Research Article

Coalitional Game Theoretical Approach for VANET Clustering to Improve SNR

Selo Sulistyo [ID], **Sahirul Alam, and Ronald Adrian**

Department of Electrical Engineering and Information Technology, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

Correspondence should be addressed to Selo Sulistyo; selo@ugm.ac.id

Clustering is considered as the potential approach for network management in vehicular ad hoc network (VANET). The performance of clustering is often assessed based on the stability of the clusters. Hence, most of the clustering methods aim to establish stable clusters. However, besides the stability of cluster, good link quality must be provided, especially when reliable and high-capacity transmission is demanded. Therefore, this paper proposes a clustering method based on coalitional game theory with the purpose to improve the average of vehicle-to-vehicle (V2V) signal-to-noise ratio (SNR) and channel capacity while maintaining the stability of the cluster. In the proposed method, each vehicle attempts to form a cluster with other vehicles according to coalition value. To attain the purpose of clustering, the value of coalition is formulated based on the V2V SNR, connection lifetime, and speed difference between vehicles. In fast-changing network topology, the higher average of SNR can be achieved but the stability of cluster becomes hard to be maintained. Based on the simulation results, SNR improvement can be adjusted in order to balance with the cluster stability by setting the parameters in the proposed method accordingly. Further simulation results show that the proposed method can obtain a higher average of V2V SNR and channel capacity than other relevant methods.

1. Introduction

Vehicular ad hoc network (VANET) is a new form of mobile ad hoc networks comprised of vehicles as the nodes. Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications are enabled by equipping vehicles with dedicated transceiver device. VANET is provisioned to contribute to intelligent transportation systems by providing a communication link among vehicles to support safety and traffic management purposes as well as infotainment [1]. The higher network capacity and reliable transmission link are demanded to support these purposes. However, the environments of VANET bear the dynamics of the networks and raise several technical challenges. The wide range of vehicle speed causes the fast-changing network topology, and hence, the V2V connection is difficult to be maintained. The sparse-density nodes such as in rural road cause limited connectivity

[2]. Meanwhile, the high-density nodes such as in urban road cause the burst of data traffic that leads to higher transmission collision and delay [3].

Node clustering with the hierarchical topology is considered as the potential approach for network management in VANET. A cluster consists of several vehicle nodes with some similar characteristics which form a communication group. A cluster commonly has one cluster head that acts as the centre of the group and manages the cluster. Other than the cluster head, the ordinary vehicle nodes are called cluster members. Despite the effectiveness of clustering for network management in VANET, clustering has some adversities mainly due to the dynamic environment. In a survey [4], the performance of clustering is generally assessed according to the cluster stability. It is good to have a stable cluster. However, it is insignificant unless the reliable and high-capacity communication link can be provided.

Therefore, in this paper, a clustering method for VANET is proposed with the aim to improve the signal-to-noise ratio (SNR) and channel capacity while maintaining the stability of the cluster. The proposed clustering method uses the distributed approach based on coalitional game theory. In general, the coalitional game exhibits the model of interaction between a set of players who attempt to form a group or coalition to reinforce their standing in the game [5]. This concept matches with the investigated problem in this paper, where a vehicle interacts with the other nearby vehicles to establish connection and to form a cluster.

As the implementation of the coalitional game in VANET, each vehicle node is considered as a player who attempts to form a coalition (cluster) with other vehicle nodes based on coalitional value. The value of coalition is determined by the revenue and cost. Based on the purpose of the proposed clustering method, the revenue is the link quality or the SNR of V2V connection. Meanwhile, the cost is the variable that decreases the revenue. To obtain the higher SNR, the vehicles need to change the V2V connection frequently. Thus, the stability of the cluster becomes harder to be maintained. For this reason, the cost in the coalitional game is aimed to maintain the stability of the cluster. The cost is formulated based on the connection lifetime and the speed difference between vehicles. The main contributions of this paper and the advantages of the proposed clustering can be described as follows.

(1) A new clustering method for VANET based on coalitional game theory is proposed with the aim to improve SNR and channel capacity. The proposed clustering method also provides the flexibility on balancing between high link quality and stability of the cluster by adjusting some parameters. Thus, the impact of parameters in the proposed method and the guide for adjusting the parameters are also presented in this paper.

(2) The proposed clustering method uses a distributed approach which is favorable due to the flexibility, especially for VANET environment. It is also expected to reduce the workload of the network central, i.e., the cluster head. In this distributed clustering, each vehicle only requires information about the link quality and the speed of neighborhood vehicles. This information is more conveniently provided than the exact position of vehicles. Thus, some assumptions such as the use of GPS can be avoided. The use of GPS is still disputable due to the accuracy problem.

(3) This paper presents the simulation of clustering in VANET using a real-world map and realistic vehicle mobility based on Simulation of Urban Mobility (SUMO). The steps to build the simulation are presented in detail. Simulations of clustering are performed using dedicated program built in MAT-LAB. Thus, this work can be focused in the physical layer of communication by specifically defining the signal propagation model and the metrics for performance evaluation.

2. Related Work

Some research studies related with clustering and game theory in ad hoc networks, especially VANET, are presented in this section. Some technical differences between those research studies and this paper are also presented. The technical differences can be outlined based on the criteria to form a cluster, cluster control (centralized or distributed), cluster head selection procedures, and goal of clustering.

The metric to form a cluster used by most of the clustering methods is the distance between vehicles such as in [6–12]. The calculation of distance is based on the vehicle's position, and thus, it is assumed that each vehicle is equipped with GPS. However, the usage of GPS is still disputable due to the accuracy problem. In this research, the proposed method avoids the use of distance between vehicles as the metric to form a cluster. Another metric used to form a cluster is the vehicle speed, where the adjacent vehicles with lower speed difference are grouped into the same cluster. Vehicle speed is used as one of the metrics to form a cluster in [6, 11, 13], and even the proposed method in this paper also uses this metric. The reason is that the vehicle speed is the dominant factor affecting the stability of the cluster. Another metric to form a cluster affecting the stability of cluster is the connection lifetime, such as that used in [6, 8] and the proposed method in this paper. The other distinct metrics to form a cluster are the node hierarchy [14] and neighborhood follow [15].

Most of the clustering methods including this research perform clustering in a distributed manner. However, clustering can also be performed using a centralized approach with the aid of infrastructures, such as in [12], where the clustering is performed by the cellular base station (BS). The clustering method in [13] is also centralized since the fuzzy system requires information from all vehicles to be processed by the centre of the networks. Most of the clustering methods also assume the single-hop cluster structure for the sake of simplicity. The multihop cluster structure, despite the complexity, has either advantages or disadvantages. A multihop cluster enables the flexibility and increases the coverage of the cluster. However, the delay of transmission may increase due to the multihops transmission. Therefore, the number of hops is limited to a certain number to maintain effectiveness. The proposed method in this paper assumes the multihop clustering.

Cluster head selection in some clustering methods uses the same metrics as in cluster formation, while other methods use more specific metrics for cluster head selection. The specific metric mostly used for cluster head selection is the node connectivity, such as in [13, 14]. Node connectivity is equal to the number of neighbor vehicles within the transmission range. Thus, the vehicle which has most neighbor vehicles is considered as the best candidate of the cluster head. The other metrics to select the cluster head are as follows. In [6], the cluster head is selected based on the geographical position, i.e., the vehicle that is located at the centre of a cluster. The speed of vehicle closest to the average speed of cluster is selected as the cluster head such as in

[12, 13]. In [7], the cluster head is selected based on the lowest vehicle ID, while in [8], it is selected based on the node weight which is calculated based on the history of link quality. In [10], the cluster head is the first vehicle that forms a cluster. Afterwards, the concept of secondary cluster head (SCH) is introduced. When the cluster head leaves the cluster, the SCH can switch position to be the cluster head. The SCH is selected based on the minimum speed difference and nearest distance to the primary cluster head. In [11], the cluster head is selected based on eligibility which is defined by a fuzzy logic controller. In [16], the cluster head is selected based on the connection with minimum overhead. In [12], besides based on the average cluster speed, the cluster head is selected from the vehicle which has the best channel quality to establish a connection with the base station. In [15], the cluster head is selected based on neighborhood follow, that is, the vehicle followed by other vehicles. In this proposed clustering method, cluster head selection is performed based on the electability, i.e., the vehicle which is selected by most of the cluster members. The further description of electability is presented in the next section.

The goal of most clustering methods is to form the stable cluster [6–9, 11, 13–15]. The clusters with high stability can support the performance of routing protocol. However, the stability of the cluster becomes less significant if the high link quality cannot be provided, especially when the reliable or high-capacity transmission is demanded. Therefore, some clustering methods aim different goals such as high data packet delivery ratio (DPDR) and low delay [16], reducing signal overhead and improving communication quality by aggregating V2I traffic transmission to the cluster head [12]. Meanwhile, the goals of clustering in this proposed method are to improve the average of V2V SNR and channel capacity. This goal actually can be achieved by establishing V2V connection with the higher SNR. However, in the environment of VANET which is very dynamic, this can cause the vehicles change the connection frequently and thus ruining the stability of the cluster. Therefore, a special approach is needed to deal with this complex problem.

This research proposes the coalitional game theory as the approach for clustering in VANET. Game theory is known as the potential approach for wireless communication and expected to give contribution in VANET development. In [17], game theory is proposed for ad dissemination in VANET. In [18], game theory is used to observe the cooperation in VANET according to the network conditions. In [19], game theory is proposed to jointly control the power level and channel selection in VANET. The coalitional game theory was also implemented in other wireless communication systems such as cellular and WiMAX. In [20], coalitional game was utilized in multiple-input multiple-output (MIMO) cellular network to form the network and to mitigate the interference at once. Each established communication link has interference relationship with other links. Hence, the establishment of communication link was modeled as coalitional game with an aim to mitigate the interference. The proposed algorithm converged to the Nash equilibrium to enable network formation with lower interference and higher data rate. In [21], coalitional game was utilized to arrange the merge process in cooperative communications with aim to attain energy efficiency. The proposed merge process consisted of three stages, i.e., transmission request, merge, and cooperative transmission. With the coalitional game, the merge process was directed to attain energy efficiency of each node in the group. Therefore, the total power (transmission and processing power) in cooperative transmission can be lower than direct transmission.

In this paper, the coalitional game theory is selected due to flexibility to formulate the rule in forming a coalition. Moreover, the concepts of coalition and clustering are compatible. The coalitional game theory has been used in clustering for mobile ad hoc networks (MANETs) [22]. However, besides the difference of system model between MANET and VANET, there is a key difference between research in [22] and this research, that is, the formulation of the coalitional rule. Another difference is that research in [22] does not have a cluster head in cluster structure, whereas this research has.

Research in this paper utilizes the trace data from real-world map. The usage of trace data has become an alternative for simulating VANET environment, and it has an advantage due to the proximity with the real condition. Another related work in VANET using the trace data was presented in [23]. In this paper, the trace data from Google maps is used to simulate vehicle mobility while forming the clusters. Meanwhile, the trace data in [23] were used for simulation of route selection which considers the data rate in route selection. Since the aim of route selection is to optimize throughput of the network using TV white space, the usage of trace data and Google spectrum data set is appropriate in this case.

3. Proposed Approach

This paper proposes a multihop clustering method. In multihop clustering, the stability of connection between vehicles is the key factor which determines the stability of the cluster. Therefore, the establishment of V2V connection must be arranged appropriately to attain the longer connection duration and the better link quality. Furthermore, the formation and maintenance of the cluster rely on the established V2V connection. In this proposed method, coalitional game theory is employed to improve the mechanism of V2V connection establishment due to the relevance of the cluster formation model and coalition formation model. The vehicles aim to form cluster by establishing V2V connection with other vehicles which can give mutually good link quality and connection lifetime. This concept is similar to the concept of coalition formation game, where each player selects other players to form group which can give mutual benefit to the group members. Therefore, the coalitional game is considered as the potential approach to be implemented in VANET clustering.

The problem of clustering investigated in this paper can be described as an optimization problem as follows. There are two objectives of the optimization, namely, to improve the SNR of V2V connection and to establish the stable

clusters. However, each of those objectives is a trade-off for the other objective. This is due to the dynamic environment in VANET, thus the fluctuation of SNR is high and the vehicles tend to change connection frequently in order to obtain the higher value of SNR. Meanwhile, the frequent change of connection disrupts the stability of cluster. Therefore, this is a complex optimization problem. However, as mentioned above, the coalitional game can be the appropriate approach to model this optimization problem. Here, the variables related with SNR value such as transmission distance and power are included in the revenue of coalition. Meanwhile, the variables influencing the stability of V2V connection as well as the stability of cluster such as the speed difference and connection lifetime are included in the cost of coalition. Finally, the vehicles will decide whether to maintain current connection or change the connection based on the value of coalition which is the difference between revenue and cost of coalition. Therefore, the objective of optimization is directed to maximize the value of coalition. Furthermore, the details about the proposed clustering method including the implementation of coalitional game approach in the clustering process are described as follows. The role of coalitional game approach in the clustering process can be seen in Figure 1.

3.1. Vehicle-to-Vehicle Establishment.
Coalitional game theory approach is used for establishment of vehicle-to-vehicle (V2V) connection. Therefore, the proposed clustering approach is called coalitional game clustering (CGC). In the coalitional game, players select any other player for cooperating or forming a coalition based on the value they expect to get. The value of the coalition has two elements, i.e., revenue and cost. High value is implied by the high revenue and low cost. The implementation of this approach in V2V connection establishment is described as follows.

3.1.1. Revenue.
The main purpose of this clustering approach is to obtain a higher transmission link quality. Since this research is done in the physical layer of communication, the transmission link quality observed is the signal-to-noise ratio (SNR) and the channel capacity obtained from the establishment of a V2V connection. Channel capacity is proportional to the SNR based on the Shannon–Hartley theorem. Therefore, the proposed revenue function is based on SNR as given by

$$u_j\left(S_j^i\right) = \Gamma_{ij}, \tag{1}$$

where $u_j(S_j)$ is the revenue of vehicle j for using strategy S_j^i and Γ_{ij} is the SNR of downlink transmission from vehicle i to vehicle j. The value of SNR for the revenue is presented in watt instead of decibel since the minimum revenue is expected to be equal to zero instead of negative. The term strategy in game theory means a set of actions that can be selected by a player. In this case, the strategy is the action to select any other vehicles than itself to form a coalition (establish a V2V connection). The list of strategies that can be chosen by vehicle j can be denoted as
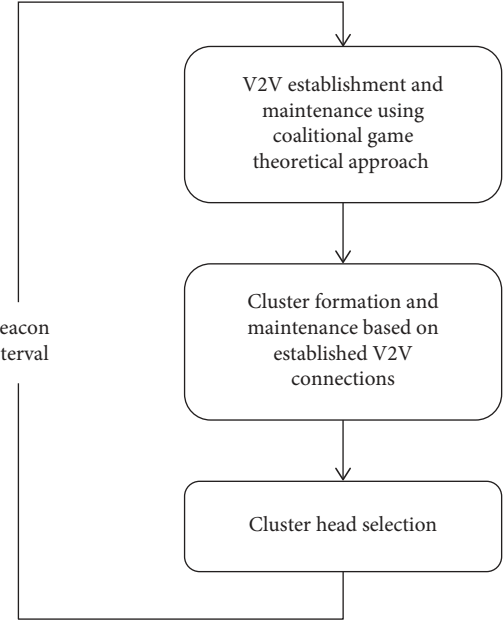

Figure 1: Procedure in the proposed clustering method.

$$S_j = S_j^1, S_j^2, \ldots, S_j^i, \ldots, S_j^N, \tag{2}$$

where i and N are, respectively, the index of transmitter vehicle and the total number of other vehicles within the transmission range of vehicle j. It can be noted that the number of strategies depends on the traffic around vehicle j. In a normal or sparse traffic, the number of strategies can still be handled by the vehicle. However, in high-density traffic, the number of strategies can be numerous and difficult to handle. Therefore, it is suggested to limit the number of strategy especially in a dense traffic condition, e.g., by listing the vehicles within a certain radius from vehicle j or by defining the minimum threshold of SNR so that the vehicles will be included in the list.

3.1.2. Cost.
Cost is considered as the amount or price that reduces the revenue. In this proposed clustering approach, the cost is formulated to increase the lifetime of a V2V connection. In other words, it is aimed to maintain the stability of clusters. In this CGC, the cost consists of two elements as follows.

(i) Connection lifetime (t_c) represents the duration of V2V connection that has been established. The duration increases as long as the vehicle is connected with the same transmitter vehicle. Otherwise, if the vehicle changes connection to another vehicle, then t_c is reset. It can be understood that t_c directly represents the lifetime of a V2V connection. Therefore, in cost formulation, when the value of t_c is lowest (e.g., after reset), then the cost must be highest. Hence, the vehicle hardly changes the connection to another vehicle if it has just established a connection with a vehicle. Afterwards, as t_c increases, the cost will decrease. Cost element 1 (c_1) based on t_c is defined as follows:

$$c_1 = \begin{cases} 0, & t_c \geq \delta_c, \\ \dfrac{\delta_c - t_c}{\delta_c}, & t_c < \delta_c, \end{cases} \quad (3)$$

where δ_c is the threshold for t_c. In other words, δ_c is the minimum time required by a vehicle to change V2V connection to another vehicle with c_1 equal to zero.

(ii) Speed difference (Δv) between two vehicles is considered as one of the factors affecting the lifetime of a V2V connection. Vehicles with the higher difference of speed will terminate the connection sooner since they leave their transmission range. On the contrary, vehicles with the lower difference of speed can maintain their connection for the longest time. In cost formulation, the absolute value of Δv is proportional to the cost element 2 (c_2) as given by

$$c_2 = \max\left[1, \left(\frac{|\Delta v|}{\Delta v_{max}}\right)\right], \quad (4)$$

where Δv_{max} is the defined maximum speed difference, and thus, c_2 will be more than 1 if $|\Delta v|$ is higher than Δv_{max}. However, the value of c_2 is limited to 1. The purpose of using absolute operator is due to the vectorization of speed, where vehicles that move in the opposite direction have opposite sign of speed. For example, a vehicle moves from south to north with speed v_1, thus another vehicle moving from north to south has speed $-v_2$. Therefore, two vehicles moving in direction opposite to each other have a higher absolute value of speed difference. This is relevant to the fact that two vehicles moving in different directions can only maintain the V2V connection in a short time. The higher cost can avoid the establishment of a V2V connection between vehicles, which are moving in the opposite direction, as long as there is at least another vehicle moving in the same direction within their transmission range.

The two cost elements in (3) and (4) are formulated so that the values are normalized, i.e., between 0 and 1. Afterwards, the combination of those cost elements to define the final cost for CGC is given by

$$c_j\left(S_j^i\right) = \max[c_1, c_2]. \quad (5)$$

The reason for using max operator is to allow each element to stand out depending on the condition. For example, when the speed difference is above the threshold (c_2 is maximum), c_j will have a maximum value regardless of the value of c_1. In another case, when some neighbor vehicles move with nearly same speed (thus the speed difference is very low), then c_1 can stand out to determine which vehicle to connect. Certainly, the connection with the same vehicle is tried to be maintained than looking for a connection with another vehicle. However, the decision to use max operator is based on intuition, and the comparisons with the usage of other operations such as mean or weighted value are not presented in this paper.

3.1.3. Value. The value obtained by vehicle $j (v_j)$ for selecting strategy S_j^i is defined as the difference between the revenue obtained by forming a coalition (establishing a connection) with vehicle $i \, u_j (S_j^i)$ and the cost of the coalition with vehicle $i \, c_j (S_j^i)$ as given by

$$v_j\left(S_j^i\right) = u_j\left(S_j^i\right)\left(1 - c_j\left(S_j^i\right)\right). \quad (6)$$

Based on the above formulation, v_j will have the maximum value (equal with the revenue u_j) if the cost c_j is zero. On the contrary, v_j is minimum (zero) if the cost c_j is equal to 1, regardless of the revenue. In CGC, vehicle j selects a cooperation with vehicle i based on the value $v_j (S_j^i)$. Vehicle i is selected to establish V2V connection if vehicle i gives the best value to vehicle j. However, there is a constraint in selecting a vehicle to form V2V connection, i.e., the transmission range. Hence, the vehicles out of the transmission range cannot be selected due to the constraint. This can be done by setting the cost of connection with vehicles out of transmission range to the maximum. Although the connection with a vehicle, which is far away, has very a low SNR or even zero, setting the cost to maximum is aimed to assure that the vehicle will not be selected.

3.2. Cluster Formation. Once the vehicles select the paired vehicle to establish V2V connections, the clusters are formed. Since V2V connection establishment is arranged in a distributed manner, the formed clusters could be dynamic. The topology of clusters could be multihops and overlapping clusters are possible. Figure 2 shows the illustration of cluster formation using CGC. Two clusters are called overlap each other when there is at least one member of a cluster within the coverage of cluster head from the other cluster. For example, in Figure 2, the head of the cluster I is vehicle 5. Vehicles 7 and 8 belong to cluster II; however, they are within the coverage of vehicle 5. Thus, clusters I and II overlap each other. The overlapping clusters are possible since each vehicle is allowed to select another vehicle independently to establish a V2V connection. Although the vehicles aim to establish the connection with high a SNR, due to the coalitional game rule, the connection with the highest SNR is probably not selected. For example, vehicle 8 has five neighborhood vehicles as in Figure 2. Let us assume that V2V connection between vehicles 8 and 9 has the highest instantaneous SNR. However, due to the cost factors in the coalitional game such as connection lifetime and speed difference, vehicle 8 chooses vehicle 7 to establish V2V connection instead of vehicle 9. Thus, vehicles 7, 8, and 10 form their own cluster instead of joining cluster I. In the worst case, CGC allows a cluster to be formed although it has only two members like cluster III. It seems better if vehicle 12 connects to vehicle 13. However, due to the coalitional value, vehicle 12 forms its own cluster with vehicle 11.

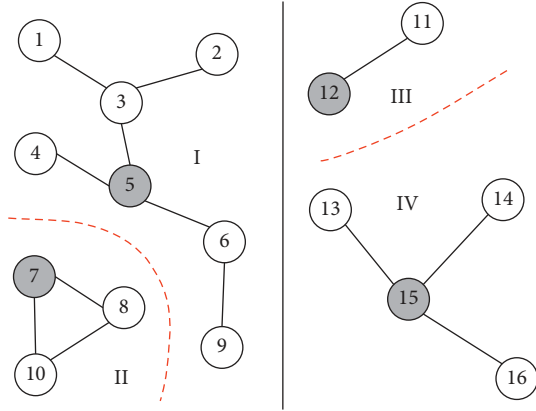One of the cost elements in CGC is the speed difference (Δv). This cost prevents two vehicles with high speed

FIGURE 2: Structure of clusters using CGC.

```
Input: E_j (vehicle electability), v_j (vehicle speed)
CH_found ⟵ 0
for j = 1, 2, . . . , N_j (number of cluster members)
    if j^{t-1} = CH & E_j ≥ 2 then
        j becomes CH
        CH_found ⟵ 1
        break
    end if
end for
if CH_found == 0 then
    for j = 1, 2, . . . , N_j
        CH_candidacy ⟵ 1
        for k = 1, 2, . . . , N_k; k ≠ j; N_k = N_j − 1
            if E_j < E_k then
                CH_candidacy ⟵ 0
            elseif E_j == E_k & v_j > v_k then
                CH_candidacy ⟵ 0
            end if
        end for
        if CH_candidacy == 1 then
            j becomes CH
            break
        end if
    end for
end if
```

ALGORITHM 1: Cluster head selection.

difference to establish a V2V connection. In the calculation of Δv, the speed of vehicles is vectorized. Thus, Δv can become higher if the two vehicles have a different direction. From Figure 2, it can be noticed that vehicles 11 to 16 form different clusters than vehicles 1 to 10. That is because vehicles 11 to 16 move in direction opposite to that of vehicles 1 to 10. However, cluster formation by vehicles with opposite direction is still possible, i.e., if the vehicles move very slowly so that the speed difference is below the threshold (Δv_{max}). It is possible in some situations such as in traffic jam, junction with the traffic light, or in the tollgate area.

3.3. Cluster Head Selection.

In CGC, each vehicle selects another nearby vehicle to establish a V2V connection. In this case, a vehicle can be selected by more than one vehicle. Therefore, in CGC, the electability of vehicles is used as one of the criteria to select the head of a cluster. The vehicle with the highest electability among the cluster members can become the cluster head (CH). Another criterion to select the cluster head is the speed of the vehicle. When there are two or more vehicles with the same electability, then the vehicle with the slower speed becomes the cluster head. Another alternative can be used, i.e., by selecting the vehicle with smaller speed difference to the cluster. However, it requires more computation and adds the transmission overhead. In spite of this, it is not a big problem since the main factor to select the cluster head is the electability, and the additional parameter is used only when there are two or more equal cluster head candidates. Due to the dynamic environment of VANET, the electability of a vehicle can also fluctuate. Therefore, a procedure is added in cluster head selection to increase the lifetime of a cluster head. A vehicle can maintain the state as cluster head as long as there are at least two vehicles connected (has electability at least 2), even though there is another vehicle in the cluster with higher electability. Algorithm 1 describes the process of cluster head selection in CGC.

4. Simulation

4.1. System Model

4.1.1. VANET Environment.
In this system model, the environment of VANET is in toll road with the longest route

has length 9 kilometers. The road is traced from the real map, i.e., city toll at Semarang, Indonesia. The longest route starts from (−6.954628, 110.451622) at the north to (−7.026183, 110.432788) at the south. In the middle of the main route around (−6.973131, 110.450006), there is a tollgate where vehicles from either north or south direction must pass through the gate, and thus, the flow of vehicles is slower around this point. The road has 3 lanes for each direction; meanwhile, the tollgate has 4 lanes for each direction. As shown in Figure 3, other than the main road (A to B), there are two points (C and D) where vehicles can enter or leave the road. Thus, there are 6 possible routes in this map: A-B, B-A, A-C, C-B, B-D, and D-A.

The traced coordinates of the roads from Google Maps [24] are in decimal degrees system. Therefore, a conversion is done for the convenience in the distance calculation. Coordinates in decimal degrees can be converted to Cartesian by multiplying with 111320. This multiplier is a particular value for coordinate at the equator (within 23°N/S).

4.1.2. Vehicle Mobility and Distribution.
The mobility of vehicles is simulated using an open source, microscopic and continuous road traffic simulator, namely, Simulation of Urban Mobility (SUMO) [25]. SUMO provides some conveniences in simulating the mobility of the vehicle by allowing the user to plot the road manually, to define the types and the flow of vehicles, and to record the result of simulations. The movement of vehicles is defined by Krauss' car-following model, which is influenced by the surrounding vehicles. The decisions of lane-changing and overtaking are
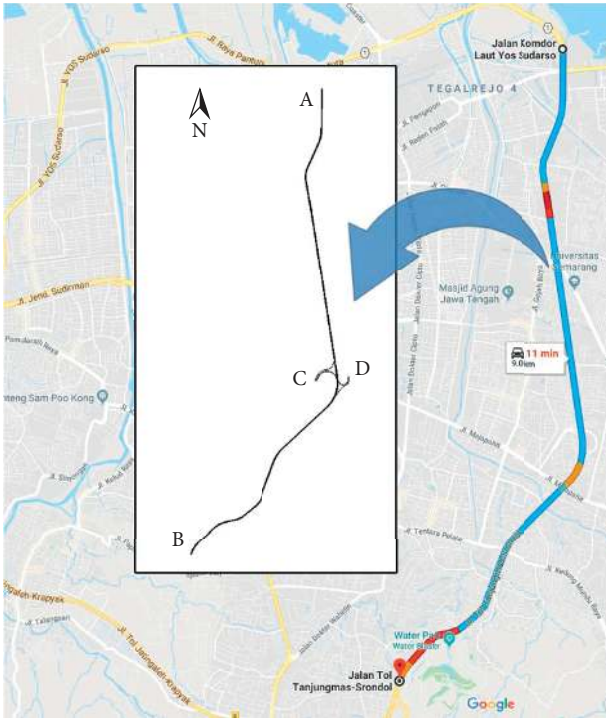
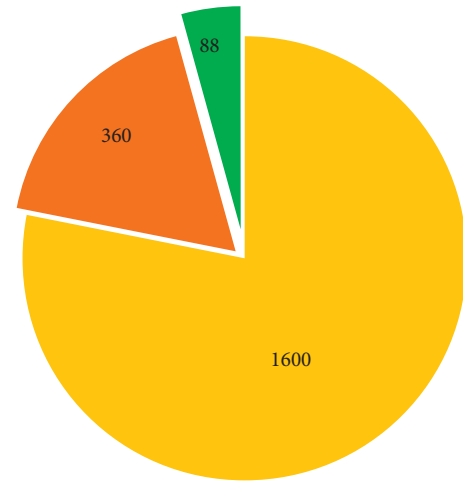FIGURE 3: Map of the road for VANET simulation.

TABLE 1: Vehicle speed attributes.

Vehicle	Maximum speed (m/s)	Speed factor	Speed deviation
Car	40	0.75	0.2
Coach	30	0.375	0.1
Trailer	20	0.5	0.3

Unit in vph (vehicles per hour)

■ Car
■ Coach
■ Trailer

FIGURE 4: Distribution of vehicle based on the type of vehicle.

defined by Krajzewicz's model [26]. In this VANET simulation, the results of SUMO are saved as an XML file consisting of information about vehicles position and speed along with the identity of vehicles. In this work, the simulation of vehicle mobility using SUMO uses the following settings. There are three types of vehicles: car, coach, and trailer. The properties of the vehicle are shown in Table 1. The distribution of vehicles based on the type and the route are presented in Figures 4 and 5, respectively.

Apart from the maximum speed and speed deviation, the flow of vehicle is also influenced by the lane speed limits. The speed factor as in Table 1 defines the expected multiplicator for the lane speed limits. In the road network, like the aforementioned, there is a tollgate around (−6.973131, 110.450006). Normally, the flow of vehicles is slower around the tollgate area. Therefore, a road edge with length 100 m centered at the tollgate coordinate is defined with lane speed limit 4 m/s. Other than this tollgate area, the road edges have lane speed limit 40 m/s. Figure 6 shows the screenshot of SUMO taken from the graphical user interface (GUI). From Figure 6, it can be noticed that the density of the vehicle is different between the area around the tollgate and other areas on the map. The density of the vehicle at the tollgate is higher, and the flow is slower as also shown by the red color on the road in Figure 3.

4.1.3. Signal Propagation Model. The propagation of the signal from the transmitter vehicle to receiver vehicle is defined by path loss and fading. Path loss defines the attenuation of the signal proportional to the distance between the transmitter and receiver. Fading meanwhile defines the
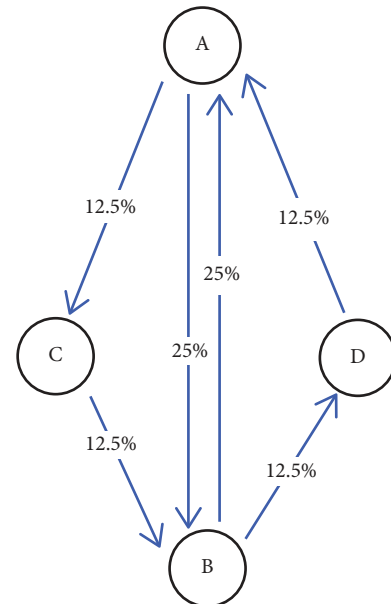

FIGURE 5: Distribution of vehicle based on the route.

variation either gain or attenuation of the signal due to the surrounding environment. Based on the path loss characterization for vehicular communication in [27], communication link gain (G) determined by path loss and fading in the highway environment is given by
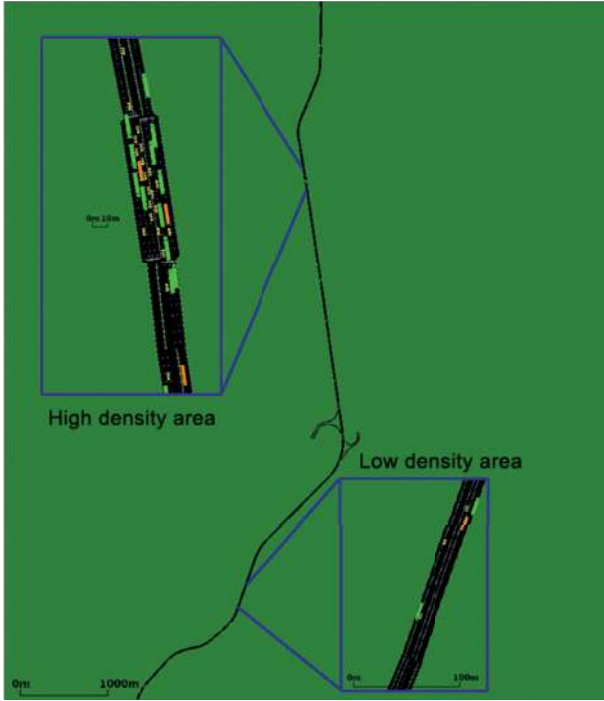
FIGURE 6: Screenshot of vehicle mobility simulation in SUMO.

$$G = -(37.92 + 16.4 \log_{10} d) + S_\sigma, \qquad (7)$$

where d represents the distance between the transmitter and receiver in meters. Fading is represented by S_σ, i.e., random variable, which has a normal distribution; zero mean; and standard deviation of 4.46. The link gain in (7) determines the strength of the signal received by the receiver vehicle as given by

$$S = Gp, \qquad (8)$$

where p is the transmission power level at the transmitter. In this work, all vehicles are assumed to have a fixed transmission power level, i.e., 20 dBm or equal with 100 mW.

In wireless communication, especially when the reused spectrum is used, the interference becomes a generic issue. However, this work has not dealt with the channel allocation. Therefore, the interference is not modeled specifically. For the sake of simplicity, the interference is represented by using additional noise with Gaussian distribution; mean 20 dB; standard deviation of 5 dB. In addition to interference, the communication channel has noise, namely, white noise that has equal intensity at different frequencies. The noises are summed and then used to calculate the SNR (Γ) of V2V connection. For the known SNR value in watts and the bandwidth of channel (B) in hertz, the capacity of the channel (C) is calculated based on Shannon–Hartley theorem as given by

$$C = B \log_2 (1 + \Gamma). \qquad (9)$$

4.1.4. Simulations Setup. The simulation of vehicle mobility and VANET clustering is performed separately. The vehicle mobility is simulated in SUMO, and the result of the

simulation is an extensible markup language (XML) file with information about vehicle ID, speed, coordinate position, and lane position at every second. Meanwhile, the simulations of VANET clustering are performed in MATLAB. A dedicated program is compiled to simulate the proposed CGC in MATLAB. As the input of simulation, the output file from SUMO is converted from XML to MATLAB data file (.mat). In addition to the vehicle position file, random numbers to represent fading (S_σ) are generated and saved as the input of simulation. This is to ensure that the same environment is used in VANET clustering simulation. Therefore, the simulations of clustering using different algorithms can be performed and the results can be fairly evaluated. Along with the clustering simulation using CGC, the simulations using lowest ID (LID) [7], density-based clustering (DBC) [8], and mobility-based dynamic clustering (MDC) [6] are performed for the purpose of comparison and evaluation. LID is the widely known clustering algorithm for mobile ad hoc network (MANET) as well as VANET. Meanwhile, DBC is the clustering method that has the most similarity with the proposed CGC in terms of selection metrics for clustering. Moreover, DBC is the last clustering method that uses SNR as one of the selection metrics according to a survey [28]. The number of clustering methods which use SNR as the selection metrics is very limited, thus one more comparison is done with a recent clustering method, namely, MDC. Although MDC does not use SNR as the metric to form the cluster, it has very good performance in terms of cluster stability. Therefore, MDC is also used for comparison in this research.

The simulation in SUMO runs for 800 seconds. However, the output file of simulation only records the information during $t_{sim} = 501$ to $t_{sim} = 710$. The early time of simulation is considered as warming up time, where the vehicles start to enter the road. Within $t_{sim} = 501$–710, some vehicles leave the road and some new vehicles enter the road. Thus, there are around 200 vehicles in total at 1 t_{sim}. The output file from SUMO is then used as the input of clustering simulations in MATLAB. Based on the input data, the duration of clustering simulation is 210 seconds. However, the results of the simulation are recorded for 200 seconds, i.e., from $t_{sim} = 11$–210. The reason is that the DBC algorithm needs some warming up times, i.e., the first 10 seconds of clustering simulation.

4.2. Performance Metrics. The performance of clustering methods are observed and evaluated using the following metrics:

(i) Number of clusters denotes the number of clusters formed at every unit of time during simulations. Usually, the fewer number of clusters is more desired since the fewer number of clusters is expected to increase the transmission efficiency such as reducing the overhead and communication between clusters.

(ii) Average size of the cluster represents the average number of members in a cluster. The higher number of cluster members is more desired since in the

worst case, a cluster can only have one or two members. However, sometimes the number of cluster members should be limited. For example, if the number of channels is limited, then the fewer number of members can reduce the channel congestion. Regardless of this opinion, it is assumed that the higher number of cluster members is better.

(iii) Cluster head change rate denotes the number of cluster head changes per second. Cluster head change can be in the existing clusters or the newly formed clusters. The fewer number of this metric is desirable since it implies a more stable cluster.

(iv) Clustering coverage defines the percentage of vehicles that join any clusters. Normally, 100% coverage is better. However, if the clustering method allows a single vehicle to form a cluster, then the coverage can be 100%.

(v) Average V2V SNR denotes the average of SNR from all V2V connections established at the time.

(vi) Average channel capacity denotes the average channel capacity based on the V2V SNR. Channel capacity denotes the maximum bit rate that the channel can support.

4.3. Results and Discussion. The results of simulations are organized into two subsections. The first subsection presents the results of simulations by alternating the value of parameters in CGC, namely, connection time threshold (δ_c) and maximum speed difference (Δv_{max}). Afterwards, it presents the results of the simulation using the proposed CGC and three other clustering methods, namely, LID, DBC, and MDC.

4.3.1. Impact of CGC Parameters Setting. Figures 7–12 display the results of clustering simulation by alternating δ_c (in seconds) and Δv_{max} (in meters per second). As shown in Figure 7, the number of clusters increases when δ_c is increased. It is because that if δ_c is increased, the vehicles become more difficult to change connection, and thus, they will form clusters with the vehicles with lower speed difference. It can be said that due to the increased δ_c, the vehicles tend to form more clusters but with the higher stability. On the contrary, the number of clusters decreases when Δv_{max} is increased. The Δv_{max} is analogous with the tolerance of speed difference. If the tolerance is low, the vehicles will be separated into more clusters. On the contrary, if the tolerance is high, the vehicles have more chance to unite into the same clusters. Hence, fewer clusters are formed. The average of cluster size as shown in Figure 8 is inversely proportional to the number of clusters. Hence, the higher the number of clusters, the smaller the average cluster size and vice versa. As mentioned above, if δ_c is increased, the vehicles tend to form more clusters but with the higher stability. One of the metrics to evaluate the stability of the cluster is the cluster head change rate. The more stable the cluster, the lower the change rate of the cluster head. In Figure 9, it can be noticed that if δ_c is increased, the cluster
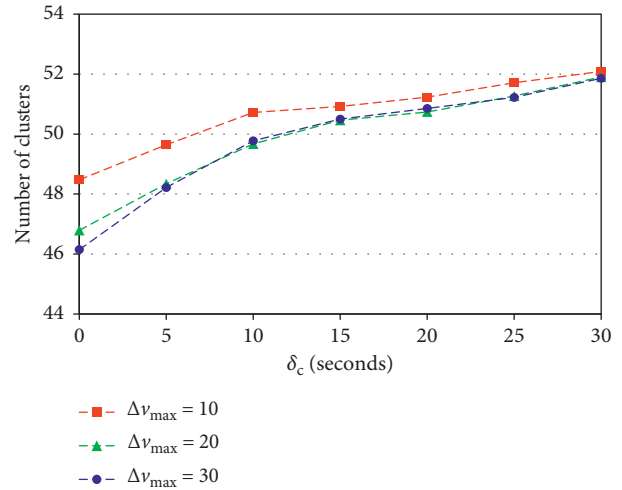

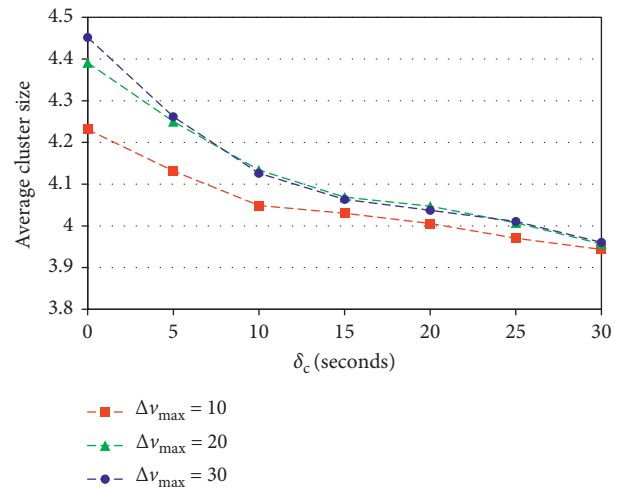FIGURE 7: Number of clusters based on the variation of δ_c and Δv_{max}.

FIGURE 8: Average cluster size based on the variation of δ_c and Δv_{max}.

head change rate decreases. The cluster head change rate also decreases if Δv_{max} is increased. This is because Δv_{max} is relevant to the cluster stability. The lower value of Δv_{max} forces the vehicles to maintain a connection with the vehicles that have nearly similar speed. The coverage of CGC is shown in Figure 10. It can be noticed that either Δv_{max} or δ_c do not have any significant impact toward the coverage of clustering. The coverage of CGC is mainly influenced by the transmission range of vehicles. As long as there are at least two vehicles within the transmission range of each other, those vehicles can form a cluster. However, some vehicles on the low-density area may not have another vehicle within their transmission range, and hence, they cannot form a cluster. Since those vehicles do not belong to any cluster, the coverage of clustering is not 100%.

Figure 11 shows the average of V2V SNR. When δ_c is increased, the average of V2V SNR decreases. This is because that when δ_c is lower, the chance that the vehicles change
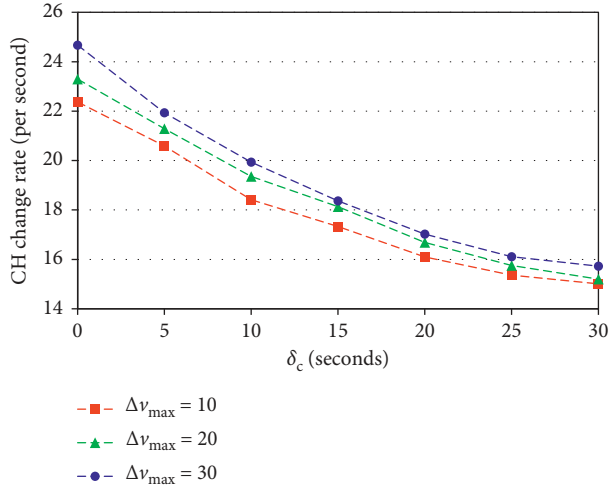
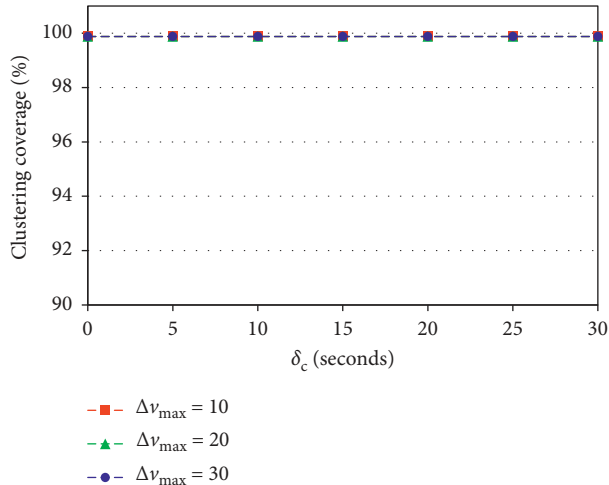FIGURE 9: Cluster head change rate based on the variation of δ_c and Δv_{max}.

FIGURE 11: Average V2V SNR based on the variation of δ_c and Δv_{max}.

FIGURE 10: Clustering coverage based on the variation of δ_c and Δv_{max}.
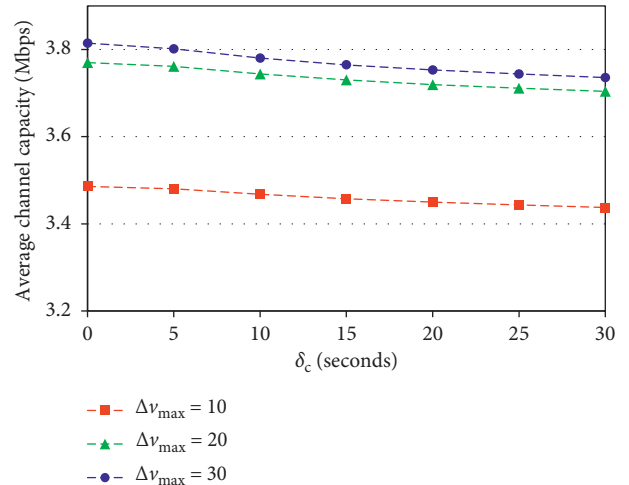
FIGURE 12: Average channel capacity based on the variation of δ_c and Δv_{max}.

connection to obtain a higher SNR is higher since the cost element of connection lifetime is lower. The lower cost at early connection lifetime enables the vehicles to change connection more frequently. On the contrary, the higher δ_c implies that connection change at early connection lifetime is higher, and hence, the vehicles choose to stick the connection with the current vehicle although the connection with other vehicles probably has the higher SNR. Meanwhile, the average of V2V SNR increases if the Δv_{max} is increased. This is due to the flexibility of vehicles to select V2V connection is higher. Since the vehicles have more options to establish a V2V connection, the vehicles have a higher chance to select a V2V connection with a higher SNR. The channel capacity is proportional with the SNR. Thus, the average channel capacity as shown in Figure 12 and the relation with the variation of δ_c and Δv_{max} can be equally described as in the explanation of average V2V SNR.
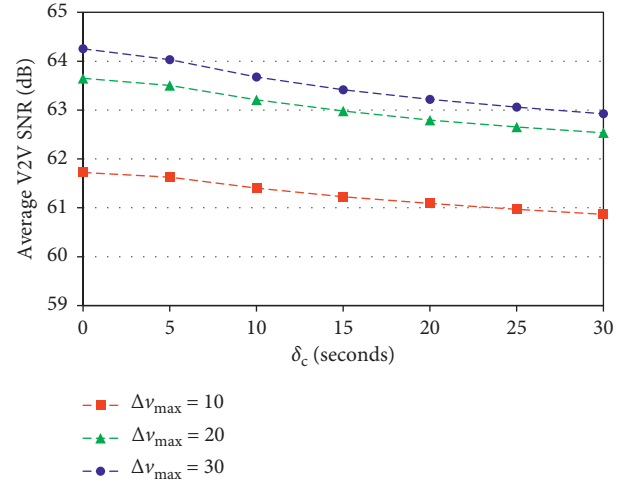
4.3.2. Comparison of Clustering Performance. The results of simulations using LID, DBC, MDC, and the proposed CGC are presented here. For the balance between connection link quality and cluster stability, the parameters of CGC are defined as follows: δ_c = 30 s and Δv_{max} = 20 m/s. Meanwhile, to obtain the higher link quality, DBC uses a minimum SNR 40 dB and minimum group membership lifetime (GML) 3 seconds to be the member of a cluster. For the simulation using the MDC method, the following parameter values are used. The clustering distance (D_t) or the maximum radius of cluster from the cluster head is 250 m. The beacon interval (BI) and merge interval (MI) are 0.5 s and 5 s, respectively. The maximum duration of temporary cluster head (T_{CHt}) and unclustered node (T_{UN}) are 5 s and 3 s, respectively.

Figures 13(a) and 13(b), respectively, display the average number of clusters and the average of cluster size from the simulation using LID, DBC, MDC, and CGC. It can be seen
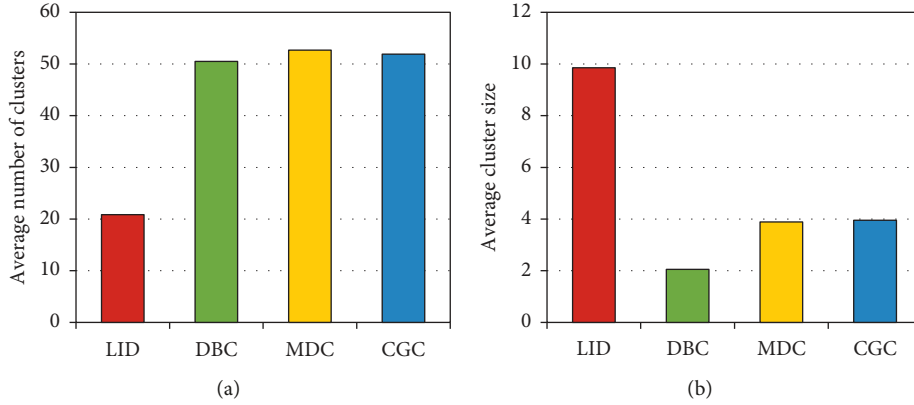
FIGURE 13: Cluster structure based on the number of clusters and cluster size. (a) Average number of clusters and (b) average size of cluster.

that LID has the lowest number of clusters and the biggest size of the cluster. This is because LID allows any vehicles to join a cluster, provided the vehicles are within the transmission range of the cluster head. Even the vehicles can join a cluster from the opposite direction of the road. Since the selection criteria are not strict, more vehicles can join a cluster, and hence, the size of the cluster becomes bigger. Therefore, only a few clusters are formed. Meanwhile, DBC, MDC, and CGC have more stringent criteria in forming a cluster. Thus, their average cluster size is smaller compared to LID. Moreover, DBC has the lowest average of cluster size due to the GML minimum requirement. To become a member of a cluster, a vehicle has to maintain link quality above the threshold until the minimum duration of GML is reached. Because of this, a vehicle cannot directly join a cluster and probably remain unclustered for some moments.

Figure 14(a) shows the cluster head change rate from simulations using the three methods. LID and MDC have lower cluster head change rate. This is because LID selects the cluster head based on the lowest ID of the vehicles. The cluster head will remain as long as no new vehicle with lower ID joins the cluster. This method surely has an advantage in toll road environment, where only a few possible routes exist. Moreover, almost no vehicle stops at the edge of the road. Therefore, the cluster head can remain the same for the longer duration. DBC has the highest cluster head change rate since the cluster head selection is based on weight, which is basically the average of link quality. However, this method is vulnerable, especially in a highly dynamic environment. Unfortunately, in toll road environment, the vehicles have a wide range of speed unlike in urban road. Therefore, the communication network is highly dynamic and the link quality can change frequently. CGC performs cluster head selection based on the electability. Each member of the cluster elects the candidate of the cluster head based on the highest value of coalition. In this case, the dynamic environment may also affect the performance of this method. However, in CGC, a vehicle is allowed to maintain the status as cluster head as long as it is elected by at least two vehicles. As the result, the cluster head change rate can be reduced. In Figure 14(a), it can be noted that MDC has the lowest cluster head change rate as expected, since MDC has a strong point in terms of cluster stability. The duration of

cluster head can be maintained for a longer duration in MDC, and hence, the cluster head change rate is very low. In MDC, a cluster head can retain the status as cluster head as long as there is at least one member, although there is another event that a cluster head must relinquish the status, i.e., during cluster merging.

The coverage of clustering is shown in Figure 14(b). LID has 100% coverage since the single vehicle can be counted as a cluster. Meanwhile, DBC and CGC do not count the single vehicle as a cluster. However, in CGC, the cluster can be formed by at least two vehicles located at their transmission range of each other. Therefore, CGC has great coverage since it only excludes the single vehicles. MDC also has high cluster coverage and it is comparable with LID and CGC. This is because in MDC, cluster formation with only two vehicles is also allowed. The coverage of DBC is the lowest. This is because DBC prevents a vehicle to join any clusters directly. As results, more vehicles are in unclustered status, and hence, the coverage of DBC decreases.

The average of V2V SNR and channel capacity are presented by Figures 15(a) and 15(b), respectively. DBC and CGC certainly have the better average of V2V SNR than LID since they consider SNR as one of the criteria in forming a cluster. MDC also has a higher V2V SNR average than LID although SNR is not included in criteria to form a cluster. This is because MDC limits the cluster size within a certain radius. Therefore, the higher value of SNR can still be obtained. However, CGC has the highest average of V2V SNR among the four methods. This is because CGC selects the best link quality to establish connection although indirectly through the concept of the coalitional game (revenue, cost, and value). Even the higher V2V SNR can still be reached by selecting the best connection. However, the stability of the connection cannot be maintained for a longer duration. In this case, coalitional game rule performs the work to balance between the higher link quality and the more stable connection. The capacity of the channel is proportional with the SNR according to (9). However, in Figure 15(b), the average channel capacity using DBC is the lowest among the three methods. This is due to the averaging, where the channel capacities from all vehicles are summed and then divided by the number of vehicles. Meanwhile, in DBC, there are some vehicles that are in an unclustered state. Since those vehicles
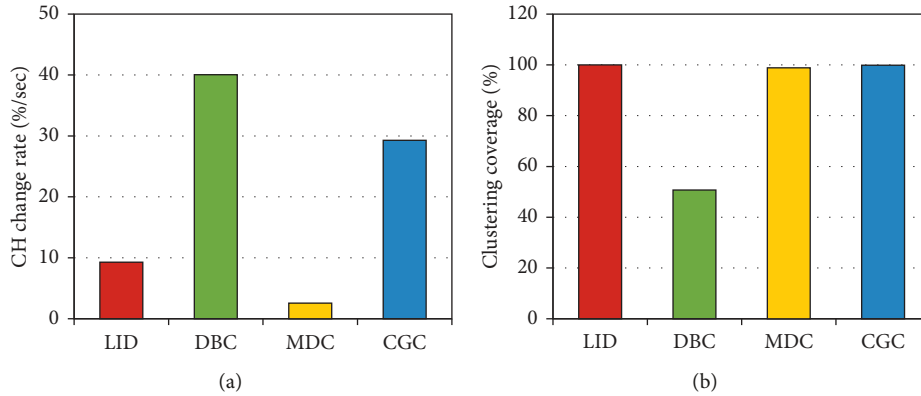
FIGURE 14: Cluster head change rate (a) and clustering coverage (b).
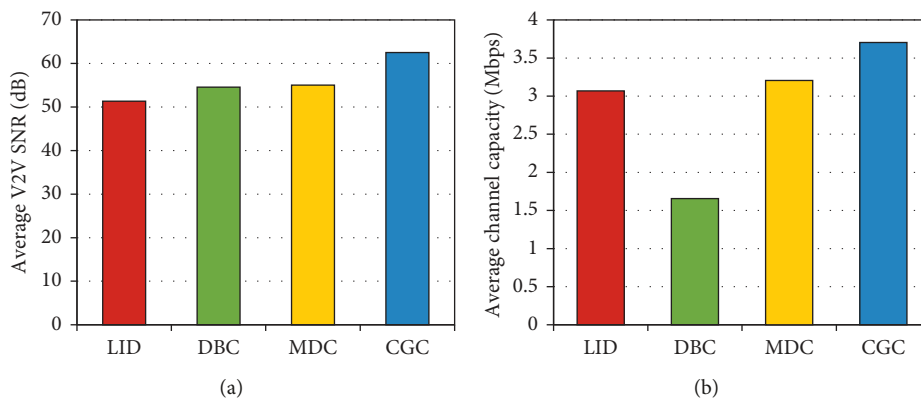

FIGURE 15: Link quality based on V2V SNR and channel capacity. (a) Average V2V SNR and (b) average channel capacity.

cannot establish a V2V connection, their channel capacity is counted as zero.

5. Conclusion

A distributed clustering method for VANET based on coalitional game theory, namely, CGC is proposed in this paper. Each vehicle attempts to form a cluster with other vehicles according to the concept of coalition value. Since the purpose of clustering is to improve the V2V SNR while maintaining the stability of the cluster, the coalition value is formulated based on this purpose. The value of coalition is defined by the revenue (V2V SNR) and the cost (connection lifetime and speed difference). In fast-changing network topology, the higher average of SNR can be obtained but the stability of the cluster becomes hard to be maintained. Based on the simulation results, SNR improvement can be adjusted in order to balance with the cluster stability by setting the parameters in CGC accordingly. Further simulation results show that CGC can obtain a higher average of V2V SNR and channel capacity than the other relevant methods.

Data Availability

This study is based on the datasets generated by the SUMO software upon vehicle mobility model we had constructed,

which are available from the corresponding author upon request.

Disclosure

There are no other persons who satisfied the criteria for authorship but are not listed. The authors further confirm that the order of authors listed in the manuscript has been approved by all of them.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

SS contributed mainly to the defining problem formulation, mobility model, and system modeling as well as discussion parts. SA contributed for implementing the simulation and graphical production while RA contributed for preparing the mobility models dataset using the SUMO software. The authors also confirm that the manuscript has been read and approved by all named authors.
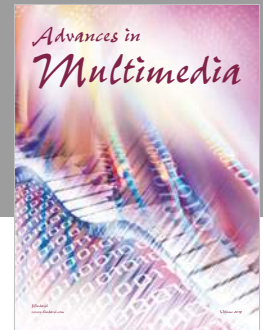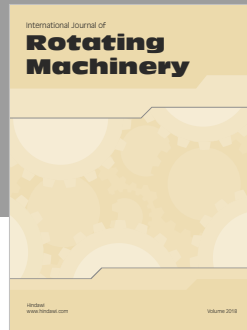
Acknowledgments

References

[1] H. Hartenstein and K. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 6, pp. 164–171, 2008.

[2] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. Tonguz, "Routing in sparse vehicular ad hoc wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 8, pp. 1538–1556, 2007.

[3] N. Wisitpongphan, O. K. Tonguz, J. S. Parikh, P. Mudalige, F. Bai, and V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks," *IEEE Wireless Communications*, vol. 14, no. 6, pp. 84–94, 2007.

[4] R. S. Bali, N. Kumar, and J. J. P. C. Rodrigues, "Clustering in vehicular ad hoc networks: taxonomy, challenges and solutions," *Vehicular Communications*, vol. 1, no. 3, pp. 134–152, 2014.

[5] W. Saad, Z. Han, M. Debbah, A. Hjorungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Processing Magazine*, vol. 26, no. 5, pp. 77–97, 2009.

[6] M. Ren, L. Khoukhi, H. Labiod, J. Zhang, and V. Vèque, "A mobility-based scheme for dynamic clustering in vehicular ad-hoc networks (VANETs)," *Vehicular Communications*, vol. 9, pp. 233–241, 2017.

[7] C. R. Lin and M. Gerla, "Adaptive clustering for mobile wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 7, pp. 1265–1275, 1997.

[8] S. Kuklinski and G. Wolny, "Density based clustering algorithm for VANETs," in *Proceedings of the 5th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops*, pp. 1–6, Washington, DC, USA, April 2009.

[9] W. Li, A. Tizghadam, and A. Leon-garcia, "Robust clustering for connected vehicles using local network criticality," in *Proceedings of the 2012 IEEE International Conference on Communications (ICC)*, pp. 7157–7161, Ottawa, Canada, June 2012.

[10] A. Ahizoune and A. Hafid, "A new stability based clustering algorithm (SBCA) for VANETs," in *Proceedings of the 37th Annual IEEE Conference on Local Computer Networks—Workshops (LCN Work)*, pp. 843–847, Clearwater Beach, FL, USA, October 2012.

[11] I. Tal and G. Muntean, "User-oriented fuzzy logic-based clustering scheme for vehicular ad-hoc networks," in *Proceedings of the 2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pp. 2–6, Dresden, Germany, June 2013.

[12] X. Duan, Y. Liu, and X. Wang, "SDN enabled 5G-VANET: adaptive vehicle clustering and beamformed transmission for aggregated traffic," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 120–127, 2017.

[13] B. Jinila and Komathy, "Rough set based fuzzy scheme for clustering and cluster head selection in VANET," *Elektronika Ir Elektrotechnika*, vol. 21, no. 1, pp. 54–59, 2015.

[14] E. Dror, C. Avin, and Z. Lotker, "Fast randomized algorithm for hierarchical clustering in vehicular ad-hoc networks," in *Proceedings of the 2011 10th IFIP Annual Mediterranean Ad Hoc Networking Workshop*, pp. 1–8, Sicily, Italy, June 2011.

[15] Y. Chen, M. Fang, S. Shi, W. Guo, and X. Zheng, "Distributed multi-hop clustering algorithm for VANETs based on neighborhood follow," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–12, 2015.

[16] S. Ucar, S. C. Ergen, and O. Ozkasap, "Multihop-cluster-based IEEE 802.11p and LTE hybrid architecture for VANET safety message dissemination," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2621–2636, 2016.

[17] X. Tang, H. Sun, L. Sun, C. Tan, and G. Xu, "Game theoretical approach for ad dissemination in cluster based VANETs," in *Proceedings of the 2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2013)*, pp. 1–6, Kunming, China, August 2013.

[18] S. Shivshankar and A. Jamalipour, "An evolutionary game theory-based approach to cooperation in VANETs under different network conditions," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 2015–2022, 2015.

[19] S. Sulistyo and S. Alam, "Distributed channel and power level selection in VANET based on SINR using game model," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 9, no. 3, pp. 432–438, 2017.

[20] T. Zhou, Y. Chen, and K. J. R. Liu, "Network formation games in cooperative MIMO interference systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 1140–1152, 2014.

[21] H.-Q. Lai, Y. Chen, and K. J. R. Liu, "Energy efficient cooperative communications using coalition formation games," *Computer Networks*, vol. 58, pp. 228–238, 2014.

[22] R. Massin, C. J. Le Martret, and P. Ciblat, "A coalition formation game for distributed node clustering in mobile ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3940–3952, 2017.

[23] C. Jiang, Y. Chen, and K. J. R. Liu, "Data-driven optimal throughput analysis for route selection in cognitive vehicular networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 11, pp. 2149–2162, 2014.

[24] G. Maps, Jl. Tol Tanjungmas-Srondol, https://www.google.com/maps/place/Jl.+Tol+Tanjungmas-Srondol,+Sawah+Besar,+Gayamsari,+Kota+Semarang,+Jawa+Tengah+50163/@-6.99753 16,110.4328032,13.5z/data=!4m5!3m4!1s0x2e70f3355f88d1430x 1c3760724e0a9890!8m2!3d-6.9730197!4d110.4500701.

[25] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO—simulation of urban mobility," *International Journal on Advances in Systems and Measurements*, vol. 5, no. 3, pp. 128–138, 2012.

[26] D. Krajzewicz, M. Bonert, and P. Wagner, "The open source traffic simulation package SUMO traffic management projects," in *Proceedings of the Robotics and Automation*, 371–380, Orlando, FL, USA, May 2006, https://elib.dlr.de/46740/1/RoboCup2006_dkrajzew_SUMO.pdf.

[27] H. Fernandez, L. Rubio, V. M. Rodrigo-Penarrocha, and J. Reig, "Path loss characterization for vehicular communications at 700 MHz and 5.9 GHz under LOS and NLOS conditions," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 931–934, 2014.

[28] C. Cooper, D. Franklin, M. Ros, F. Safaei, and M. Abolhasan, "A comparative survey of VANET clustering techniques," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 657–681, 2017.