A Simple and Robust Clustering Scheme for Large-Scale and Dynamic VANETs

MUSTAFA BANIKHALAF [ID]1 AND MOAIAD AHMAD KHDER2, (Senior Member, IEEE)
1Department of Computer Science, Yarmouk University, Irbid 21163, Jordan
2Department of Computer Science, Applied Science University, East Eker 3201, Bahrain

Corresponding author: Mustafa Banikhalaf (mbanikhalaf@yu.edu.jo)

ABSTRACT Clustering is a promising technique to manage network resources efficiently and, in vehicular communications it is used to group vehicles with similar characteristics managed by a selected vehicle called a Cluster Head (CH). Due to the highly dynamic topology in vehicular networks, a CH selection process becomes a challenging task. Thus, this paper presents a new clustering scheme, namely, Efficient Cluster Head Selection (ECHS) scheme to select the most suitable CHs. The proposed ECHS scheme introduces important conditions pertaining to the methods deployed in constructing clusters before starting the CH selection. For instance, based on the ECHS rules the ideal CH is the one that centralizes the cluster. This is because it will remain connected as long as possible with its neighbors. The ECHS scheme also guarantees proper clustering distribution in the network, so that the distance between two consecutive clusters are adjusted carefully. Such conditions are guaranteed to effectively cluster vehicles in the road and make the ECHS scheme works better than its counterpart. Simulation experiments are conducted to examine the performance of the ECHS and the results demonstrate that the ECHS scheme achieves the design objectives in terms of CH lifetime, Cluster Member Lifetime (CML), Packet Loss Ratio(PLR), Overhead for Clustering(OC), Average Packet Delay (APD), and Cluster Number (CN).

INDEX TERMS Clustering, cluster head, cluster gateway candidate, vehicular ad hoc networks.

I. INTRODUCTION

Nowadays, shortcomings of the traditional transportation systems are eliminated significantly by employing intelligent Vehicular Ad hoc Networks (VANETs). Due to the rapid development of wireless sensors and Internet of Vehicles (IoVs) [1], VANETs can be integrated with other technologies such as Cloud and Fog computing [2]–[4]. This sort of integration makes VANETs easy to deploy and include more traffic management applications. Basically, communication in VANETs can be divided into two categories, depending on the types of the running applications and requested services. The first one, is Vehicle-to-Vehicle (V2V) communication, which is commonly used when vehicles share local traffic information among each other without using infrastructure configuration [5]–[7]. The second one is Vehicle-to-Everything (V2X) communication which combines any type of communication between vehicles and infrastructure nodes such as a Road-Side Unit (RSU), a Fog Node,

a Base Station, and a Cloud Center. (V2X) communication usually helps vehicle to collect information about different zones inside a city to deduct traffic congestions and to discover free-congested roads [8], [9]. It is also used for controlling and managing virtual and physical traffic light systems [10], [11].

In contrast to V2X communication, end-to-end channel connectivity among vehicles in V2V communication cannot be always guaranteed. This is because vehicles are moving at high speeds and different directions, which leads to a frequent network partitions problem. In fact, discontinued connection results in poor network performance, packet dropped and further affects exchanging traffic safety information. Since V2V model is categorized as a decentralized self-organizing network, bandwidth and channel contention are not managed by a centralized vehicle. Because a single vehicle has a limited transmission coverage, it cannot maintain a global knowledge for a large-scale and dynamic network as VANETs. Thus, clustering technique has been proposed in which V2V communication can be managed by selected vehicles based on certain criteria [12]–[17].

The associate editor coordinating the review of this manuscript and approving it for publication was Yan Huo [ID].

The basic principle of clustering technique is to divide vehicles into virtual groups. Each group represents a cluster and contains vehicles having similar characteristics such as vehicles density, velocity and geographical locations [18]. Before creating the cluster, vehicles that relatively move in the same directions and speeds agree on selecting one CH. It serves as a local coordinator between its members, collects and disseminates traffic information. Communication between clusters is established via Cluster GateWay (CGW) vehicles that belong to two clusters at the same time. Generally, the established cluster must make sure its running for long time and reducing network overhead [19].

In this paper, we present robust clustering scheme which is called an Efficient Cluster Head Selection (ECHS) scheme, to construct reliable and stable clusters in VANETs. The ECHS scheme applies the following rules to meet its proposed requirements:

- The CH is preferred to be positioned at the middle of the cluster to avoid cluster gap problem.
- Each CH should determine its Cluster Gateway Candidates (CGCs), for the sake of proper selection of the next CH, and for adaptive cluster distribution among vehicles.
- Reducing coverage overlap between adjacent clusters to eliminate unnecessary CGWs.
- Eliminating number of redundant retransmissions from CGWs that have the same additional coverage area.

The rest of paper is organized into following sections. Section 2 discusses related work. Section 3 presents the system model of the proposed scheme. Section 4 describes the simulation environment and shows the obtained results. Section 5 concludes the paper with possible directions for the future work.

II. RELATED WORK

In VANETs, unpredictable vehicular mobility either in terms of a variety of vehicle speeds or vehicle movements in different directions, imposes extra constraints in designing cluster formation schemes. Therefore, several clustering techniques based on mobility features have been proposed with the aim of mitigating challenges posed by creating clusters. Since most of them adopted similar techniques, we shed the light on the most important ones, and interested readers can refer to other works [20]–[23].

A new dynamic mobility-based clustering scheme is proposed in [24] for urban environment. The scheme builds clusters in VANETs based on some several important metrices such as vehicles' relative velocity, and link lifetime estimation. In [25], a cluster protocol namely Clustering Formation for Inter-Vehicle Communication (CF-IVC), is used to cluster group of vehicles based on different speed intervals, where each vehicle joins a cluster of similar velocity. The Affinity PROpagation for VEhiclar networks (APROVE) scheme is proposed in [26], which it employs a similarity function to create stable clusters. The similarity function is defined as a combination of vehicular position (current and

future positions) and mobility. Nevertheless, APROVE scheme is a sort of a distance-based clustering algorithms, which often suffers from frequent re-clustering when the speed of vehicles changes dramatically. Moreover, APROVE scheme incurs an extra delay time for cluster formation process due to need several iterative loops. A lane-based clustering algorithm is presented in [27] for a CH selection based on the majority traffic flow and mobility information of vehicles. Vehicles that will turn to the left or to the right line are not allowed to become CHs, but only vehicles that continue driving in the same direction. Another clustering based direction approach is presented in [28], where CHs selection process takes place based on vehicles travelling directions. Vehicles that move in opposite direction are not allowed to be elected as CHs due the short communication period between the CH and its CMs. This assists to reduce the cluster reconfiguration cost. Similarly a novel algorithm to form stable clusters for a highway environment was presented in [13]. Clusters are formed by vehicles traveling in the same direction and with the same speed level. Hence, vehicles are classified their neighbors into stable neighbors and non-stable neighbors based on their relative speed. The authors in [29] presented a concept of having two CHs inside the same cluster. The primary cluster head (PCH) and a secondary cluster head (SeCH). The PCH has the highest weight value that is aggregated from different metrices such as mean speed and distance of a vehicle to its neighbors. The SeCH acts as a backup to the PCH to improve the cluster stability and takes its responsibility when leaving the cluster.

The Double Head Clustering (DHC) method for VANETs has been introduced in [30], which used new metrices for a CH selection to increase the cluster stability and efficiency. Besides speed of the vehicle, direction, and position, it considers the link quality and the link expiration time (LET).

Furthermore, a considerable number of important clustering techniques has been proposed with the aim of efficient routing and data dissemination. In [31], after selecting CHs based on the direction and distance metrices, it is used for route discovery and to deliver data packets. The authors in [32] proposed a Novel Real Time Vehicular Communication (RTVC) scheme for VANET. Clusters are developed between vehicles based upon average speed and direction, and multicast routing protocol is presented to route data packet from a source to a destination vehicle in two phases. In the first phase, the CH is established the route from the source vehicle, while in the second phase the CH is formed the route to destination via intermediate CHs. Cluster-based On-demand DElay tolerant routing (CODE) algorithm for a highway scenario is proposed in [33]. CODE elects CHs on the basis of vehicles direction and relative speeds to establish route discovery between pairs of source and destination. Similarly, in [34] a cluster is formed based on vehicles direction and location to disseminate data adaptively and to optimize network bandwidth. A clustering scheme based on the driving directions of vehicles is presented in [35]. In the formed cluster, each CM is employed a non-deterministic approach

based on the number of received packet to disseminate data between vehicles, and the CH forwards the received packet toward the transmission direction.

Benefit of the Clustering technique is also extended to efficient dissemination safety messages in VANETs [36]. A Novel Segment based Safety message broadcasting in Cluster (NSSC) has proposed with three main functions, namely, Cluster Formation, Collision Avoidance and Safety Message Broadcasting. Variant based Clustering (VbC) scheme based in Chaotic Crow Search (CCS) algorithm is used in NSSC to elect a CH based on the two different metrics that are mobility and connectivity metrics.

Obviously, most of the suggested schemes in the related work handle clustering problem in VANETs by exploiting vehicles mobility characteristics. However, some critical factors in creating clusters have been neglected and have not been explicitly defined and modeled, such as clusters distribution pattern among vehicles. In addition, CHs have been selected by applying simple calculation comparison between vehicles in terms of speed, direction and distance without considering a CH location inside its cluster. Furthermore, in the above related works a CGW is nominated only if it is related to two clusters concurrently, but reducing duplicate retransmissions from neighboring CGWs have not been addressed clearly. Therefore, to fill these gaps this paper presents a new ECHS scheme with the aim of finding an optimized solution for clustering technique in VANETs.

III. SYSTEM MODEL
A. MODELING OF EFFICIENT CLUSTERS
The VANET topology can be modeled as undirected graph G (V, E). V is defined as a set of vertices representing the vehicles in the network, and E is the set of edges representing the communication links among vertices (i.e. vehicles). There is a direct communication link $(i, j) \in E$, if and only if vehicles i and j are in each other's transmission range:

$$\exists (i, j) \in E \rightarrow dis(i, j) \leq min(Tr_i, Tr_j) \quad (1)$$

where Tr_i, Tr_j represents the maximum transmission range for vehicle i and j, respectively. dis(i,j) is the distance between vehicles i and j. Suppose that N_i is a set of one-hop neighbors for vehicle i then, N_i can be represented as follows:

$$N_i = \{j \in V | \exists (i, j) \in E\} \quad (2)$$

The one-hop neighbors of vehicle i, is defined as:

$$N_{if}(v_i) = \{v_j | v_j \in V, dis(i, j) \leq min(Tr_i, Tr_j)\} \quad (3)$$
$$N_{ib}(v_i) = \{v_j | v_j \in V, dis(i, j) \leq min(Tr_i, Tr_j)\} \quad (4)$$

$N_{if}(v_i)$ and $N_{ib}(v_i)$ represent the cardinality of the 1-hope neighborhood set of vehicles i from forward and backward direction, respectively. A vehicle can determine a neighborhood direction by calculating the angel degree from received messages. The coordination of current position and previous of each vehicle is defined as (X_c, Y_c) and (X_p, Y_p), respectively.
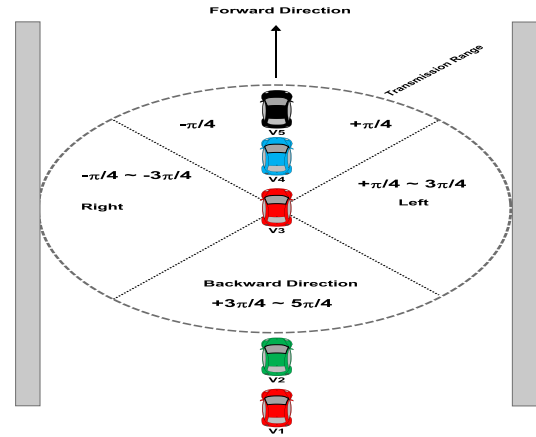

FIGURE 1. Relative backward and forward directions.

Then, the direction angle θ is calculated as follows:

$$\theta = \tan^{-1} \frac{Y_p - Y_c}{X_c - X_p} \quad (5)$$

Then, forward direction is confined to the closed interval:

$$\left[\theta - \frac{\pi}{4}, \theta + \frac{\pi}{4}\right] = \left\{\theta_f | \theta - \frac{\pi}{4} \leq \theta_f \leq \theta + \frac{\pi}{4}\right\} \quad (6)$$

and similarly, backward direction:

$$\left[(\theta + \pi) - \frac{\pi}{4}, (\theta + \pi) + \frac{\pi}{4}\right]$$
$$= \left\{\theta_b | (\theta + \pi) - \frac{\pi}{4} \leq \theta_b \leq (\theta + \pi) + \frac{\pi}{4}\right\} \quad (7)$$

Fig.1 shows the relative backward and forward directions for vehicle v3. Cluster head should be selected carefully to reduce selection overhead, and to prolong its lifetime. In addition to speed, direction and density factors which are generally used in election of a cluster head, the following factor must be considered. It is crucial for each vehicle to know if it's both directions are full, empty or one direction is semi-full of neighbors, so it can be decided if it is a potential CH or not. This is due to high likely to leave its cluster or to join another one quickly. A vehicle that is located at a tail or at a front of a cluster and does not have one-hope neighbor communication from front or backward directions, should be excluded from being a CH. From (3) and (4), we can know if a vehicle is semi-full of neighbors if:

$$N_{if}(v_i) = 0 \quad or \quad N_{ib}(v_i) = 0 \quad (8)$$

But a vehicle is considered full of neighbors if both of its directions are:

$$N_{if}(v_i) \quad and \quad N_{ib}(v_i) > 0 \quad (9)$$

and if both directions are empty:

$$N_{if}(v_i) = 0 \quad and \quad N_{ib}(v_i) = 0 \quad (10)$$

A CH is always preferably to be located approximately at a center of the cluster. To achieve this requirement a CH should meet the following condition:
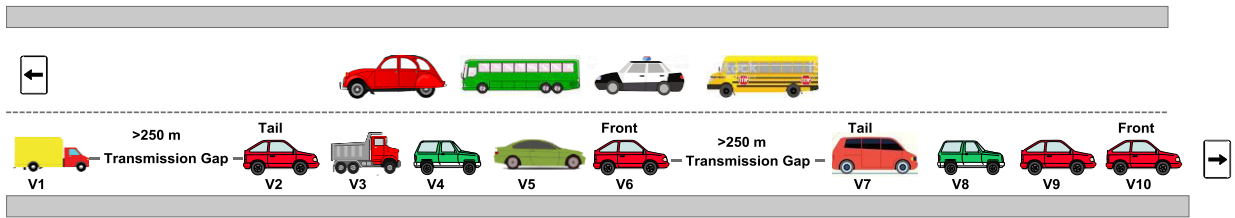
$$N_{if}(v_i) \cong N_{ib}(v_i) \quad (11)$$

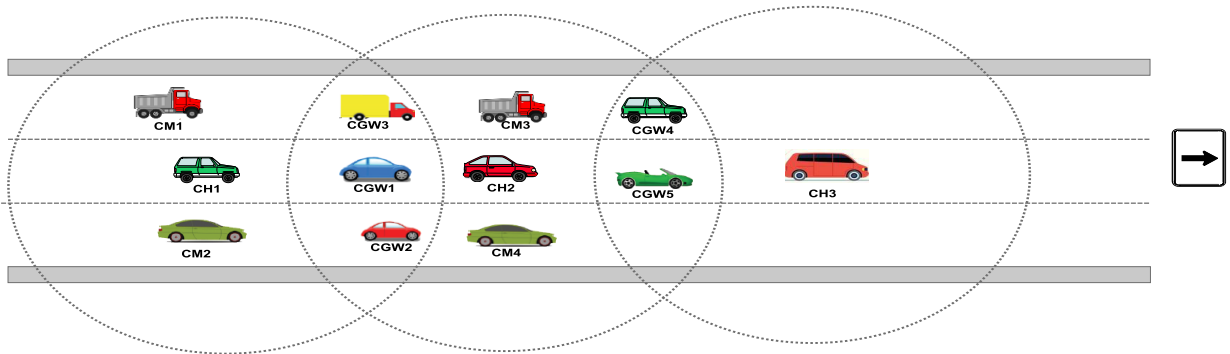**FIGURE 2.** A scenario where tail and front of the cluster should not be CHs.

FIGURE 3. Illustration of a scenario where unnecessary CHs should not be formed.

This means if a vehicle has an equal or an approximately similar number of neighbors in both directions, it is more likely to become a CH. The vehicle that is located at a center of the cluster is always the best one to manage it. It is high probably to stay with its cluster for long time because its speed and movement are related to vehicles that surround it.

As show in Fig.2, v2 is located at a tail of the cluster and does not have neighbor's knowledge from behind, since v1 is located beyond its transmission range of 250m. Similarly, v6 is located at a front of the cluster and maintain neighboring information from only behind neighbors (i.e. v5, v4, v3, v2). However, v6 cannot receive neighborhood information from front vehicles due to transmission gab between it and v7. Hence, v2 could slow down its speed and join another cluster, and v6 could speed up and join the next cluster. This will cause frequent network fragmentations and repeat a CH election producer several times. To handle this problem, tail and front vehicles should not be elected as CHs. In this case, each vehicle is responsible to monitor the direction of its neighbors.

If a vehicle does not receive any hello message from at least one direction, this indicates the vehicle is located at a tail or at a front of the cluster with a transmission gap. Then, this vehicle will be unable to send a Cluster Head Announcement (CHA) to its neighbors. As shown also in Fig.1, V3 cannot send CHA as it receives hello message from only forward direction v4 and v5, since v1 and v2 are located outside of V3's transmission range, and it cannot receive any information from backward direction. Hence, V3 has one direction semi-full neighbor, and does not meet

CH requirements. When the same rules are applied on the scenario of Fig.2, we can conclude that v4 and v9 are privilege to become CHs as their both directions are full of neighbors, while v2, v6 and v7 are not.

B. CLUSTER GATEWAY SELECTION

In some cases, as shown in Fig.3 a few CMs of a given cluster have already declared themselves as CGWs for two CHs. Rest of CMs can be reached from both clusters' CGWs. For instance, assume that there are three CHs; CH1, CH2 and CH3. CH1 has CGW1, CGW2 and CGW3, and CH3 has CGW4 and CGW5, which are already CMs of CH2. When CH1 disseminates a message, its CGWs (i.e. CGW1, CGW2 and CGW3) are responsible to deliver this message to CH2. Any transmission from CH2 must be suppressed, since these CGWs can cover all CMs of CH2. Hence, cluster 2 which managed by CH2 does not serve any benefit and should not be created initially. The following methodologies are implemented in this paper to prevent the unwanted cluster problem such as CH in Fig.2.

It is clearly appeared when CMs of a given CH can be covered by CGWs of surrounding clusters. Once a CH is anticipated it immediately selects its Cluster Gateway Candidates (CGC) based on the following conditions. CGCs must be located at the CH's boarder as it the most potential candidate of being a future CGW. Each CM keeps a record for a distance between itself and its CH and shares it with surrounding neighbors. If a CM does not receive a distance value larger than its to his CH, it announces himself as a CGC. Before any vehicle sending CHA, it collects
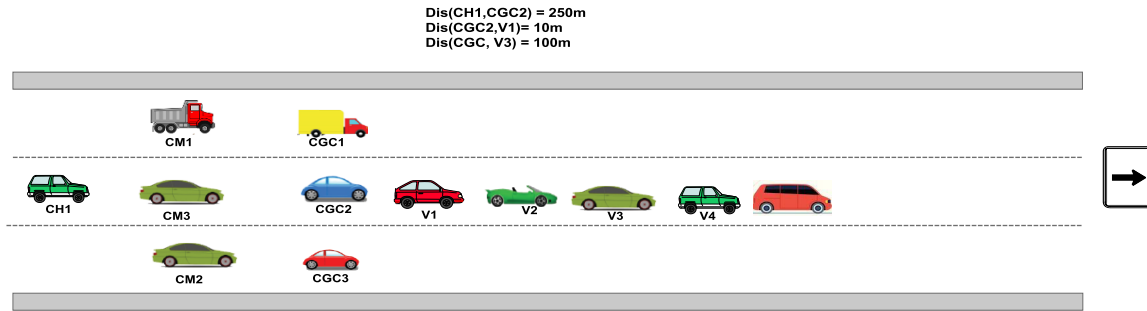
Dis(CH1,CGC2) = 250m
Dis(CGC2,V1)= 10m
Dis(CGC, V3) = 100m

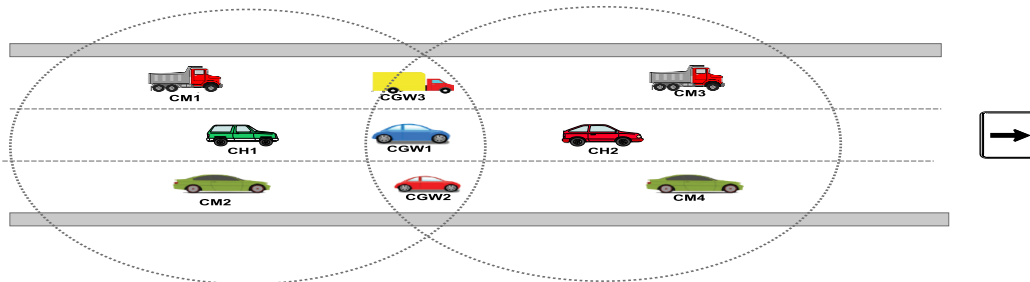**FIGURE 4.** A scenario shows how to determine CGC and the next CH.

FIGURE 5. Duplicated retransmission from CGW1, CGW2 and CGW3.

neighbor's information and starts a *backoff* timer based on the following equation:

$$backoff = T_{max} \left(1 - \frac{D_{CGC,vi}}{2D_{max}} \right) T_{random} \qquad (12)$$

where Time_{max} is defined as the maximum differed time, $D_{CGC,v}$ indicates the difference distance between a vehicle v_i and CGCs, and D_{max} is the maximum allowed transmission range. T_{random} is added to avoid a contention between two vehicles close to each other. If a vehicle very close to any CGC, it will initiate a long *backoff* timer which decrease its opportunity to become a CH. Applying this methodology can minimize number of unwanted clusters and optimize a network performance. In Fig.4, suppose that CH1 is announced himself as a cluster head, and CGC1, CGC2 and CGC3 are anticipated as Cluster Gateway Candidate, since they are located at CH1's border. CGCs broadcast their statues to surrounding neighbors including v1, v2, v3 and v4.

Each vehicle initiates a *backoff* timer based on (12), before sending CHA to its neighbors. Vehicles close to CGCs such as v1 and v2 wait long time before sending CHA and this reduce their opportunity of becoming CHs. On the other hand, *backoff* timer for vehicles furthest from CGCs such as v3 and v4 is set to short, and they are highly possible to be elected as CHs. v1 is not also appropriate CH candidate because it is approximately covering the same area that CH1 has already covered. Furthermore, several CGWs will be created due to increasing number of CMs that will hear CH declaration from CH1 and v1 if it becomes CH2. In this scenario to guarantee efficient distribution and selection of CHs among vehicles,

rules described in (11) and (12), are employed. Hence, v3 is considered as the best CH because it is a center vehicle, and it keeps a reasonable distance to CH1.

C. REDUCTION NUMBER OF UNNECESSARY RETRANSMISSON

CGWs in most cases cannot be a single vehicle, since several vehicles inside the same cluster can be nominated as CGWs. As shown in Fig.5, three vehicles are nominated as CGWs; CGW1, CGW2 and CGW3 because they belong to two the CH1 and CH2 at the same time. When the CH1 or the CH2 propagates a message only CGW1, CGW2 and CGW3 are privilege to forward it to another CH. It is clear that one broadcast from CGW1, CGW2 or CGW3 can cover the target area. However, all CGWs will perform the same task with no additional coverage area. To handle this problem each CGW initiates a *backoff* timer relative to its distance from its CH as the following equation:

$$backoff = \text{Time}_{max} \left(\frac{D_{CGW,CH}}{D_{max}} \right) + T_{random} \qquad (13)$$

where Time_{max} is defined as the maximum differed time, $D_{CGW,CH}$ indicates the difference distance between a CGW and its CH, and D_{max} is the maximum allowed transmission range. T_{random} is added to avoid a contention between two vehicles close to each other. CGW_1 forwards the received message from the CH_1 as it is the furthest one and its timer will be expired first. Once CGW_2 and CGW_3 hear the same message from CGW_1 during the *backoff* time, they cancel their retransmission immediately. If the *backoff* timer

is expired and CGW_2 and CGW_3 do not receive the same message from CGW_1, they take place and perform forwarding the message.

D. CLUSTERING VEHILCES BASED ON SPEED LIMIT

In RTVC scheme [32], vehicles are grouped to a cluster based on a Cluster Speed Limit (CSL) as follows:

- Clusters of slow lane (CS): CSLS < 60 km/h.
- Clusters of medium lane (CM): $60 \leq$ CSLM < 90 km/h.
- Clusters of fast lane (CF): $90 \leq$ CSLF $\leq$ 120 km/h.

where CSLs, CSLm, and CSLf are related to clusters of slow, medium, and fast lane vehicles, respectively. Every two minutes each vehicle computes the average speed (V_{avgspd}) parameter, to decide whether it can join or leave a cluster. Let C_f is the cluster of fast speed lane, C_m is the cluster of medium speed lane and C_s is the cluster of slow lane. Then, Clusters construction model can be defined based on CSL as follows:

$$C_f = \begin{cases} f_1 \in [\alpha f_1 | \theta_1] \\ f_2 \in [\alpha f_2 | \theta_2] \\ \cdot \\ \cdot \\ \cdot \\ f_n \in [\alpha f_n | \theta_n] \end{cases} \quad (14)$$

$$C_m = \begin{cases} m_1 \in [\alpha m_1 | \theta_1] \\ m_2 \in [\alpha m_2 | \theta_2] \\ \cdot \\ \cdot \\ \cdot \\ m_n \in [\alpha m_n | \theta_n] \end{cases} \quad (15)$$

$$C_s = \begin{cases} s_1 \in [\alpha s_1 | \theta_1] \\ s_2 \in [\alpha s_2 | \theta_2] \\ \cdot \\ \cdot \\ \cdot \\ s_n \in [\alpha s_n | \theta_n] \end{cases} \quad (16)$$

where (f_1 to f_n), (m_1 to m_n) and (s_1 to s_n) represents cluster members of C_f, C_m, and C_s with a direction of θ, respectively.

E. CLUSTER FORMATION PROCEDUER

In the beginning all vehicles are declared themselves as Norma Vehicles (NVs). Only NVs which meet aforementioned cluster head selection requirements can send CHA to their surrounding neighbors which summarized as follows:

- A NV before sending CHA should check if it verifies condition of (11), to make sure it is centered at the cluster.
- If the NV deducts itself as a neighbor to CGC, it should start *backoff* timer equal to the one described in (12).
- Each NV has neighbor list of all vehicles with V_{avgspd} and CTV values. Only NV that its average speed similar or equal to CTV $\pm$ u value can send CHA [32].

Once a NV achieves above requirements it sends out CHA and announces itself as a CH_i, if it has not yet received any CHA from other vehicles during CH_{timer}. Any NV receives

CHA from CH_i, tries to join the cluster by sending Cluster Member Request (CMR) to it. If CH_i accepts its request, the NV becomes a CM, and it should check if it is a potential CGC. CM or CGC is considered left the cluster if it does not receive any information from its CH with an interval time. Then, it should change its status again to a NV. Algorithm 1,2 and 3 describe CH election, cluster formation and CM and CGC formation steps, respectively.

Scheme 1 CH Election

NV v_i collects its neighborhood information.
Set CH_Flag = True
If $N_{if}(v_i) \cong N_{ib}(v_i)$ ***then***
   If v_i *is_neighbour of CGC* ***then***
      Set backoff timer for v_i as in equation (12)
     Else
        Set random backoff timer for v_i
    End IF
End_IF
While *backoff timer for* v_i *>0* ***do***
   If v_i *receives the CHA from another* CH_j ***then***
      CH_Flag = False
      Go to scheme 2
End_IF
End While
If CH_Flag = True ***then***
v_i is elected as the CH_i
CH_i broadcasts the CHA
End_IF

Scheme 2 Cluster Formation

Set CH_Flag = 1
If NV v_i receives the CHA from CH_j ***then***
   v_i sends CMR message to CH_j
End_IF
 If CH_j receives the CMR from NV v_i ***then***
   Check if the v_i from C_s, C_m or C_f
   If v_i belongs to CH_j ***then***
      Send acceptance confirmation to v_i
    Else
      Discard CMR message from v_i
   End IF
End_IF

IV. PERFORMANCE EVALUATION

In this paper, the proposed scheme (i.e. the ECHS is compared with three well-known previously clustering scheme adopting the same technique in clustering vehicles, which are the RTVC scheme [32], the CF-IVC [25] and the APROVE scheme [26]. The experimental results are generated by using Network Simulator 3 (NS3) version 3.21 [37]. Simulator of Urban Mobility (SUMO) which is known as a micro-traffic simulator is used to generate realistic mobility traces of vehicles at different densities [38]. For communication between

Scheme 3 CM and CGC Formation

If NV v_i *receives the confirmation message from CH$_j$* ***then***
    v_i becomes CM$_j$
  Else
    Go to scheme 1
   End_IF
  If CM$_j$ *is located at CH$'_j$s boarder* ***then***
    CM$_j$ becomes CGC$_j$
 End_IF

TABLE 1. The basic parameters of simulation environment.

Parameters	Value
Simulation time	500 s
Simulation area	5000m x 50m
Vehicle speed	0-120km\h
Vehicle density	50-150/km
Simulation runs	30
Transmission range	100-300m
No. of lanes per direction	4
MAC Protocol standard	IEEE 802.11p
Propagation model	Two-ray Ground

vehicles we use DSRC channels, which implement the WAVE module with IEEE802.11p standard for both the physical and MAC layers [39]. In all simulated scenarios three lanes highway and two-lanes urban per direction are used with length's road of 3km and 10km, respectively. The vehicles speed varied from 0km/h where vehicles stopped at traffic light intersections or paused due to traffic congestions inside urban city, and 120k/h for highways speed. The total number of vehicles for each direction ranges from 50 to 150 vehicles. Summary of the simulation parameters are shown in Table.1.

To compare efficiency of the ECHS scheme with its counterpart the following six important metrices are considered, since they are widely used in previous related works [13], [24]–[28], and show the capability of the proposed scheme in handling cluster problems in VANETs:

- **Cluster Head Lifetime (CHL):** represents the interval time starting from when a vehicle changes its state to CH to when it becomes non-CH.
- **Cluster Member Lifetime (CML):** represents the interval time starting from when a vehicle changes its state to CM (when it joins the cluster) to when it changes its state (when it leaves the cluster).
- **Cluster Number (CN):** the number of clusters created during the simulation time.
- **Packet Loss Ratio (PLR)**: represents a ratio of a number of sent packets to a number of received packets at the destination in a unit of time.
- **Overhead for Clustering (OC):** shows additional signaling overhead to form and maintain the cluster structures
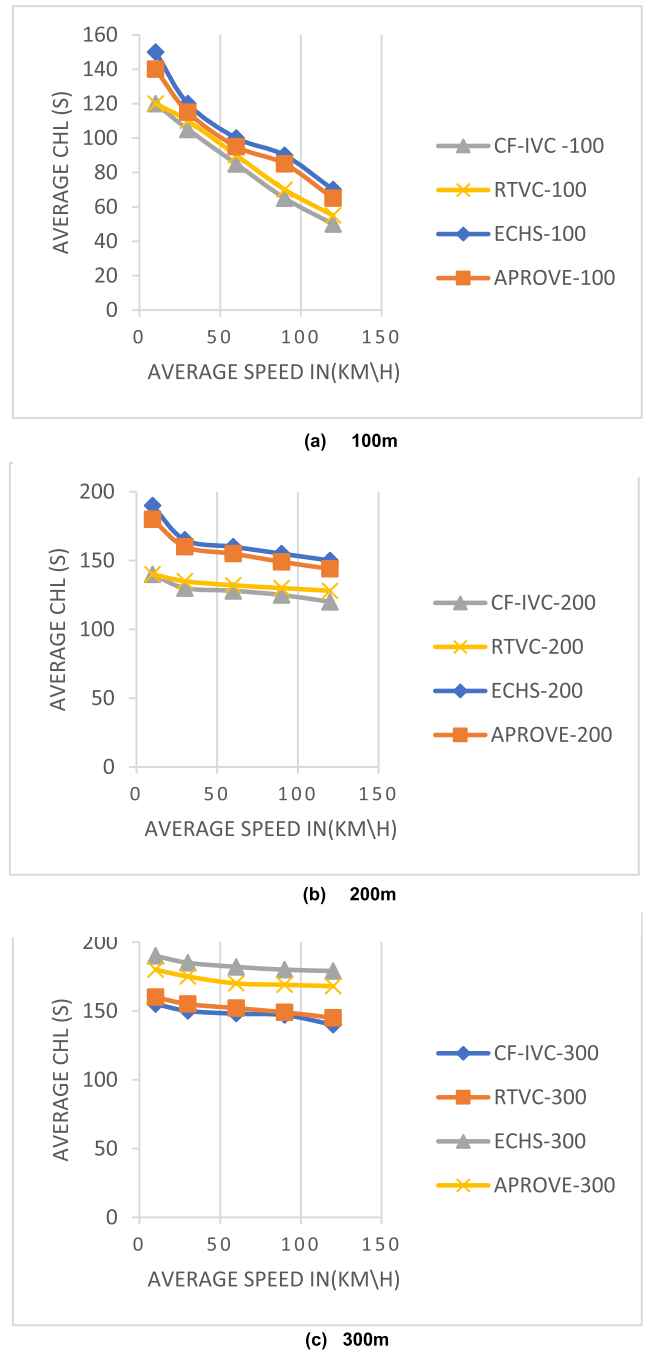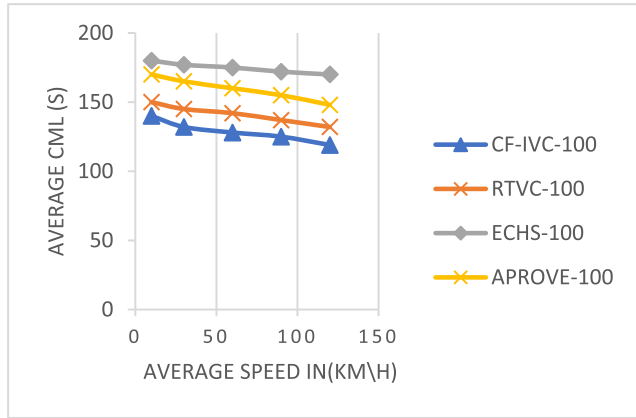

(a)   100m

(b)   200m

(c)   300m

FIGURE 6. Cluster Head Lifetime (CHL) VS Average Speed under different transmission range. (a) 100m. (b) 200m. (c) 300.

- **Average Packet Delay**: refers to as the average delay between times at which the data packet was transmitted from the source vehicle until the time it is received at the destination vehicle.
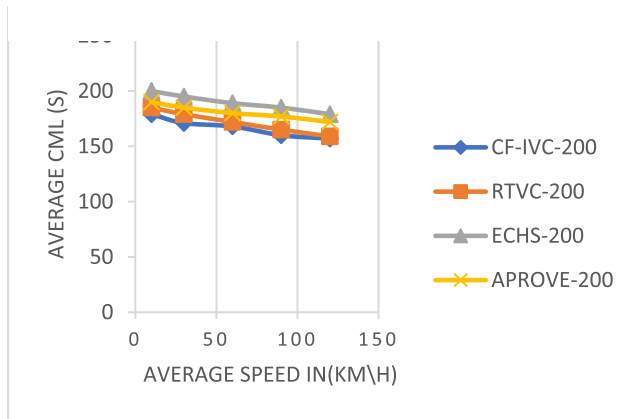
A. CLUSTER HEAD LIFETIME (CHL)

Fig.6 show the average CHL for the proposed scheme against the APROVE, the RTVC and the CF-IVC schemes, under different vehicles speed and different transmission ranges
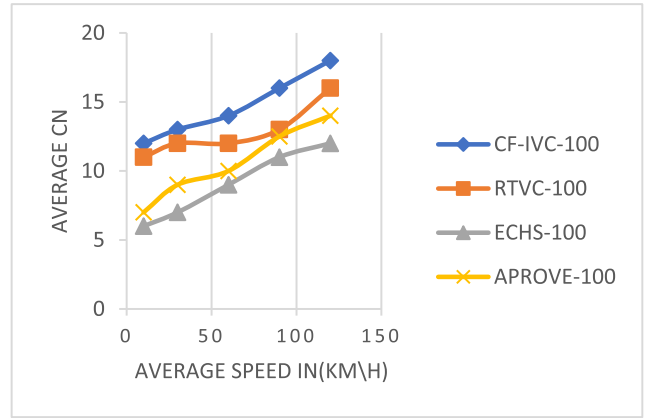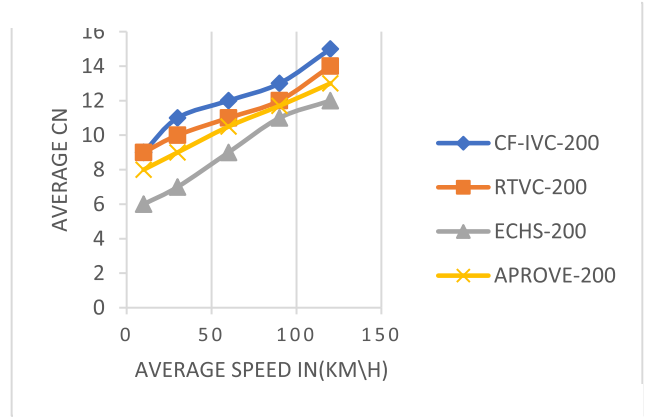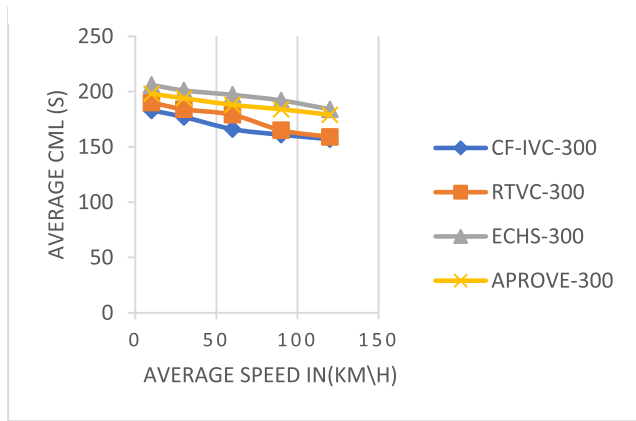
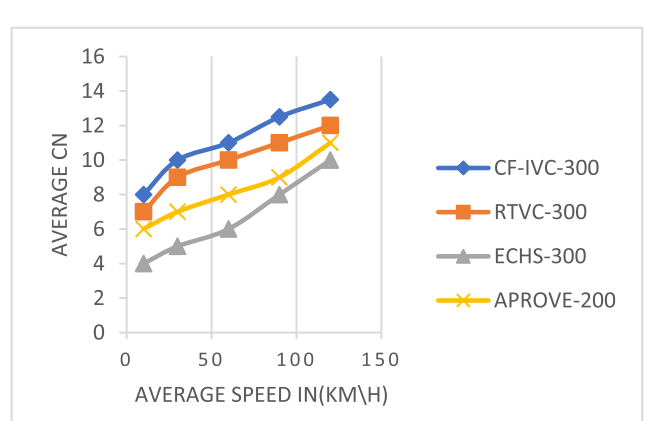**FIGURE 7.** Cluster Member Lifetime (CML) VS Average Speed under different transmission range. (a) 100m. (b) 200m. (c) 300.

FIGURE 8. Cluster Number (CN) VS Average Speed under different transmission range. (a) 100m. (b) 200m. (c) 300.

(100m, 200m, 300m). The results demonstrate that when a vehicle's speed increases the average CHL for all proposed schemes decreases relatively. This is due to when the vehicles move faster, the network topology changes dramatically, and as a result it's very difficult for the CH to keep long connectivity with its CMs. The results also reveal that when the transmission range increases the CHL also increases. It is clearly that wider coverage always enables the CH to connect longer time with CMs and reduces number of joining/leaving times for vehicles. The ECHS scheme performs better than

its counterparts as it imposes strict conditions when electing the CH that always is located at the central of group vehicles.

B. CLUSTER MEMBER LIFETIME (CML)

Fig.7 also shows the average CML for the ECHS scheme compared to the APROVE, the RTVC and the CF-IVC schemes under different vehicles speed and transmission rages. Similarly, based on the discussion on previous section we can conclude both vehicles speed and the transmission range affect dramatically on CML. The lifetime of CM is
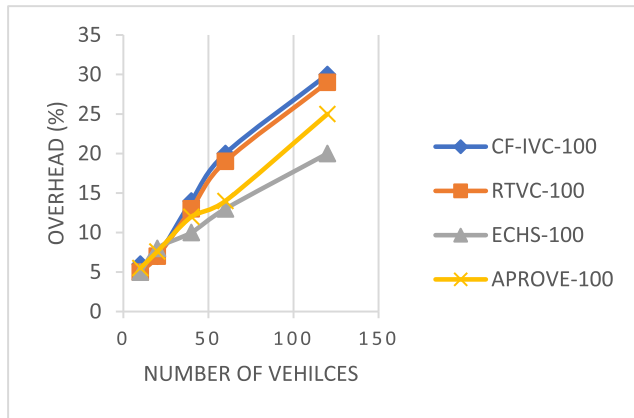
FIGURE 9. Overhead for Clustering (OC) VS Number of Vehicles with 100m transmission range.
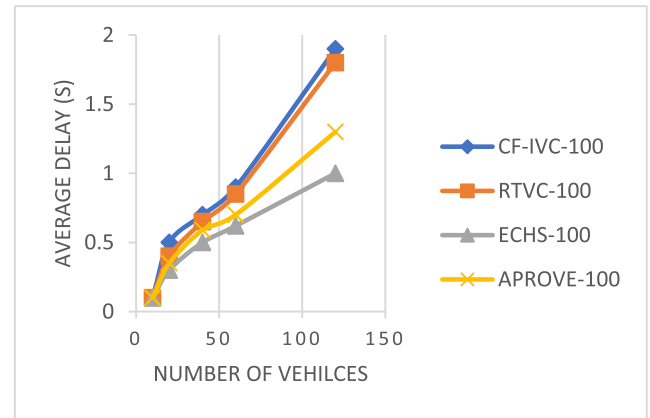

FIGURE 10. Packet Loss Ratio (PLR) VS Number of Vehicles with 100m transmission range.

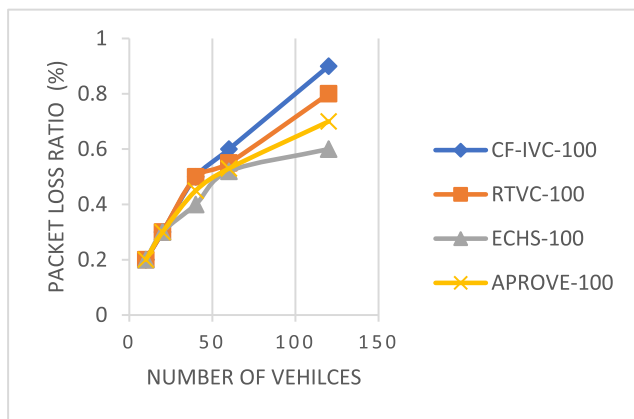**FIGURE 11.** Average Delay VS Number of Vehicles with 100m transmission range.

decreased by increasing its speed, and it is increased when the transmission range increases. It is clearly from the Fig.7, the ECHS scheme outperforms APROVE, RTVC and CF-IVC schemes in terms of CML.

C. CLUSTER NUMBER (CN)

The performance of three clustering schemes in terms of CN is illustrated in Fig.8. As CN decreases network performance increases. We observed from the results in the figures, that increasing the transmission range reduced CN to be formed. The logical explanation of this fact is the coverage area of the current cluster increases in a wide range of transmission, and that makes the CH to include several CMs to serve them. The proposed approach exhibited better performance for all the different transmission ranges.

D. AVERAGE PACKET OVERHEAD, LOSS RATIO AND DELAY

Results in Fig.9 shows the average CO by changing the number of vehicles, while the transmission range is set to 100m and vehicles speed to 60km/h. The obtained results are expected because increasing the number of vehicles is normally increases the number of clusters created and total messages exchanged. Generally, CHA and CMR messages are increased when the number of vehicles increases, since when a cluster is formed normal vehicles will send more CMR messages. Obviously, the overhead generated by the ECHS scheme is less than the overhead generated the APROVE, the RTVC and CF-IVC schemes, since the ECHS scheme keeps long communication links with its member thus decreasing an unnecessarily re-clustering process. As a result, total clustering related messages are reduced by the ECHS scheme as a vehicle does not need to switch to another cluster frequently.

PLR for all schemes under different vehicle densities is presented in Fig.10. As the number of vehicles is increased the packet loss ratio is increased also. This is due to frequent collisions between packets and high contention between vehicles on the same wireless channel. Since the ECHS generates lowest average CN with maximum CHL and CML compared to other schemes, it also incurs minimum packet loss ratio. In general, a cluster instability can increase the number of lost packet and total overhead.

Fig.11 shows the average delay at maximum vehicle densities for the proposed clustering algorithm. The results show that the proposed ECHS achieves lower delay compared to the referenced schemes. High clustering stability and low communication overhead among the clusters and cluster members are the main reasons for low end-to-end delay in the proposed scheme. Another reason for lower network delay in the proposed scheme ECHS is the careful selection of CGWs with using adjusted *backoff* timer.

V. COCLUSION AND FUTURE WORK

In this paper we developed a new robust clustering scheme for large-scale and dynamic VANETs. Based on the general mobility information and neighborhoods knowledge for vehicles, the ECHS scheme outperforms existing counterparts in different aspects. For instance, it reduces unnecessary clusters by introducing new conditions when nominating CHs. Vehicles located at the both edges of the cluster are not allowed to become CHs, but only those at the middle as they

have longer connectivity with their neighbors. The ECHS also groups vehicles that have similar speeds and directions which guarantees cluster stability and reliability. CGWs in the ECHS scheme are selected very carefully and performing retransmissions is assigned to the most appropriated CGW in the cluster. This can reduce duplicated retransmissions and utilize network channel.

The performance of the ECHS scheme is compared with the RTVC [32] and the CF-IVC [25] schemes, in terms of important metrics such as Cluster Head Lifetime, Cluster Member Lifetime and Cluster Number. The ECHC scheme shows superior performance over its counterparts and demonstrates that vehicles speed and direction are not always enough to build stable clusters in the dynamic VANETs topology.

In our future work, we will investigate the performance of the proposed scheme in more complex scenarios and other VANETs challenges. For instance, the ECHS scheme can be incorporated with routing protocols to enhance routing strategy under a high dynamic environment. It can be also used for fast data dissemination in several safety related applications. Thus, the advantages of the ECHS scheme can be explored further with our previous works such as in [40].

REFERENCES

[1] O. Kaiwartya, A. H. Abdullah, Y. Cao, A. Altameem, M. Prasad, C.-T. Lin, and X. Liu, "Internet of vehicles: Motivation, layered architecture, network model, challenges, and future aspects," *IEEE Access*, vol. 4, pp. 5356–5373, Apr. 2016.

[2] A. Boukerche and R. Grande, "Vehicular cloud computing: Architectures, applications, and mobility," *Comput. Netw.*, vol. 135, pp. 171–189, Apr. 2018.

[3] R. Shrestha, R. Bajracharya, and S. Y. Nam, "Challenges of future VANET and cloud-based approaches," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–15, May 2018.

[4] H. A. Khattak, S. U. Islam, I. U. Din, and M. Guizani, "Integrating fog computing with VANETs: A consumer perspective," *IEEE Commun. Standards Mag.*, vol. 3, no. 1, pp. 19–25, Mar. 2019.

[5] J. Fukumoto, N. Sirokane, Y. Ishikawa, T. Wada, K. Ohtsuki, and H. Okada, "Analytic method for real-time traffic problems by using contents oriented communications in VANET," in *Proc. 7th Int. Conf. ITS Telecommun.*, Sophia Antipolis, France, Jun. 2007.

[6] T. Nadeem, S. Dashtinezhad, C. Liao, and L. Iftode, "TrafficView: Traffic data dissemination using car-to-car communication," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 8, no. 3, pp. 6–19, Jul. 2004.

[7] L. Wischhof, A. Ebner, H. Rohling, M. Lott, and R. Halfmann, "SOTIS–a self-organizing traffic information system," in *Proc. 57th IEEE Semiannual Veh. Technol. Conf. (VTC-Spring)*, Jeju-do, South Korea, Apr. 2003.

[8] R. Doolan and G. Muntean, "EcoTrec—A novel VANET-based approach to reducing vehicle emissions," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 3, pp. 608–620, Mar. 2017.

[9] J. Sanguesa, J. Barrachina, M. Fogue, P. Garrido, F. Martinez, J.-C. Cano, C. Calafate, and P. Manzoni, "Sensing traffic density combining V2 V and V2I wireless communications," *Sensors*, vol. 15, no. 12, pp. 31794–31810, Dec. 2015.

[10] J. Li, Y. Zhang, and Y. Chen, "A self-adaptive traffic light control system based on speed of vehicles," in *Proc. IEEE Int. Conf. Softw. Qual., Rel. Secur. Companion (QRS-C)*, Vienna, Austria, Aug. 2016, pp. 382–388.

[11] N. S. Nafi and J. Y. Khan, "A VANET based intelligent road traffic signalling system," in *Proc. Australas. Telecommun. Netw. Appl. Conf. (ATNAC)*, Nov. 2012, pp. 1–6.

[12] J.-H. Kwon, H. S. Chang, T. Shon, J.-J. Jung, and E.-J. Kim, "Neighbor stability-based VANET clustering for urban vehicular environments," *J. Supercomput.*, vol. 72, no. 1, pp. 161–176, Jan. 2016.

[13] Z. Y. Rawashdeh and S. M. Mahmud, "A novel algorithm to form stable clusters in vehicular ad hoc networks on highways," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, no. 1, Dec. 2012, Art. no. 15.

[14] A. Ziagham and M. NooriMehr, "MOSIC: Mobility-aware single-hop clustering scheme for vehicular ad hoc networks on highways," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 9, pp. 424–431, 2016.

[15] C. Caballero-Gil, P. Caballero-Gil, and J. Molina-Gil, "Self-organized clustering architecture for vehicular ad hoc networks," *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 8, Aug. 2015, Art. no. 384869.

[16] L. Rivoirard, M. Wahl, P. Sondi, M. Berbineau, and D. Gruyer, "CBL: A clustering scheme for VANETs," in *Proc. 6th Int. Conf. Adv. Veh. Syst., Technol. Appl.*, Nice, France, 2017, pp. 19–25.

[17] K. Abboud and W. Zhuang, "Stochastic modeling of single-hop cluster stability in vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 226–240, Jan. 2016.

[18] C. Cooper, D. Franklin, M. Ros, F. Safaei, and M. Abolhasan, "A comparative survey of VANET clustering techniques," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 657–681, 1st Quart., 2017.

[19] R. S. Bali, N. Kumar, and J. J. Rodrigues, "Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions," *Veh. Commun*, vol. 1, no. 3, pp. 134–152, 2014.

[20] C. J. Joshua, R. Duraisamy, and V. Varadarajan, "A reputation based weighted clustering protocol in VANET: A multi-objective firefly approach," *Mobile Netw. Appl.*, vol. 24, no. 4, pp. 1199–1209, Aug. 2019.

[21] A. Ishtiaq, S. Ahmed, M. F. Khan, F. Aadil, M. Maqsood, and S. Khan, "Intelligent clustering using moth flame optimizer for vehicular ad hoc networks," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 1, Jan. 2019, Art. no. 155014771882446.

[22] R. Hajlaoui, E. Alsolami, T. Moulahi, and H. Guyennet, "An adjusted K-medoids clustering algorithm for effective stability in vehicular ad hoc networks," *Int. J. Commun. Syst.*, vol. 32, no. 12, pp. 1–16, 2019.

[23] O. Senouci, Z. Aliouat, and S. Harous, "MCA-V2I: A multi-hop clustering approach over vehicle-to-Internet communication for improving VANETs performances," *Future Gener. Comput. Syst.*, vol. 96, pp. 309–323, Jul. 2019.

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

[25] O. Kayis and T. Acarman, "Clustering formation for inter-vehicle communication," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Seattle, WA, USA, Sep. 2007.

[26] B. Hassanabadi, C. Shea, L. Zhang, and S. Valaee, "Clustering in vehicular ad hoc networks using affinity propagation," *Ad Hoc Netw.*, vol. 13, pp. 535–548, Feb. 2014.

[27] S. A. Mohammad and C. W. Michele, "Using traffic flow for cluster formation in vehicular ad-hoc networks," in *Proc. IEEE LCN*, Denver, CO, USA, Oct. 2010, pp. 631–636.

[28] A. S. K. Mammu, U. Hernandez-Jayo, and N. Sainz, "Direction aware cluster-based multi channel MAC protocol for vehicular ad hoc networks," in *Proc. IEEE 11th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Abu Dhabi, United Arab Emirates, Oct. 2015, pp. 549–556.

[29] A. B. Tambawal, R. M. Noor, R. Salleh, C. Chembe, and M. Oche, "Enhanced weight-based clustering algorithm to provide reliable delivery for VANET safety applications," *PLoS ONE*, vol. 14, no. 4, pp. 1–19, 2019.

[30] G. H. Alsuhli, A. Khattab, and Y. A. Fahmy, "Double-head clustering for resilient VANETs," *Wireless Commun. Mobile Comput.*, vol. 2019, Mar. 2019, Art. no. 2917238.

[31] M. D. Venkata, M. M. M. Pai, R. M. Pai, and J. Mouzna, "Traffic monitoring and routing in VANETs—A cluster based approach," in *Proc. 11th Int. Conf. ITS Telecommun.*, Saint Petersburg, Russia, Aug. 2011, pp. 27–32.

[32] W. Farooq, M. Ali Khan, and S. Rehman, "A novel real time framework for cluster based multicast communication in vehicular ad hoc networks," *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 4, Apr. 2016, Art. no. 8064908.

[33] J. Zheng, H. Tong, and Y. Wu, "A cluster-based delay tolerant routing algorithm for vehicular ad hoc networks," in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Sydney, NSW, Australia, Jun. 2017, pp. 1–5.

[34] N. Maslekar, J. Mouzna, H. Labiod, M. Devisetty, and M. Pai, "Modified C-DRIVE: Clustering based on direction in vehicular environment," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Paris, France, Jun. 2011.

[35] L. Liu, C. Chen, T. Qiu, M. Zhang, S. Li, and B. Zhou, "A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs," *Veh. Commun.*, vol. 13, pp. 78–88, Jul. 2018.

[36] I. S. Alkhalifa and A. S. Almogren, "NSSC: Novel segment based safety message broadcasting in cluster-based vehicular sensor network," *IEEE Access*, vol. 8, pp. 34299–34312, 2020.

[37] *NS3: The Network Simulator*. Accessed: Nov. 2019. [Online]. Available: https://www.nsnam.org/

[38] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, "SUMO—Simulation of urban mobility," in *Porc. SIMUL*, Barcelona, Spain, 2011, pp. 55–60.

[39] S. Eichler, "Performance evaluation of the IEEE 802.11p WAVE communication standard," in *Proc. IEEE 66th Veh. Technol. Conf.*, Kauai, HI, USA, Sep. 2007, pp. 1–5.

[40] M. Banikhalaf, R. Hammad, A. Al-dubai, and B. Ghaleb, "An adaptive warning message scheme for emergency vehicles using vehicular ad hoc communication," *Int. J. Eng. Technol.*, vol. 7, no. 4, pp. 3182–3189, 2018.

MUSTAFA BANIKHALAF received the Ph.D. degree from the School of Computing, Edinburgh Napier University, in 2013. He is currently an Assistant Professor with the Computer Science Department, Yarmouk University. He has been working on the area of group communications, mobile and vehicular ad hoc networks, and high-performance networks. In 2010, he received the Merit Competition-Award for his Ph.D. studentship.

MOAIAD AHMAD KHDER (Senior Member, IEEE) received the Ph.D. degree from the Faculty of Information Science and Technology, The National University of Malaysia, in 2015. He is currently an Assistant Professor with the Computer Science Department, Applied Science University, Bahrain. He has been working on the area of mobile environment, mobile and vehicular ad hoc networks, and mobile database.

• • •