in each direction. The distance between any two neighboring intersections was 400 m. We introduced the Nakagami propagation model to simulate a realistic fading vehicular environment (see Table III) [37]. The transmission range was approximately 250 m, and the packet reception probability for various distances was as shown in Fig. 8. Other simulation parameters were the default setting of ns-2.34 [33]. The proposed protocol was compared with “AODV-ETX” [38] (no-cooperative approach) and “CDS-SVB” [17]. “Proposed” and “Proposed-V” show the proposed protocol without and with multi-hop data delivery virtualization respectively. Each simulation was conducted 50 times for different topologies, and the error bars of the following figures show the 95% confidence intervals.

**TABLE II**

<table>
<thead>
<tr>
<th>Simulation Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology</td>
</tr>
<tr>
<td>Number of nodes</td>
</tr>
<tr>
<td>Maximum velocity</td>
</tr>
<tr>
<td>Mobility generator</td>
</tr>
<tr>
<td>MAC</td>
</tr>
<tr>
<td>Fading model</td>
</tr>
<tr>
<td>Simulation time</td>
</tr>
</tbody>
</table>

**TABLE III**

<table>
<thead>
<tr>
<th>Parameters of Nakagami Model: Freeway (Street)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_1$</td>
</tr>
<tr>
<td>1.9 (2.0)</td>
</tr>
</tbody>
</table>

![Packet Reception Probability](image)

Fig. 8. Packet reception probability for various distances (left: freeway scenario, right: street scenario)

A. Effect of coalition

To show the performance of the game theory-based forwarding algorithm, we first used a static environment with constant bit rate (CBR) traffic to evaluate the protocols. There were 100 nodes in this simulation, and the route length between the source node and the destination node was 2.

Fig. 9 shows the packet delivery ratio for various numbers of CBR flows. The performance of “Non-cooperative” approach (“AODV-ETX”) drops significantly when the number of traffic flows increases to 8. The reason can be explained by Fig. 10 and Fig. 11.

Fig. 10 shows the number of collided MAC frames (including hello messages) per transmission range per second for various numbers of CBR flows. Without cooperation, different traffic flows are likely to use different nodes to forward data...
packets and thus increase the probability of collisions between MAC frames. The collisions result in a waste of wireless resources due to the exponential backoff procedure at the MAC layer, which finally leads to the drop of packet delivery ratio. The proposed protocol can keep a very low number of packet collisions even when the number of traffic flows is large by stimulating cooperation among multiple sender nodes. The utilization of the same cluster head node can achieve a significantly lower end-to-end delay as shown in Fig. 11.

B. Performance for various numbers of TCP flows

We also evaluated the performance of the proposed protocol for various numbers of TCP flows. The TCP segment size was 1000 bytes. Fig. 12 and Fig. 13 show the TCP throughput for various numbers of TCP flows in the freeway scenario and the street scenario, respectively. The number of hops from the source node to the destination node was 2.

The proposed protocol can achieve a notable improvement of TCP throughput over “AODV-ETX” and “CDS-SVB” in both the freeway scenario and the street scenario. In the street scenario, the number of TCP flows has less impact on the throughput because the simulation area is larger, which results in a lower contention probability between different flows. Since “CDS-SVB” does not consider link quality in the back-
bone selection, the generated end-to-end route is unsatisfactory in terms of end-to-end packet delivery ratio. In “CDS-SVB”, the backbone selection is conducted by one by one, where each backbone node selects the next backbone node. This incurs a higher overhead when the network topology changes, which explains the high number of MAC collisions in Fig. 14.

Since the proposed protocol considers vehicle mobility and signal qualities between vehicles in the cluster head selection, a highly reliable end-to-end route can be established. The cluster head-based forwarding approach ensures that the proposed protocol can efficiently use the wireless resources by reducing the number of concurrent sender nodes in the network. The performance advantage of “Proposed-V” over “Proposed” explains the importance of the multi-hop data delivery virtualization approach. These three elements, specifically the reliable end-to-end route selection, cluster head-based forwarding, and multi-hop data delivery virtualization, jointly contributed to the lowest end-to-end delay shown in Fig. 15.

C. Performance for various numbers of hops

We conducted simulations for various numbers of hops. There were five TCP flows from random source nodes to random destination nodes. Fig. 16 and Fig. 17 show the TCP throughput for various numbers of hops in the freeway scenario and the street scenario respectively.

Fig. 14. Number of collided MAC frames (including hello messages) per second per transmission range for various numbers of TCP flows (freeway).

Fig. 15. End-to-end delay for various numbers of TCP flows (freeway).

Fig. 16. TCP throughput for various numbers of hops (freeway).

Fig. 17. TCP throughput for various numbers of hops (street).

We can observe that the cluster head-based forwarding and the multi-hop data delivery virtualization scheme become more effective in terms of throughput and delay when the number of hops increases. This is because TCP is an acknowledgment-based packet delivery approach that is sensitive to the end-to-end packet loss ratio. In the street scenario, all protocols are more sensitive to the number of hops. This is due to a higher probability of route breaks due to the changes of moving directions at intersections. Since the proposed protocol can improve the end-to-end packet delivery probability by selecting a better route and reducing the MAC collision ratio at each hop, a higher end-to-end performance becomes possible (see Fig. 18 and Fig. 19). When the average number of hops of TCP flows increases, the number of concurrent TCP
traffic flows could increase and the advantage of the proposed protocol over other protocols thus becomes more significant, as was previously discussed (see Fig. 12).

V. CONCLUSIONS AND FUTURE WORK

We proposed a V2R cluster-based routing protocol where a coalitional game theory-based routing metric is used to stimulate the vehicles to use cluster heads for the data transmission, and a Q-learning approach is used to evaluate a route in terms of multi-hop efficiency. The cluster heads are selected using a fuzzy logic-based algorithm by considering vehicle movement, vehicle distribution, and channel qualities between vehicles. A multi-hop data delivery virtualization scheme is integrated with the routing protocol to ensure a high throughput and low delay in multi-hop lossy vehicular environments. We showed and analyzed the performance advantage of the proposed protocol over possible alternatives by conducting extensive computer simulations.

Currently, the fuzzy membership functions are designed based on our experience. In future work, we will consider adaptive tuning of the fuzzy parameters using online learning algorithms such as reinforcement learning. Since it possible to achieve better quality-of-service with data-driven optimization, we will also discuss how to improve the performance of V2I communications based on the data collected from vehicles and roadside units.

ACKNOWLEDGMENT

This research was supported in part by the open collaborative research program at the National Institute of Informatics (NII) Japan (FY2018), JSPS KAKENHI Grant Numbers 16H02817, and 16K00121.

REFERENCES
