

A Novelty of Hypergraph Clustering Model (HGCM) for Urban Scenario in VANET

MAYS KAREEM JABBAR^{1,2} AND HAFEDH TRABELSI¹

¹CES Laboratory, École Nationale d'Ingénieurs de Sfax (ENIS), University of Sfax, Sfax 3029, Tunisia

²Faculty of Engineering, University of Misan, Amarah, Misan 62001, Iraq

Corresponding author: Mays Kareem Jabbar (m_mays85@uomisan.edu.iq)

ABSTRACT A vehicular ad hoc network is a dynamic and constantly changing topology that requires reliable clustering to prevent connection failure. A stable cluster head (CH) prevents packet delay (PD) and maintains high throughput in the network. This article presents a two-fold novel scheme for stable CH selection. In the first part of the proposed scheme, the vehicle network is considered a one-to-many connection network, which is near to a practical scenario. The cluster generation is handled using a newly proposed vehicular-hypergraph-based spectral clustering model. In the second part, the CH is selected considering the criteria for maintaining a stable connection with the maximum number of neighbours. The new rewarding/penalising relative speed and neighbourhood degree fulfil the condition. Eccentricity assesses that the vehicle should be at the centre of the cluster. Another metric with deep learning spectrum sensing is introduced for CH selection. Trust calculation is performed using deep learning-trained spectrum sensing as a model. The primary vehicle in noisy and noiseless environments is recognised using layers of long short-term memory. A high trust score is awarded to the vehicle which vacates the spectrum in the sensing of the primary vehicle. The stable CH selected by these metrics reduces the overhead that occurs due to the frequent shifting of the CH from one vehicle to another. This has been validated by the improved CH stability; increased cluster member (CM) lifetime and reduced rate of change of CH. The proposed scheme also demonstrates a considerable improvement in PD and throughput.

INDEX TERMS Cluster head stability, eccentricity, hypergraph, trust, VANET.

I. INTRODUCTION

A. BACKGROUND

A vehicular ad hoc network (VANET) is an amalgamation of existing ad hoc networks, wireless LAN and cellular technology to achieve road traffic safety and efficiency. VANETs stand out due to their hybrid network architecture, node movement and various application scenarios. Therefore, these networks pose several unique networking research challenges [1]. VANETs are amongst the significant applications of intelligent transportation systems. They include various applications, such as cooperative traffic monitoring, control of traffic flows, blind crossing, prevention of collisions, nearby information services and real-time detour route computation[2].

VANETs have a highly dynamic topology, with a high relative speed of vehicles and frequent discontinuity in the networks. The topology management in VANETs can be

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Tucci¹.

achieved through clustering. To manage the topology, similar moving speed vehicles are grouped. The more effective is the clustering, the better is the topology control [3]. Each cluster has a CH that communicates with the external controlling stations such as the road side unit (RSU) [4]. The vehicles are grouped based on either their location or user information, such as direction and speed. Both approaches are based on mathematical formulation, leaving behind the sociological aspects of how the driver will change the lane and his destinations. Cluster generation and maintenance are a distributed approach. However, cluster generation is considered centralised by many researchers [5]. Likewise, many challenges are faced in an urban scenario as the heavy intersection of roads, concrete infrastructure and heavy density of vehicles lead to variable topology and mobility with time dependency. The weak communication link is another major issue.

B. OUR CONTRIBUTION

This work proposes a novel approach for clustering formation and maintenance of a VANET structure in an urban scenario.

It is called the hypergraph clustering model (HGCM). The CH stability is governed by a cumulative multimetric factor inclusive of relative speed, eccentricity, neighbourhood and spectrum sensing based on cooperative trust. The contributions of this work are listed below:

- As per our knowledge, this work is the first of its kind to introduce a formulation of VANET through a hypergraph. The construction of the hypergraph is designed using the distance proximity amongst the vehicles in the network.
- Practical and optimal partitioning of the hypergraph through tensor trace maximisation (TTM) is proposed. A high order has all the edges but with negligible weights. Thus, the adjacency matrix is nearly sparse, and the overall computational complexity is effectively reduced.
- Optimal clusters are selected in accordance with the Calinski–Harabasz concept. This method is an external criterion for selecting optimal clusters. Hence, the information is independent, and the structure of the information is inherited. Such a method is also preferred for convex clustering.
- The network’s performance, especially in an urban scene, can be improved by installing auxiliary facilities, such as RSUs. Here, an evolving graph structure of the traffic is conceived using betweenness centrality.
- Spectrum sensing is redesigned as a classification problem. The method proposed for sensing is long short-term memory (LSTM), which is extensively trained for all signal types, including noise. Thus, it can sense an untrained signal and classify a vehicle as primary or secondary.
- The scheme of a cumulative multimetric for selecting a CH is presented, through which strong connectivity and stable link lifetime are maintained. The stability of the CH enhances the routing performance of the designed approach. Extensive simulation and comparison of the proposed scheme with existing state-of-the-art techniques are presented to show its supremacy.

C. PAPER ORGANISATION

This paper discusses a new hypergraph-based multimetric CH selection algorithm that increases the CH stability. The algorithm is tested on the Baghdad City map. The rest of this paper is organised as follows. Section II presents a literature review of some clustering algorithms and research problems. The proposed hypergraph generation method is introduced in Section III. Section IV introduces the CH selection metrics, cluster maintenance phase and time complexity for the proposed method. A simulation and a comparative analysis with existing state-of-the-art techniques are presented in Section V, and concluding statements with future prospects are provided in Section VI.

II. LITERATURE REVIEW

A. RELATED WORK

An intensive literature survey is conducted, especially on the user information-based methods employed for clustering.

In 2001, Basu *et al.* [6] primarily designed the MOBIC method for MANETs on the basis of the ratio of the power levels received at each node from its neighbours. The authors proposed a distributed clustering algorithm for CH selection. The algorithm was tested on a simulated area in NS-2 but not in an actual scenario, and the method considered only the radio signal. Code division multiple access (CDMA) was proposed by Kayis and Acarman [7]. It was an intervehicle communication scheme, in which nodes were assigned a specific task autonomously on the basis of their speed, then clusters were formed. Only a theoretical approach was projected here. Su *et al.* [8] designed a scheme based on a multichannel communication with three primary protocols for reducing data congestion, QoS and efficient bandwidth over a vehicle-to-vehicle (V2V) network. The method proposed by Rswshedh *et al.* [9] grouped vehicles in accordance with mobility patterns and minimised the total number of created clusters. Another work introduced by Maslekar *et al.* [10] also reported using the location and direction of vehicles for CH selection. Here, a direction-based clustering algorithm was used for data dissemination. Through this approach, the protocol density of vehicles on a given road was estimated. A novel concept of node precedence algorithm was proposed by Goonewardene *et al.* [11]. Robust mobility adaptive clustering (RMAC) identified a one-hop neighbour and selected a CH using the relative node mobility metrics of speed, location and travel direction. An evolving structure was created via neighbourhood analysis. With the same mobility concept, another scheme using the affinity propagation algorithm was proposed in a distributed manner by Shea *et al.* [12]. The authors claimed the existence of clusters with high stability.

The stability of CH has been the main focus of researchers. For instance, a multilevel clustering algorithm that mainly concentrated on the stable, long lifetime of clusters was designed by Vodopivec *et al.* [5]. The cluster formation was based on factors, such as the density of the connection graph, link quality and traffic conditions. Mohammad and Michele [4] designed another method to improve cluster lifetime, especially in an urban scenario, based on the lane concept. The technique was effective in reducing the overhead of reclustering and led to an efficient hierarchical network topology. A three-based passive clustering was introduced by Wang and Lin [13] to improve intervehicle communication and eventually cluster stability. The method also reported enhanced performance in network analysis. A beacon-based clustering algorithm was proposed by Souza *et al.* [14] to re-establish cluster stability in the case of reclustering. Another proposed method in [15] formulated a weighted approach that included such parameters as the number of neighbours based on dynamic transmission range, the direction of vehicles, entropy and distrust value. The authors tried to increase

stability and connectivity with the reduction in overhead. They also tested the adaptive allocation of transmission range technique. An adaptable mobility-aware clustering algorithm based on destination positions (AMACAD) was introduced by Morales *et al.* [2]. The authors presented the concept of adaptive mobility. Wolny [16] primarily focused on improving cluster stability by modifying the DMAC algorithm. An adaptive service provider infrastructure (ASPIRE) architecture was designed by Koulakezian [17] by using the concept of vehicular mobility in a highway scenario. The author allowed vehicles to connect to the network through regular mobile IP nodes, thereby increasing the connectivity and decreasing the overhead by caching in clusters. The same ideology of mobility was reintroduced using the concept of multihop, and a new clustering scheme based on multihop was presented by Zhang *et al.* [18]. Considering that vehicle direction is also an important parameter, Maslekar *et al.* [19] introduced the concept of CH switching. The speed difference between neighbouring nodes was taken to obtain a stable clustering structure [20], [21]. The interest in the multihop approach is evident amongst researchers. The establishment of a multihop-based clustering scheme using neighbourhood has been exploited [22], [23]. Relevant approaches allow an optimal set of nodes to join the same cluster and increase stability. With the advances in technology, researchers have also proposed a complete solution for VANETs, which includes CH selection, cluster formation and maintenance. A link reliability-based clustering algorithm (LRCA) was designed by Ji *et al.* [24] to filter out unstable neighbours on the basis of the link knowledge providing a complete solution of VANET. Alsuhi *et al.* [25] proposed another approach called double-head clustering (DHC). This approach conceived a complete solution by using metrics, such as vehicle speed, position and direction, in addition to other metrics related to the communication link quality, such as the link expiration time and the signal-to-noise ratio. A novel concept of primary and secondary CHs, called enhanced weight-based clustering algorithm (EWCA), was introduced by Bello Tambawal *et al.* [26] to improve stability. Short-range vehicle communication-based clustering in VANETs, named centre-based secure and stable clustering (CBSC) algorithm, was proposed by Cheng and B. Huang [27].

The literature survey comparatives based on the traffic scenarios, achievements, limitations, and various metrics used for cluster formation are presented in Table 1.

B. RESEARCH PROBLEMS

This complete literature survey is for two decades, in which the primary focus has been the stability of CH. However, VANET clustering requires cluster generation/maintenance and CH selection as two research areas. During the study, the following few issues have been identified:

- The work on the real scenario is mostly limited to the highway; analysis on the urban scenario is missing.

- The mobility and neighbourhood are the most metrics taken, and these metrics are lost in the urban scenario as the vehicle speed is low and there is huge congestion in peak hours.
- VANET is a continuously changing topology, which creates challenges in establishing a connection from one source to the destination vehicle. If the connection is multihop, then the data loss can be high as the carrying vehicle may change the direction and speed. Accordingly, the information should be transmitted in a single hop, which is feasible by making a reliable cluster. A CH is elected amongst the neighbouring nodes. Along with efficient cluster generation, the CH-annotated vehicle is responsible for improving the network performance. Sustaining a CH for a long period is difficult.
- Researchers have made clusters and selected the CH by calculating the vehicles' behaviour in the network, velocity, moving direction and position in lanes. Fuzzy logic has been used, such as in [29], [30], along with other heuristic algorithms, such as in [33], [34]. Fuzzy logic schemes require tuned membership functions to decide for CH selection, which necessitate considerable experience and behaviour analysis of vehicles on a particular road [35]. Owing to fast-changing topology and distributed VANET architecture, heuristic algorithms hardly make decisions due to several iterations in the calculation. They cannot cope up with the changing frequency of a VANET environment in a crowded city. However, we assume that in highways with few vehicle densities, they may present good performance.

With all these gaps in the literature studied, our primary goal is to design a complete solution for VANET, especially for the urban scenario. The designed approach has cluster formation, CH selection and maintenance. The evolving nature of VANETs is meritoriously captured using the concept of hypergraphs and the clusters are formed through the designed hypergraph algorithm. For CH selection, a cumulative multimetric is designed to consider four factors: relative speed, trust, neighbourhood and eccentricity. The trust metric is introduced with deep learning spectrum sensing. Spectrum sensing using LSTM is introduced. This deep learning solution enables the detection of signals. It allows effective utilisation of the spectrum, especially in an urban scenario where excessive congestion occurs during peak hours. The spectrum can be used by primary users (PUs, e.g. ambulance, police or emergency). A cumulative multimetric maintains strong connectivity and stable link lifetime. The proposed scheme is tested on the real map of Baghdad.

III. PROPOSED MODEL

A multilane road structure in an urban scenario is considered. Fluctuating density of building infrastructure and vehicular mobility with total number of vehicles N are infused into the real map scenario with corresponding speed and locations. This scenario is depicted in Figure 1. Each on-board

TABLE 1. Comparative literature survey on cluster formation metrics in the existing state of the art.

Year	Algorithm	Traffic Scenario	Achievements	Limitations	Metric
2001	MOBIC [6]	Highway	Low rate of change of CH is realised.	It was experimented on MANET, and dynamic VANET experimentation was not performed.	Radio Signals
2007	CDMA [7]	-	Vehicles' orthogonal codes avoid data collision.	Simulated validation is lacking.	Speed
2007	Su et al.[8]	Highway	Safety message delivery by CH is improved.	Real-time delivery of messages was not evaluated.	Directions
2009	Rswshedh et al.[9]	Highway	The CH stability is independent of speed.	It is claimed to minimise the number of clusters, but validation is missing.	Speed, Location , Direction
2009	Maslekar et al. [10]	-	The data dissemination problem is solved by optimal CH selection.	Only vehicles at constant speed were considered.	Location , Direction
2009	RMAC [11]	Highway	The algorithm is robust to a highly dynamic VANET.	Frequent change of CH is the issue, which was not dealt with in the article.	Speed, Location , Direction
2009	APROVE [12]	Highway	It achieves longer cluster stability compared with the state of the art.	The security or bandwidth sharing was not explored in the work.	Distance Speed
2010	Wang[13]	-	It reduces the overhead in routing through clustering by using vehicle link metrics in clustering.	The relative velocity metric was ignored, which is widely used in cluster stability.	Density, Communication Link
2010	VWCA [15]	Highway	Communication overhead is decreased in linking a new cluster due to membership stability.	Algorithm validation with various vehicle densities was not done.	Trust, Direction
2011	MDMAC [16]	Highway	The CH stability is improved.	The algorithm's dependency on the node's actual position and velocity is a disadvantage.	Location, Speed, Final Destination Range
2011	Zhang et al.[18]	Highway	N-hop clustering performs better than single-hop clustering.	Bandwidth access and security for n-hop were not discussed.	Mobility
2011	Maslekar et al. [19]	Highway	Density estimation-based clustering exerts less overhead than linkage clustering.	The lifetime of the vehicles in the cluster was ignored.	Direction
2012	TB [20]	Highway	It increases the cluster lifetime and reduces the overhead within a cluster by clustering the fast- and slow-moving vehicles in the same cluster.	Vehicle velocity was considered uniform throughout the simulation, which could match to a highway scenario to some extent but would fail in an urban environment.	Speed, Location, Direction
2014	Arkian et al. [21]	Highway	The new traffic generalisation method improves the CH stability.	Bandwidth congestion in the cluster and sparse network is still a challenge for this work.	Speed, Neighbours
2015	DMCNF [22]	Highway	Neighbourhood relative speed and clustering history are the advantages of this algorithm.	The routing of messages in this algorithm was not explored.	Speed, Neighbours
2018	PMC [23]	Highway	The number of neighbours, expected transmission count and time to leave are added to multi-hop clustering.	Routing in passive clustering was not discussed.	Speed, Neighbours, Link lifetime ,Position
2018	LRCA [24]	Urban	Link reliability-based clustering and routing are the advantages for the CH stability.	End-to-end delay in the routing algorithm was not evaluated.	Speed, Link lifetime ,Position
2019	DHC [25]	Highway, urban	Two CH selection strategies were investigated and advanced in stability.	Multiple features are used in CH selection, which increases the complexity of selection of the optimal set of parameters for enhanced stability.	Signal Strength, Relative Speed, Link Lifetime
2019	EWCA [26]	Highway	The weighted values calculated based on mobility make the stability robust to vehicle density and transmission range.	Clustering collision problem arises as a vehicle joins another cluster without leaving the previous one due to independency from the transmission range.	Speed, Position
2019	CBSC [27]	Highway	The stability scheme is independent of the driver's driving behaviour, and security is also considered.	The evaluation was done on a highway scenario in which driving behaviour poses no challenge.	Position, Relative Speed
2019	CVoEG [28]	Highway	The method is scalable, and clustering is independent of the network density.	Two nodes in graphs are considered connected, although a vehicle can be connected to many other vehicles.	Link Lifetime, Position, Relative Speed
2019	FCMS [29]	Highway	Security, relative speed and eccentricity are the advantages in clustering.	Fuzzy logic is used to assign a value on the basis of three inputs. The rule and membership function set needs extensive knowledge or practice.	Relative speed, degree, security, and trustworthiness
2020	GOA [30]	-	Grasshopper optimisation tunes the network parameters for enhanced stability.	The use of an optimisation algorithm in VANET is not feasible due to the frequently changing topology.	Grasshoppers' optimisation-based node.
2020	QMM [31]	Highway	The stability and connectivity during communications are increased.	Limits to highway only	QoS of vehicles.
2021	Maan et al.[32]	Highway	The fuzzy logic with software-defined networking improves the cluster stability.	Fuzzy logic usage imposes a knowledge constraint to define the optimal rules.	Network connectivity level, lane weight, average velocity, and average distance.
2021	Khan et al.[33]	Highway	Spectral clustering reduces the computational cost due to specified inter and intra clustering.	Extensive evaluation on various parameters is required.	Link connectivity

unit (OBU) equipment has the same transmission range R_{veh} . The same communication module is taken to communicate with auxiliary facilities, each having a transmission range R_{RSU} . Communication amongst vehicles is carried out using the V2V protocol. A Global Positioning System (GPS) unit and IEEE 802.11p radio equipment are embedded inside the OBU. By contrast, Vehicle to Infrastructure is used for communication between vehicles and RSU. Each vehicle in the network serves as a node (V) that performs as either source, destination or router.

The main task of these nodes is to broadcast information within the network. The vehicles are said to be one-hop neighbours if the distance between them is less than R_{veh} and multi-hop if the distance between them is greater than R_{veh} . In this section, the formulation for designing VANET as a hypergraph-partitioning problem is discussed in detail. The commonly used set of notations is presented in Table 2. The complete work can be divided into four major steps:

1. Neighbouring vehicle identification and adjacency matrix generation
2. Hypergraph-based spectral clustering for cluster formation
3. RSU deployment and cluster members (CMs) allotment
4. Stable CH selection

The flow diagram of the complete work is shown in Figure 2.

The pseudo code of the suggested work is jotted down in Algorithm 1.

A. WHY IS VANET A HYPERGRAPH NETWORK

As discussed previously, the problem of stable CH selection to reduce the overhead and packet drop probability is twofold. Vehicle cluster generation is not a novel but a populated concept, and various scholars have already addressed it. A network of vehicles is presented as a graph, in which a vehicular node is connected with two other vehicular nodes [28], [33], [36]. This graphical representation may be suitable in sparse density, such as highways or minimally populated cities. On the contrary, in dense urban scenarios, a vehicle is always connected with more than two vehicles, and graph theory does not fit there. Hypergraphs are a suitable representation of dense vehicle networks. The following are the key reasons to represent VANET as a hypergraph:

- VANET is a cooperative network where every decision depends on the information shared by neighbouring vehicles [37].
- In the graphical representation, a loss of information occurs in paired connections. For tasks such as analysis and learning, pairwise graph models lack the representational ability to adequately capture and show higher-order information. Higher-order interactions in these systems can be captured using hypergraphs, which show interacting components as nodes and hyperedges.
- All vehicles in the network act as either sources, destinations or routers, depending on where they are in the

Algorithm 1 Pseudo Code of the Proposed Scheme for Stable CH Detection

Input: Vehicle-to-vehicle transmission range, RSU-to-vehicle transmission range

1. For $i=1$:vehicles
 - a. All neighbouring vehicles within the vehicle transmission range are identified
2. End for
3. Calculate the similarity matrix on the basis of the distance proximity amongst vehicles
4. Generate an adjacency matrix
5. Create eigen values for each vehicle by using hypergraph-based TTM
6. Apply k-mean clustering with Calinski–Harabasz for the optimal number of clusters.
7. The centrality index gives the RSU deployment location
8. Simulate the VANET for different vehicle densities in the network and collect the vehicle moving angle, locations and other information until every vehicle leaves the network.
9. For $i=1$: clusters
 - a. For $j=1$: vehicles in cluster
 - i. Calculate the CH parameters: neighbourhood degree, relative speed, trust score and eccentricity
 - ii. Combine these matrices to obtain a single score for each vehicle
 - iii. Highest-scorer vehicle is termed as the CH
 - b. End for
10. For $i= 1$: vehicles
 - i. Selected CH sends frequent polling signals
 - ii. If (the distance between the CH and vehicle \leq CH transmission range) then
 - a. The CH receive signal in return within a stipulated time period
 - b. *CH Adds the vehicle in the RSU List&CH Local List*
Else
 - a. The vehicle does not reply within the speculated period time_span
 - b. This vehicle is considered to be disconnected and leave the cluster
 - c. *CH Removes the vehicle from the RSU List&CH Local list*
End if
 - End for
11. End for

network's hierarchy. One of the primary responsibilities of these nodes is to disseminate the data throughout the network. Vehicular mobility and the communication link amongst vehicles that is constantly breaking and

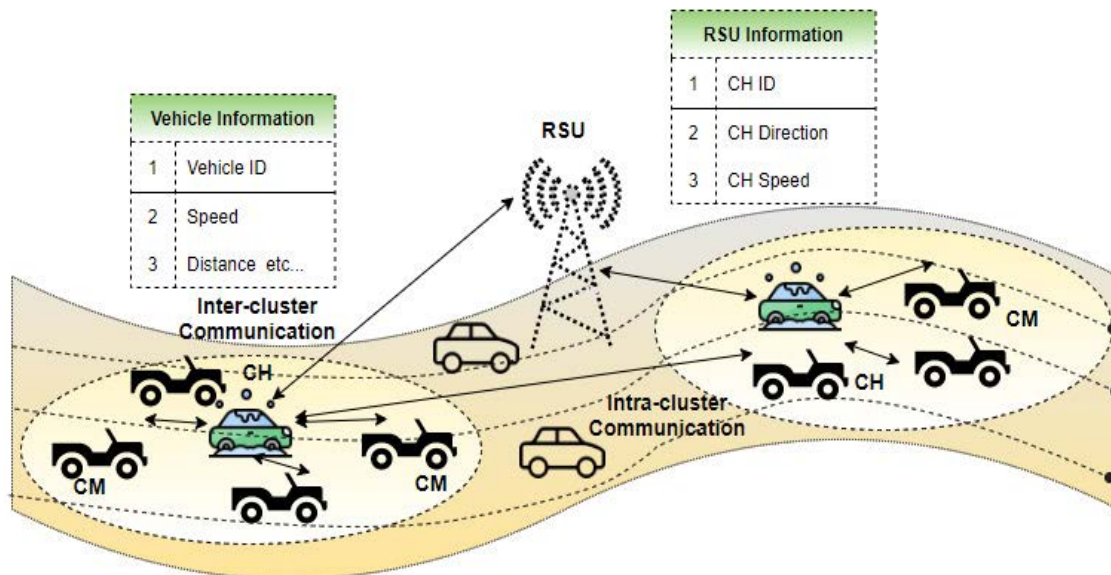


FIGURE 1. The clustering model for VANET.

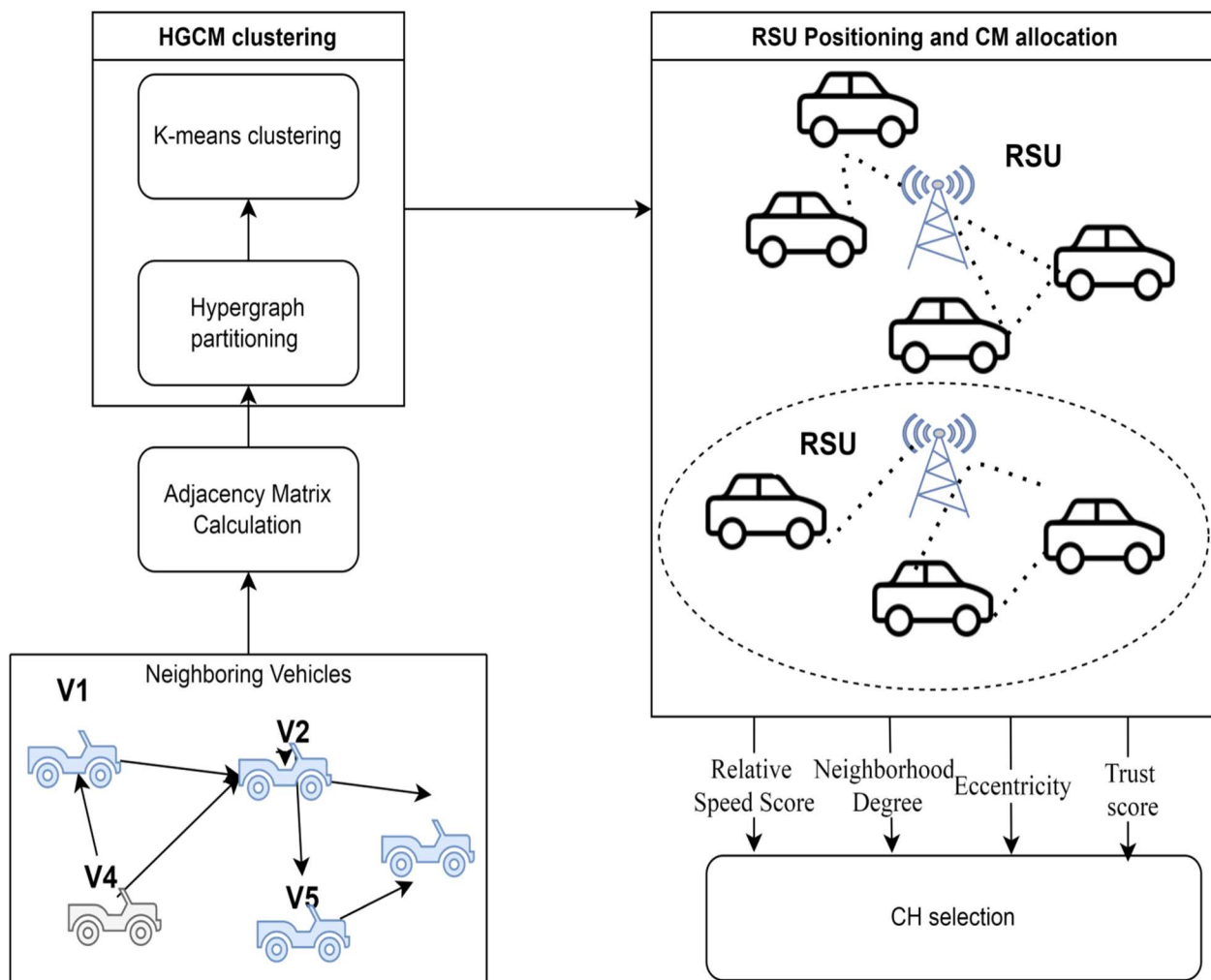


FIGURE 2. Flow diagram of the proposed HGCM scheme.

TABLE 2. Frequently used notations.

Notations	Explanation
$\mathcal{H} = (V, \mathcal{E}, W)$	Hypergraph with V hypernode cardinality of N_H and \mathcal{E} hyperedges
$V \in \{v_1, v_2, v_3, \dots, v_n\}$	V is a set of vertices: vehicles (nodes) with cardinality n
$\mathcal{E} \in \{e_1, e_2, e_3, \dots, e_m\}$	\mathcal{E} is a set of edges: the distance proximity amongst the vehicles with cardinality m
W	Weight of each edge $[0,1]$
G	Graph formulation with V vertex and \mathcal{E} edges
$Par \in \{V_1, \dots, V_k\}$	Partitions of the hypergraph into k disjoint set
A	Similarity (adjacency matrix) constructed for all hypergraph edges
Dig	Diagonal matrix constructed using the normalised cut set algorithm with unnormalised Laplacian formulation
L	Fiedler vector
$f_i(t)$	Vehicle feature
$m_i(t)$	Metric tuple with relative speed, neighbours, eccentricity and trust with $\overline{m_i^k(t)}$ as a normalised metric tuple
c_{ij}	Node connectivity; 1 when $d_{ij} \leq R_{vehi}$; else, 0
X	Set of eigenvectors achieved after the eigen decomposition of the Fiedler vector
t	Total simulation time 1500 s
N	Number of vehicles injected into the network $[100 < N < 1000]$
$Vehi_{ID}$	Vehicle ID as recorded by TraCI
R_{vehi}	Transmission range of vehicles set to 200 m
C_{num}	No. of clusters formed in the network
$tr(B_k)$	Between-group dispersion matrix for a vehicular cluster formed
$tr(Z_k)$	Within-cluster dispersion matrix for a vehicular cluster formed
$N(\mu, \sigma^2)$	Gaussian probability distribution function with mean=0 and variance =1
C_B	Betweenness centrality in a graph formulation for RSU deployment with σ_{st} short path
RSU_{Loc}	Location of auxiliary facilities placed
$C_{optimal}$	Optimal number of clusters
$Vehi_{num}$	Number of vehicles at each cluster
s	Calinski–Harabasz index for evaluating the optimal number of clusters with $tr(Z_k)$ within-cluster dispersion and $tr(B_k)$ between-cluster dispersion
no_lane	Number of lanes on the map
R_{RSU}	Transmission range of the RSU set to 350 m
(δ)	Rewarded or penalised with an absolute value of 0.01
S_{thr}	Speed threshold depending on variance and neighbours set to 2.77
V_{avg}	Average speed of vehicles in a cluster
CH_{score}	CH score
z	Tensor order
λ	Eigenvalues
h	Channel
γ	Signal-to-noise ratio
P_d	Probability of detection
P_f	Probability of false alarm
$T(Y)$	Test statistics
$\mathbb{Q}(\cdot)$	Complementary distribution function
χ^2	Chi-square probability distribution function

reconnecting cause such networks to grow in nature. The relationships amongst nodes in more diversified vehicular networks are more difficult to understand because of the networks' ever-expanding character. As a result, complex networks now utilise super networks. Hypergraph-based and network-based are super networks that exist in the literature [37]. Hypergraphs pay more attention to the dynamic evolution process, making it possible to conduct a dynamic analysis of complicated networks.

- The advantage of the hypergraph theory is that it guarantees the homogeneity of points and edges. This helps to express the relationship between nodes and edges clearly. Thus, networks' representation can be

modulated using a hypergraph, in which one vehicle can communicate with many vehicles.

B. FORMULATION AS HYPERGRAPH PARTITIONING

A connected weighted hypergraph is a three-tuple of $\mathcal{H} = (V, \mathcal{E}, W)$, where each edge \mathcal{E} links a subset of vertices V in the hypergraph and may be linked with non-negative weight. This structure is composed of n vertices, $V = \{v_i | i = 1, 2, \dots, n\}$, where each vertex is a vehicle in our study, $\mathcal{E} = c_{ij} : \{< v_i, v_j > | v_i \in V \wedge v_j \in V \wedge (|d_{ij}| \leq R_{vehi})\}$ is an edge set, and $W \rightarrow [0, 1]$ is the weight associated with each edge. The distance of each vehicle constitutes the edge of each hyperconnection. The hyperedge connection is defined in Definition 1.

Definition 1: Two vehicular nodes v_i and v_j at time t are said to be connected if

$$c_{ij} = \begin{cases} 1 & d_{ij} \leq R_{vehi} \\ 0 & d_{ij} > R_{vehi} \end{cases} \quad (1)$$

The connection between two vehicles is established using the distance proximity formulation. Here, the connection is established if the distance between the two vehicles d_{ij} is less than the transmission range of the vehicle nodes.

Using the existence or weight of edges, we want to divide V into k disjoint sets (V_1, \dots, V_k) . The related weights for m -uniform hypergraphs are archetypally called m -way affinities as each edge has distinct m vertices.

C. HGCM GENERATION MODEL

The cluster generation part of VANETs is discussed in this section. An urban scenario is considered for the simulation, and location information is shared with every neighbouring vehicle in the test case. A network hypergraph is constructed using that location information, and this section discusses the formulation of the vehicle hypergraph.

Our proposed model aims to cluster vehicles so that minimum bandwidth occupancy at any instant is achieved [38]. In the urban scenario, road congestion is unavoidable, leading to slow-moving traffic. The location, speed and vehicles in an area affect the stability of the clustering [39]. Each vehicle in a cluster is categorised as either CH or CM. Only one CH is conventionally allowed in a cluster, except in special cases of warlike fields [40]. In this article, the maximum possible vehicle density N is considered for the clustering algorithm. Using the location information, each vehicle finds its neighbour. A similarity matrix is generated, whose cell elements indicate the connection strength with another corresponding vehicle, as shown in Figure 3.

The adjacency matrix showcases the relation amongst the vehicle nodes. Spectral clustering using TTM is introduced in this subsection. The problem is partitioning the weighted hypergraph V into k disjoint sets, V_1, \dots, V_k , such that the total weight of edges within each cluster is high (dense connectivity amongst the vehicles), and the partitions are balanced [41]. The number of vehicular nodes connected to a node is defined as the degree of any node, $v \in V$, defining the total weight of edges v that is incident, i.e. $deg(v) = \sum_{e \in E: v \in e} w_e$.

Next, the volume is defined as $CM(V_1) = \sum_{v \in V_1} deg(v)$, which is the number of nodes incident on node V_1 , such that $V_1 \subseteq V$. The association amongst the edges contained within V_1 is defined as $assoc(V_1) = \sum w_e$. The normalised associativity of these individual partitions is given as

$$Par(V_1, \dots, V_k) = \sum_{i=1}^k \frac{assoc(V_i)}{CM(V_i)} \quad (2)$$

The adjacency matrix defined here is of the tensor (order z),

$$A_{i_1, i_2, \dots, i_z} = \begin{cases} w_{\{i_1, i_2, \dots, i_z\}} & \text{if } i_1, i_2, \dots, i_z \text{ are distinct} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The normalised associativity can be rewritten in terms of A and Y (inverse of the number of vehicles connected to a node) as

$$Par_{i \in \{1, \dots, k\}} = \frac{1}{z!} Trace(A \times_1 Y^{(1)T} \times_2 Y^{(2)T} \times_3 Y^{(3)T} \dots \times_z Y^{(z)T}) \quad (4)$$

where \times_l is the model- l product.

$Y^{(1)T}, Y^{(2)T}, Y^{(3)T}, \dots, Y^{(z)T} \in \mathbb{R}^{k \times z}$, which represents the number of CMs connected to each node v_i for each vertex. $Y^{i \in \{1, 2, \dots, z\}}$, as shown as follows:

$$Y^i = \frac{1}{\sum_k CM(V_i)} \quad (5)$$

The designed adjacency matrix [42] is considered for spectral clustering, and this hypergraph is now transverse to obtain the diagonal matrix (degree matrix) Dig that is the sum of runs over all the vehicle nodes that are one-hop adjacent to node v_i .

$$Dig_{ii} = \sum_{j=1}^N A_{ij} \quad (6)$$

For the Laplacian graph computation, this study utilises the unnormalised Laplacian matrix based on the Fiedler vector defined as

$$L = Dig^{-1/2} A Dig^{-1/2} \quad (7)$$

The top k eigenvector ($X = eig(L)$) is taken for k -means clustering that provides k -partitions of the VANET hypergraph structure. These partitions resemble the cluster formation in the vehicular network. They are further pruned to generate the optimal set of clusters for VANET maintenance. The pseudocode for the spectral clustering is listed in Algorithm 2.

Definition 2: The weighted hypergraph $\mathcal{H} = (V, \mathcal{E}, W)$, and its cluster is a tuple of $(C_{num}, C_{optimal})$, where $C_{num} = \{c_i | i = 1, 2, \dots, k\}$ is a cluster set, where k is the total number of clusters. $\forall c_i \in C_{num}, c_i = \{v_j | \tau \tau(v_j) = 0 \vee (\tau(v_j) = 1 \wedge \exists v_k \in N(v_j) \wedge \tau(v_k) = 0)\}$.

$C_{optimal}$ represents the optimal set of clusters, $C_{optimal} = [C_{num} : \forall \max(s)]$.

The clustering efficiency can be evaluated using the Calinski-Harabasz index (s) [43]. This index checks the closeness of vehicles in a cluster and the dispersion of all clusters by using Equation (8). The maximum value of s is the desired efficiency in the clustering.

$$s = \frac{tr(B_k)}{tr(Z_k)} \times \frac{Vehi_{num} - k}{k - 1} \quad (8)$$

Here, k represents the number of clusters, and each has the size of $Vehi_{num}$. $tr(B_k)$ is the dispersion amongst clusters, and

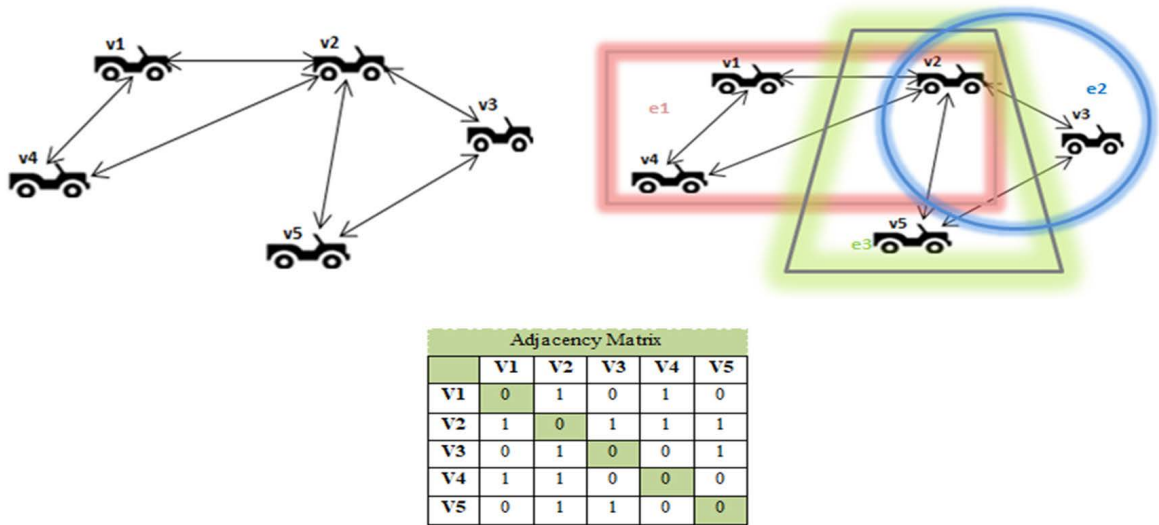


FIGURE 3. Hypergraph based VANET analysis.

$tr(z_k)$ is the dispersion amongst vehicles in a cluster. These two terms are calculated in Equations (9) and (10).

$$Z_k = \sum_{q=1}^k \sum_{x \in C_q} (x - c_q)(x - c_q)^T \quad (9)$$

$$B_k = \sum_{q=1}^k n_q(c_q - c_E)(c_q - c_E)^T \quad (10)$$

Here, c_q is a set of points in cluster q , and c_q is specifically the centre of the cluster. c_E is the centre of clusters with n_q points in them.

With the help of this index, an optimal set of clusters from the pool of formed clusters is selected on the basis of the maximum value of s .

D. RSUS DEPLOYMENT

The RSU is an integral part of the VANET. The VANET is a hierarchical architecture consisting of the main server, RSUs, and vehicles. RSU collects the data from the moving vehicles. The clustering of vehicles has been discussed in the context of RSU by many researchers as if the RSU has request congestion, then packet drop will increase. Therefore, the optimal number of RSUs has to be calculated so that maximum probable vehicles can be served without congestion. The optimal number of clusters has been calculated in the previous section. The RSU is placed at the centroid of the cluster. This way, a minimum number of RSUs can cover the maximum number of vehicles in the area. The installation cost would also be lesser (this is not evaluated in the simulation). Algorithm 3 tabulates the steps in locating the centroid for the RSUs and their deployment. The vehicles in any cluster cannot be controlled due to the nature of clustering. Consequently, few vehicles, such as three, can also lie in that precalculated cluster area. In such a case, the RSU is placed using a Gaussian probability distribution [44] as:

$$Vehi \sim N(\mu, \sigma^2) \quad (11)$$

Algorithm 2 Cluster Formation Using Hypergraph Theory

Input: Slot with maximum vehicles: N , Location of each vehicle: $Vehi_{Loc}$

1. Select the time slot t when maximum vehicles are present N
2. Form a hypergraph $H = (V, \mathcal{E}, W)$
3. Calculate the similarity matrix A on the basis of the distance proximity amongst the vehicles
A square matrix of size $N \times N$ of similarity (adjacency matrix) as $A = \sum_{i_3, \dots, i_N=1}^N A_{ij}$
4. A diagonal matrix $Dig \in \mathbb{R}^{N \times N}$ is with $Dig_{ii} = \sum_{j=1}^N A_{ij}$, and $L = Dig^{-1/2}ADig^{-1/2}$
5. Then, k dominant eigenvector of L is computed as $X \in \mathbb{R}^{N \times k}$
6. Normalise each row of $\bar{X} = X$
7. Run k-means on the rows of \bar{X}
8. Obtain C_{num} through k-means partition $Par = \{V_1, \dots, V_k\}$
9. For each cluster C_{num} , calculate the Calinski–Harabasz (s) criterion
10. Find the optimal cluster $C_{optimal} = [C_{num} : \forall \max (s)]$

Output: The optimal set of clusters: $C_{optimal}$

The distribution is defined as the mean ($\mu = 0$) and variance $\sigma = 1$.

$$f(Vehi) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(Vehi-\mu)^2 / 2\sigma^2} \quad (12)$$

A network graph $G = (V, \mathcal{E})$ amongst V vehicles' connections with \mathcal{E} set of edges. The centrality matrix for a graph is the measure of its compactness [45]. The centrality determines the most visited vertex in a graph. For a vehicle v ,

it can be calculated as

$$C_B(v) = \sum_{s \neq v \neq u \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (13)$$

Here, σ_{st} is the total number of the shortest paths from nodes to node u , and $\sigma_{st}(v)$ is the number of paths that pass through v . The vehicle with maximum centrality value is considered the cluster's centre as in (line 10). This is the location where RSU is to be installed.

Algorithm 3 RSU Deployment

Input: Set of optimal clusters $C_{optimal}$ and number of vehicles in each cluster $Vehi_{num}$.

1. **For** $i=1: C_{optimal}$
2. **If** $Vehi_{num} < 3$
3. Select a vehicle based $Vehi \sim N(0, 1)$ shown in Equation (12)
4. $RSU_{Loc} = Vehi_{Loc}$
5. **Else**
6. Construct an urban road map with a graph $G = (V, \mathcal{E})$
7. Edge is $e_{ij} \in \mathcal{E} = d_{ij}$ based on the distance amongst vehicles
8. Obtain a connection matrix

$$c_{ij} = \begin{cases} 1 & \text{if } d_{ij} \leq R_{vehi} \\ 0 & \text{otherwise} \end{cases}$$
9. Evaluate the betweenness centrality C_B by using Equation (13)
10. $Vehi = \max(C_B)$
11. $RSU_{Loc} = Vehi_{Loc}$
12. **End**
13. **End**

Output: RSU location RSU_{Loc} .

IV. CH SELECTION AND CLUSTER MAINTENANCE FOR HGCM

A. CH SELECTION

The vehicle node v_i in the network at time t has features $f_i(t) = \{\vec{s}, \vec{p}, a, \varphi, Vehi_{ID}, \eta\}$, where \vec{s} is the vehicle speed, \vec{p} is the location of each vehicle in both coordinates (x, y) , a is the acceleration, and φ is the vehicle direction. Each vehicle is assigned a unique identity $Vehi_{ID}$, and η refers to the one-hop neighbours of vehicle node v_i . Out of these nodes, a vehicle is selected as the CH. In this article, the CH selection metric $m_i(t)$ is a collection of metrics $\{\psi_{vehi}, \eta, \mathcal{E}, t\}$. The CH selection process is dependent upon the current CH selection metrics of each CM in the cluster. ψ_{vehi} is the relative speed, η is the set of neighbours of vehicle v_i , \mathcal{E} is the eccentricity, and t is the trust calculated via spectrum sensing. The selected CH should have a maximum of $\sum_{i=1,2..n} m_i(t)$ at any instant t . Given that the hypergraph network is the cooperative network, each vehicle's feature is relative to every hyperedge linked to that hypernode [46].

The novel contribution in CH selection parameters is the strength of the cooperative nature of the hypergraph. The use of deep learning in the calculation of the trust score of each vehicle is another novel contribution to the CH selection.

The CH collects all the information from the network, sends it to the RSU and maintains the communication between the cluster vehicles and RSU. The stability of the CH will be higher if it will be in a communication link with the neighbour vehicles for a longer time.

The definitions of each metric with the essential background are presented below.

1) RELATIVE SPEED SCORE (ψ_{vehi})

Definition 3: A vehicular node's v_j relative speed score ψ_{vehi} is a score that either penalises or gives reward to a vehicle if it crosses a cluster's average speed or aligns with the cluster. A high score of ψ_{vehi} indicates a high probability of election.

This parameter determines how close a vehicle's speed is to its neighbour's. The relative speed of each vehicle is calculated by differentiating its speed from the cluster's average speed at any instant of time. The moving direction of the vehicles also comes into play this way. The more vehicles are moving in the same direction, the higher ψ_{vehi} will be. The relative speed score is evaluated as shown in equation (14) [21]. The relative speed is compared with a threshold speed s_{thr} . If a vehicle is moving at higher speed than s_{thr} , its ψ_{vehi} gets penalised with δ ; else, a reward of δ is added to its score.

$$\begin{aligned} \psi_{vehi}(t+1) &= \psi_{vehi}(t) + \delta; & |V_{vehi} - V_{avg}| \leq S_{thr} \\ \psi_{vehi}(t+1) &= \psi_{vehi}(t) - \delta; & |V_{vehi} - V_{avg}| > S_{thr} \end{aligned} \quad (14)$$

δ and S_{thr} are 0.01 and 2.77 for this work, respectively.

2) NEIGHBOURHOOD DEGREE (η)

Definition 4: The connection status between the two vehicular nodes v_i and v_j at time t in the cluster formed $C_{optimal}$ with vehicle density $vehi_{num}$ is defined as

$$\eta = \sum_{j=1}^{Vehi_{num}} c_{ij}; \quad \forall i \quad (15)$$

High η ensures that the CH will not be dynamic for a long time. The neighbourhood degree defines the total number of vehicles in the vicinity. The vehicles under the transmission range of OBU are considered neighbours. c_{ij} is 1 if the distance between two vehicles at the time stamp t is less than R_{vehi} [47]. A negative correlation exists between the distance and transmission range. That is, if two vehicles are close to each other, then a more reliable connection is bound.

3) ECCENTRICITY (\mathcal{E})

Definition 5: Let A be a fundamental matrix. Then, an eigenvector $X > 0$ exists, such that $AX = \lambda_1 X$, $\lambda_1 > 0$ is an eigenvalue of an immense magnitude of A , the eigenspace associated with λ_1 is one-dimensional, and X is the only

non-negative eigenvector of A . \mathfrak{E} is the average of top k eigenvalues of A designed for $\mathcal{H} = (V, \mathcal{E},)$.

In real time, communication links break frequently due to vehicles' high speed. to maintain a link, a requisite is set for a progressing cluster model. Usually, reclustering will become inevitable once the CH resigns or loses its suitability to continue as a CH. to ensure stability, the concept of eccentricity (\mathfrak{E}) is introduced. Here, an evolving graph-based model is designed through spectral clustering [28]. A vehicular graph topology is intended to be hypergraph $\mathcal{H} = (V, \mathcal{E}, W)$ with the usual procedure as defined in section III. the adjacency matrix A is generated on the basis of the distance proximity amongst the vehicles present at each time instant t for each cluster. The eigenvalues for a vehicle i in each group are λ_i , where $i = \{1, 2, \dots, vehi_{num}\}$. Lastly, \mathfrak{E} is the mean/average eigenscore of each vehicle calculated as [28]

$$\mathfrak{E} = \frac{1}{|Vehi_{num}|} \sum_{\lambda_i \in Vehi_{num}} \lambda_i \quad (16)$$

The maximum value of \mathfrak{E} ensures a stable CH selection designed in accordance with hypergraph theory.

4) TRUST SCORE (\mathfrak{t})

Definition 6: Through channel h , the signal is received from the user, and the probability of detection is 1. Then, the user is primary (PU); else, it is a secondary user (SU).

The vehicular network may also have some VIP and emergency vehicles, which are regarded as the PUs of the communication spectrum in the network. Others are assigned as SUs. Every vehicle takes part in spectrum sensing. Given that the communication spectrum is limited, the cognitive spectrum sensing approach is used in the communication model [48]. Once the PU is detected in the network, the SU will have to vacate the spectrum for it. The SU following this protocol gains the trust, and the trust score \mathfrak{t} is increased. The model of cognitive spectrum sensing is introduced in this work to select the most trustworthy vehicle as the CH.

The SU senses the PU presence by comparing the signal energy of neighbouring vehicles with the probabilistic threshold value. The detected energy signal (test statistic) can be presented in a complex form as

$$T(Y) = \frac{1}{2N} \sum_{i=1}^N |Y_i^{re} + Y_i^{im}|^2 \quad (17)$$

where $T(Y)$ is the test statistics received on any vehicle. This $T(Y)$ is a random variable and can be estimated using the chi-square probability distribution function as

$$P_d(\varepsilon, \mathfrak{t}) = \mathbb{Q} \left(\left(\frac{\varepsilon}{\sigma_u^2} - \gamma - 1 \right) \sqrt{\frac{t f_s}{2\gamma + 1}} \right) \quad (18)$$

The vehicle is detected as the PU if the probability of detection P_d is greater than threshold ε [49]. ε is calculated by the inverse of this chi-pdf:

$$\mathbb{Q}^{-1}(P_d) = \left(\frac{\varepsilon}{\sigma_u^2} - \gamma - 1 \right) \sqrt{\frac{t f_s}{2\gamma + 1}} \quad (19)$$

Here, $\mathbb{Q}(\cdot)$ is the complementary distribution function and is Gaussian in nature, i.e.

$$\mathbb{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt \quad (20)$$

The threshold value decides the accuracy of detection of the PU. In this work, we follow the concept of deep learning to detect the presence of PU. It has proposed the stack of deep learning layers with LSTM in the focus to factor down the signal features. Threshold value estimate is comprised of two distinct stages: data generation and deep learning model training.

The spectrum sensing network is simulated in ideal conditions to generate the training data with various modulation schemes and random input data streams. Simulated modulations are BPSK, QPSK, 8-PSK and 16-PSK. With every simulation, the generated signals' energy is mapped with the results of the PU detection with a threshold value calculated using Equation (17). As a result, forty thousand samples are used to create a labeled dataset. The PU and non-PU labels are assigned to detected signals.

The LSTM network is trained on the data to teach the decision based on sensed signal energy. The network is trained with the randomly sampled 90% data for training and 10% for the testing. Two biLSTM layers with forward and backward data sequencing are used which are connected with the fully connected layer. On training, the network is able to correctly classify the absence of any PU upto 89% whereas any PU is correctly detected upto 83.5%.

The trained network is used to obtain the threshold value for $P_d(\varepsilon, \mathfrak{t})$ on the unknown signals. For every successful detection, the trust score (\mathfrak{t}) is incremented. The higher \mathfrak{t} is, the higher the probability of a vehicle to be elected as CH will be. This trust score calculation scheme is portrayed in Figure 4.

The model is divided into three subparts: sensing block, training block and PU detection block. The energy signal database is collected by simulating the network in the ideal and Rayleigh noisy channel environment. The data are fed into the LSTM training block. After training, the detected energy signal is tested with the trained model, and the vehicle is assigned to the PU or SU. The true detection increases the trust score by 1.

A comparison of the ROC curve between the theoretical analysis of threshold calculation by Equation (19) as in [50] and the proposed LSTM-trained network is presented in Figure 5. The detection probability P_d is high for a small value of P_f . The trained LSTM network performs better in a noisy environment, which means that it is efficient and can predict with appropriate accuracy.

The complete algorithm designed for CH selection is shown in Algorithm 4. All four parameters are summed and integrated to select a stable CH for a long period of time. The vehicles are firstly clustered using Algorithm 2. For all members at each cluster, the four parameters are found. Then, CH_{score} is calculated to select the stable CH.

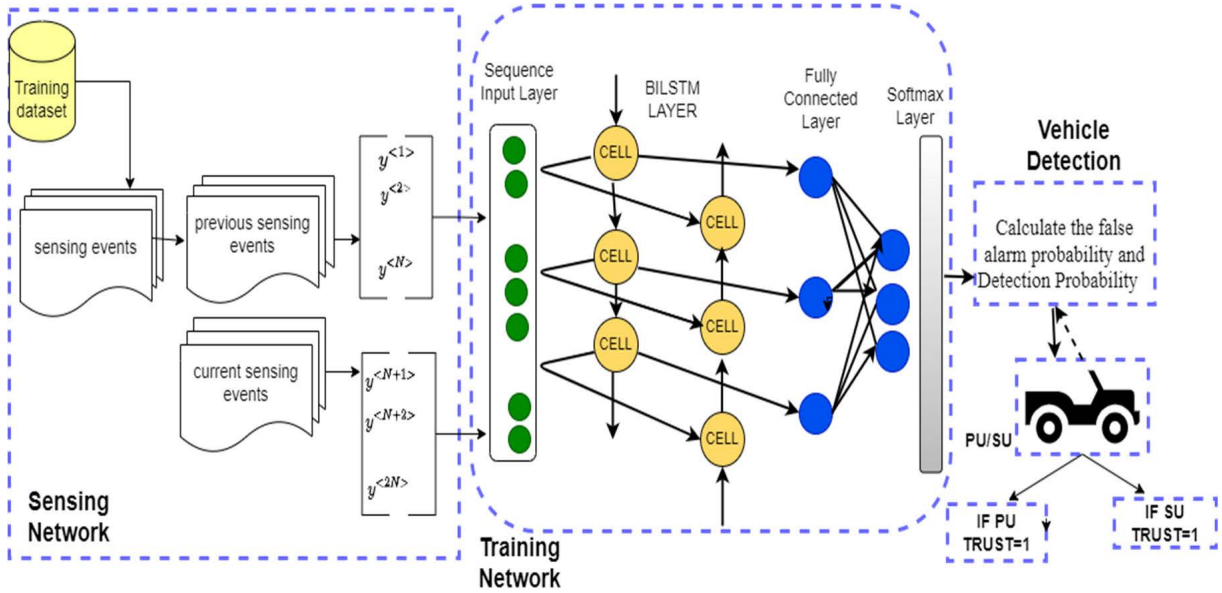


FIGURE 4. Adaptive spectrum sensing model using LSTM.

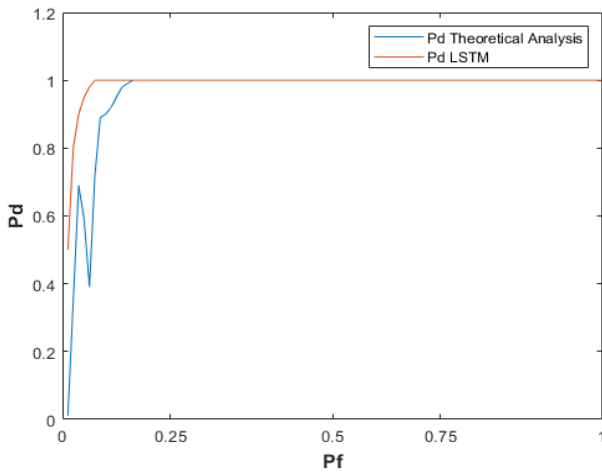


FIGURE 5. The P_d and P_f using LSTM method and theoretical analysis.

B. HGCM MAINTENANCE PHASE

The reduction of communication overhead after the selection of CH is also an important part of the designed algorithm. The cluster maintenance process ensures strong connectivity and stable link lifetime through CH. In this work, the joining of a new vehicle in a cluster and leaving of any CM are considered vital for the cluster maintenance phase. HGCM is designed with the parameters that ingest the restructuring of the topology and vehicular speed. The CH score provides a significant contribution to capturing the information of vehicle movement. The maintenance does not deal with the networking. It is designed to maintain a smooth transition of vehicles in and out over time.

1) CLUSTER ENROLMENT

Cluster formation is performed on the basis of the proximity of vehicles concerning the transmission range of RSU

Algorithm 4 CH Selection

Input: Relative speed V_{vehi} and location $[x, y]$; Number of lanes in a map (no_lane); $time_span$; Cluster member (CM); (N) no. of clusters.

1. **For** $t=1:time_span$
2. **For** $i = 1 : N$
3. **For** $j = 1 : CM$
4. Calculate ψ_{vehi} from Equation (14)
5. Find the neighbouring vehicles and calculate η by using a connection matrix Equation (15)
6. Calculate the maximum eigenvalues λ ;
7. Then obtain \mathfrak{E} by using Equation (16)
8. Signals' energy is mapped with the results of the PU detection with a threshold value calculated using Equation (17) and Equation (19).
9. Calculate t score using the LSTM trained network based on sensed signal energy.
10. Find CH_{score} for each CM
11. $CH_{score} = \psi_{vehi} + \eta + \mathfrak{E} + t$
12. **End**
13. $CH_j = Max \sum_{j=1}^k CH_{score} \forall j$
14. **End**
15. **End**

Output: CH

deployed. A small number of vehicles in a cluster with a large transmission range will lead to inconsiderably reliable networking. The selected CH starts its task by sending the polling signals and if it receives any signal in return within a stipulated time period $time_span$ with the condition that $dist_{(vehi, ch)} < r_{ch}$. A new vehicle is assigned to that cluster (each formed under RSU) and becomes CM of that particular cluster. CH will update the local database and the list of

vehicles in RSU. The arrival of a new vehicle in a cluster can also trigger CH reselection in the worst-case scenario. The algorithm designed is thus coined to formulate CH score based on four factors $\{\psi_{\text{vehi}}, \eta, \epsilon, t\}$; with this, the stability of CH is ensured.

2) CLUSTER LEAVING

CMs can leave any cluster at any moment of time. The reasons for this could be lane change or exit from a road, the ever-changing dynamics of vehicles and the topology that affects the number of CMs. Therefore, a frequent polling of signals is done between the established members and CH. If a CM does not reply within the speculated period time_span , then the CM is considered to be disconnected and leave the cluster. The CH removes the recorded vehicle, and an updated list is appended at the RSU. The complete algorithm designed for cluster maintenance is shown in Algorithm 5.

Algorithm 5 Cluster Maintenance

Input: CH; Cluster C; No. of vehicles N; CH transmission range R_{CH}

1. **For** $t=1:\text{time_span}$
2. **For** $j = 1 : C$
3. **For** $i = 1 : N$
4. **IF** $\text{dist}(\text{vehi}, CH) < R_{CH}$
5. $\text{CM}=\text{CM}+1$;
6. CH < Adds the vehicle in the RSU List&LocalList >
7. **Else**
8. $\text{CM}=\text{CM}-1$
9. CH < Removes the vehicle from the RSU List&LocalList >
10. **End**
11. **End**
12. **End**
13. **End**

Output: Cluster Update

3) TIME COMPLEXITY OF THE HGCM SCHEME

Cluster formation, RSU deployment and CH selection are the key components of the proposed algorithm. Accordingly, the algorithm's total time complexity is expressed as

$$O_{TOT} = O_{CF} + O_{RSU} + O_{CH} \quad (21)$$

where O_{CF} is the time complexity of cluster formation, O_{RSU} is for RSU deployment, and O_{CH} is for CH selection. The cluster is generated by hypergraph partitioning. The major steps involved are as follows: (1) hypergraph construction= (V, E, W) ; (2) Laplacian construction; (3) eigen problem solving; (4) vertex eigenvector computation; (5) performing k-means on \bar{X} .

In a hypergraph with pairwise similarity, the cost to construct the nearest neighbour graph is $O(N^2)d$, given that it requires d -dimensional similarity computation for each vertex pair, where N is the maximum number of vehicles in the worst-case analysis with m hyperedges. The Laplacian

construction step is directly correlated with the sparsity of adjacency matrix A through the non-zero elements NNZ (or the number of vehicles in our case). $L = NNZ(A^2)$, the eigen complexity is $E_C = O(N^3)$, and the last is the k-means complexity which is dependent on $O_{C_{optimal}} = \tau NC_{num}$, where τ is the number of iterations.

$$O_{CF} = O(N^2)d + NNZ(A^2) + O(N^3) + \tau NC_{num} \quad (22)$$

This can be reduced after removing the terms of less computational power as

$$O_{CF} = O(N^2)d + O(N^3) \quad (23)$$

The RSU deployment is done using a graph, so the computational complexity is

$$O_{RSU} = O(N^2)d \quad (24)$$

In this article, the CH selection metric $m_i(t)$ is a collection of metrics $\{\psi_{\text{vehi}}, \eta, \epsilon, t\}$. ψ_{vehi} is the relative speed, η is the set of neighbours of vehicle v_i , ϵ is the eccentricity, and t is the trust calculated via spectrum sensing

$$O_{CH} = O_{\psi_{\text{vehi}}} + O_{\eta} + O_{\epsilon} + O_t \quad (25)$$

The relative speed is a simple threshold function done on the basis of the vehicle speed; thus,

$$O_{\psi_{\text{vehi}}} = N \quad (26)$$

The next is the neighbourhood, which is a function of the c_{ij} affinity matrix for nearby vehicles.

$$O_{\eta} = O\{\log N^2\} \quad (27)$$

The eccentricity is calculated using spectral clustering methods that include the affinity matrix and eigenvalue decomposition. The entire spectral clustering complexity is

$$O_{\epsilon} = O\{N^2\}d + O\{N^3\} \quad (28)$$

The last factor is trust LSTM, which plays the primary role in this for spectrum sensing; the theoretical time complexity of the LSTM is given as

$$O_{LSTM} = 4IH + 4H^2 + 3H + HK \quad (29)$$

where I represents the number of inputs, K represents the number of outputs, and H represents the number of hidden layers. In this study, because the model is trained only once for a given vehicle signal, the LSTM detects whether the vehicle is a primary or secondary user through spectrum sensing. Thus, the time complexity condenses to

$$O_{LSTM} = O\{4H^2\} \quad (30)$$

Then, the complete time complexity is reduced to moving all the terms with less complexity than cubic and quadratic terms, as shown below:

$$O_{CH} = O\{\log N^2\} + O\{N^3\} \quad (31)$$

TABLE 3. Simulation parameters.

Parameter	Value
Scenario	Urban
Vehicle density	100–1000
Lanes	458 lanes
Vehicle average speed	10.71 m/s
R_{RSU}	350 m
Maximum Simulation time (t)	1500 s
R_{vehi}	200 m
Length of vehicle	5 and 10 m
Maximum vehicle speed	25 m/s

The overall complexity is primarily dependent on the hypergraph, i.e.

$$O_{TOT} = 2O(N^2)d + O\{\log N^2\} + 2O\{N^3\} \quad (32)$$

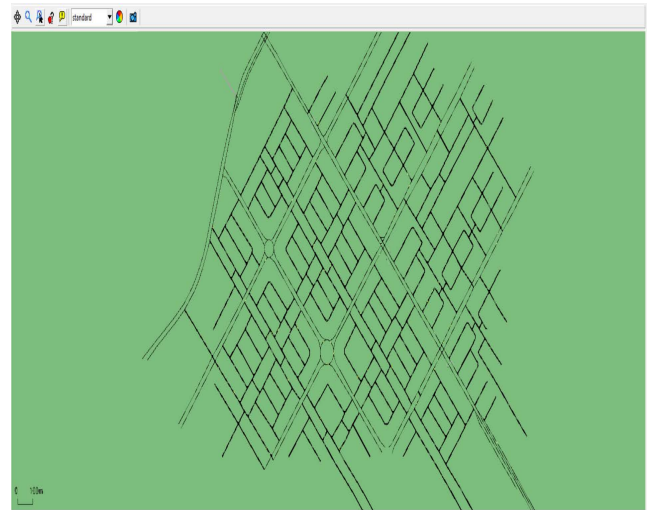
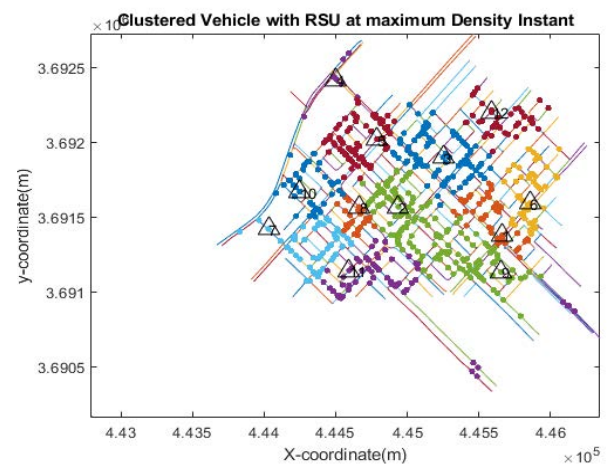
V. SIMULATION AND PERFORMANCE EVALUATION

This section describes the detailed background of the simulation tools and the various evaluation parameters utilised. The results' discussion is carried out in three phases: the effect of different traffic densities on the stability of the designed HGCM, state-of-the-art comparison, and the effect of different traffic densities on the routing performance in comparison with various existing algorithms.

A. SIMULATION TOOLS

The simulation is implemented using MATLAB (R2018b), with the processor Intel®Core™i7, 1.98 GHz, Simulation of Urban Mobility [51] (SUMO 1.7.0) and Traffic Control Interface (TraCI) [52]. SUMO is an open-source microscopic road traffic simulator licensed under the General Public License. It was developed through collaboration between the Centre for Applied Informatics, Cologne (ZAIK) and the Institute of Transportation Systems at the German Aerospace Center (DLR). TraCI is an API developed to interface SUMO with other coding platforms, such as MATLAB, Python and Java. The complete network design and the routing performance for metrics, such as throughput and PD are evaluated in MATLAB. The area considered for the study is a crowded market area of Baghdad having the latitude = 33.3573° S, 33.3730° N and longitude = 44.3960° W, 44.4190° E; it is extracted from Open Street Map. The latitudes and longitudes make a simulation area of size 1000 * 2000 m². Baghdad is the capital of Iraq and one of the largest cities in the Arab world, with massive population and a geographical area of 204 km². The traffic environment summary is provided in Table 3. These parameters are considered on the basis of extensive literature survey. Moreover, the values are minutely crafted to portray a real urban scenario with congestion, many crossroads and a large number of vehicles during peak hours [53].

The geographical region considered for the simulation is shown in Figure 6; it is a vast area with urban infrastructure. Algorithm 2 suggests the optimal number of clusters in that

**FIGURE 6.** Simulated part of Baghdad real map in SUMO.**FIGURE 7.** Cluster formation along with RSU deployment at an instant with maximum vehicular density.

region, and RSUs are deployed using Algorithm 2. Further vehicle features $f_i(t) = \{\vec{s}, \vec{p}, a, \varphi, vehi_{id}, \eta\}$ are recorded for 1000 vehicles. The number of vehicles in the simulation area varies as in real-world scenarios. Twelve clusters are optimally selected using Algorithm 2. Different colours portray each cluster and vehicles in each cluster. The triangles (in black) are the different RSUs placed, which will serve as the cluster centre (providing auxiliary facilities), Figure 7.

B. EVALUATION METRICS

The designed HGCM is also tested in terms of routing performance. The communication amongst vehicles is modelled through the Rayleigh fading channel with BPSK modulation. Owing to the vehicles' movement, the network is dynamic and fast, which introduces a Doppler effect. The effect is incorporated as the signal fades over time. The communication network parameters are listed in Table 4 [53].

Different metrics are calculated to evaluate the stability and performance of our HGCM. These metrics are as follows:

TABLE 4. Integrated network parameters for IEEE 802.11p.

Parameters	Value
Fading channel	Rayleigh Channel
Fading Parameter	Filtered Gaussian noise
Doppler shift	$V_{s,r}/\lambda$
Sample rate	10 MHz
Number of FFT points	32
Number of subcarriers	52+ D
Number of data subcarriers	52
Number of pilot subcarriers	4
Cyclic prefix	16
Modulation scheme	BPSK
Coding scheme	$1/2$
Available data rate	14Mbps

1. Cluster number: It is the number of clusters constructed during network operation. Few cluster numbers enhance the clustering algorithm efficiency [54].
2. CH Stability: CH stability is evaluated as the most prolonged interval to be crowned as CH. Mode is a statistical function that gives the interval counts for a vehicle as CH.

$$CH \text{ Stability} = Mode\left(\sum_{i=1}^t Vehi_{ID}\right) \quad (33)$$

3. CM Lifetime: It is the time interval from a vehicle joining a cluster to leaving it or changing its state to a CH.
4. CH Change Rate: It is the average number of CH changes per unit time. A stable cluster should have lower CH change rate.
5. Packet delay (PD): It is the time taken for a packet to be sent from the source to the destination vehicle. The delay in packet delivery depends on the noise, network congestion and hop travel distance. It is measured in seconds [54].

$$PD = \frac{\sum_{i=1}^N (Time \ a \ packet \ is \ received - Time \ a \ packet \ is \ transmitted)}{Number \ of \ packets} \quad (34)$$

6. Throughput: It is the number of bits successfully transmitted from the source to the destination vehicle at a given time period. It is measured in kilobits per second. High throughput can be achieved with high network stability and minimal hop count [54].

$$Throughput = \frac{(Packets \ Received \times Packet \ size)}{Total \ Time} \quad (35)$$

C. RESULTS AND DISCUSSION

The designed HGCM is analysed on a real map in an urban scenario where different densities of vehicles at various mobilities are infused into the network. The number of clusters produced throughout time affects algorithm efficiency as well. In comparison with roads, mobility is less of an issue in cities. CH is also more stable in an urban setting, where the vehicle density is higher, but the mobility is lower. The results' discussion in this section is parallel with the state-of-the-art schemes.

1) EFFECT OF DIFFERENT TRAFFIC DENSITIES ON HGCM STABILITY

The work presented in [37] by Maoli et al. opened up a way to present VANET as a hypergraph, although the authors discussed that in the context of fog computing and left the discussion gap on the network performance parameters.

The effectiveness of the designed algorithm was also gazed by the number of clusters formed over time. These numbers allow us to evaluate the quality of the formed clusters. Few clusters with vehicles having low mobility achieve efficient connection and stable clustering. On the contrary, more clusters eventually lead to high overhead and mergers. The average number of vehicles in a cluster represents the cluster size. The larger the cluster size is, the higher the clustering efficiency will be. Figure 8, shows the average number of vehicles in a cluster and the number of clusters generated at different vehicular densities for our HGCM.

In HGCM, four and 12 clusters are generated with an average of 25 and 210 vehicles in a cluster at low and high traffic, respectively.

For spectral clustering, this study employs the eigenvalues derived using VANET's hypergraph presentation. The idea is motivated by the connectivity graph eigenvalues in [28] and [33]. Both works in [28] and [33] were designed for the highway scenario, whereas our work is designed for the urban environment. The eccentricity parameter in our work is inspired by the connectivity-based CH selection in [28]. A high connectivity with vehicles represents that dense traffic and maximum CH can be connected with the maximum number of vehicles. Eccentricity is a positional parameter that can be correlated with the connectivity issue. In a graph network, the central point has the highest connectivity, as does in the hypergraph. The neighbourhood degree is another connectivity parameter. In CH selection, a relative vehicle speed denotes uniform cluster generation. The CH stability using these three parameters $\{\psi_{vehi}, \eta, \mathcal{E}\}$ is evaluated on different vehicle densities in the same network and represented in Figure 9. Given that the vehicle deployment and movement are random and near to a real environment in SUMO, $\{\psi_{vehi}, \eta, \mathcal{E}\}$ parameters are not able to conclude any concrete pattern. We hereby use a nontrivial CH selection parameter, i.e. trust score t . The trust score t , along with the remaining three CH selection parameters, improves CH stability. The novel set of CH selection parameters significantly improves the CH stability by 20% at all vehicle densities as shown in Figure 9.

Regardless of the non-uniform pattern in improving stability by the proposed set of parameters, the novel contribution shows a constant improvement compared with each parameter, as shown in Figure 10. The method designed using eccentricity only provides satisfactory stability compared with the others. This is because the network has a dynamic structure that is perfectly emulated utilising the hypergraph concept. By contrast, the rest of the parameters, such as the relative speed and neighbours, could not trace the stability

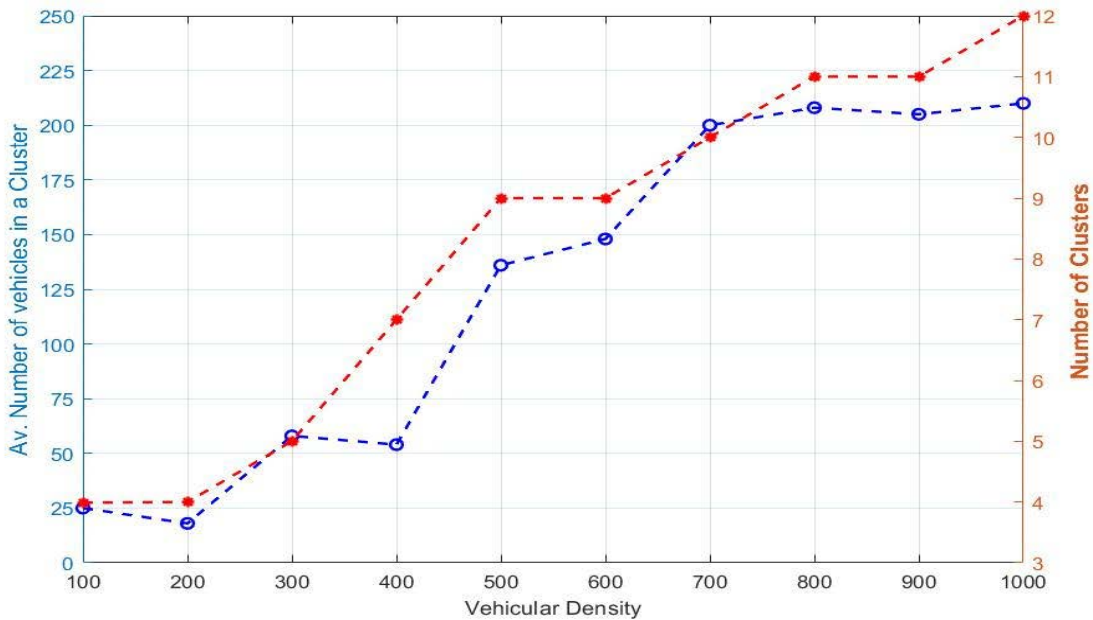


FIGURE 8. Av. number of vehicles in a cluster and cluster number of HGCM at different vehicle densities.

with increasing vehicle densities. The contribution of each at an individual level is low, but the stability provided is best when they are combined $\{\psi_{veh}, \eta, \mathcal{E}, \mathcal{L}\}$.

The efficient cluster generation in the proposed scheme leads to enhanced CH stability. The CH stability has already been validated in Figure 9 and 10. The proposed HGCM with CH selection parameters achieves 72% and 53% of stability for low and high traffic density respectively.

The CH achieves enhanced stability, as evaluated in Figure 10. However, other vehicles in the cluster are marked as CMs. The increased lifetime of CMs indicates efficient clustering by using a hypergraph. In the case of non-uniform clustering, a CM leaves clusters frequently and joins others. In such a scenario, CH stability cannot be achieved. Figure 11 presents a comparison of CM lifetime of our proposed HGCM. The HGCM scheme gains the highest lifetime compared with its counterparts, although the traffic congestion with the increase in vehicle' density imposes performance degradation. Nevertheless, it can be ignored because for a 10-fold increase in traffic from 100 vehicles to 1000 vehicles in the network, the CM lifetime decreases to 4.2% only.

Also, the lower change rate of the CH, the more stable the cluster structure. From Figure 12 we can see, that the CH change rate is the lowest due to hypergraph spectral clustered network with the CH selected considering the four selection metrics $\{\psi_{veh}, \eta, \mathcal{E}, \mathcal{L}\}$.

The cumulative multimetric reduces the overhead that occurs due to the frequent shifting of the CH from one vehicle to another. Thus, it improves CH stability and CM lifetime and reduces CH change rate in a comparison with individual metrics.

2) STATE-OF-THE-ART COMPARISON

A comparison of HGCM with some algorithms presented in the literature in terms of cluster number and CH stability is tabulated in Table 5.

Cluster-based VANET oriented Evolving Graph (CVoEG) [28] was introduced by Khan *et al.* They used a graph spectral clustering algorithm and tested it on a highway network. CVoEG [28] forms 20 clusters under a low traffic density, covering a road length of 12 km under the i-5 highway analysis of the California environment. It is expected to achieve 65.5% stability. In this study, the speed of vehicles is used to emulate graph edges. Thus, at low variance, as the speed of vehicles is nearly identical, the eigenvalues are almost the same, which eventually leads to low cluster formation.

Another work was presented by Khan *et al.* in [33]; it is the nearest peer to our work in this paper. The authors used connectivity-based CH selection and calculated the eccentricity for it. The highest eccentric vehicle in a cluster was assigned as the CH. In HGCM, the CH selection depends on a novel set of vehicle parameters. RSUs are installed at equal distances, and dynamic clusters are formed with traffic density. A 2 km road was simulated in SUMO by Khan *et al.* [33]. A maximum of 140 clusters was generated for high traffic and 20 for low traffic in 400-s simulation. The massive clusters in the network reduce the CH stability, as discussed previously in this section. The authors in [33] did not evaluate the CH stability, such that Table 5 lacks that.

In the article proposed by M. Mukhtaruzzama [55], clusters were generated by considering the moving direction of a vehicle at the junctions, vehicles' density, and transmission range. The CH was selected by relative position and time spent on the road. With 100 vehicles as a testbed, the CH stability was

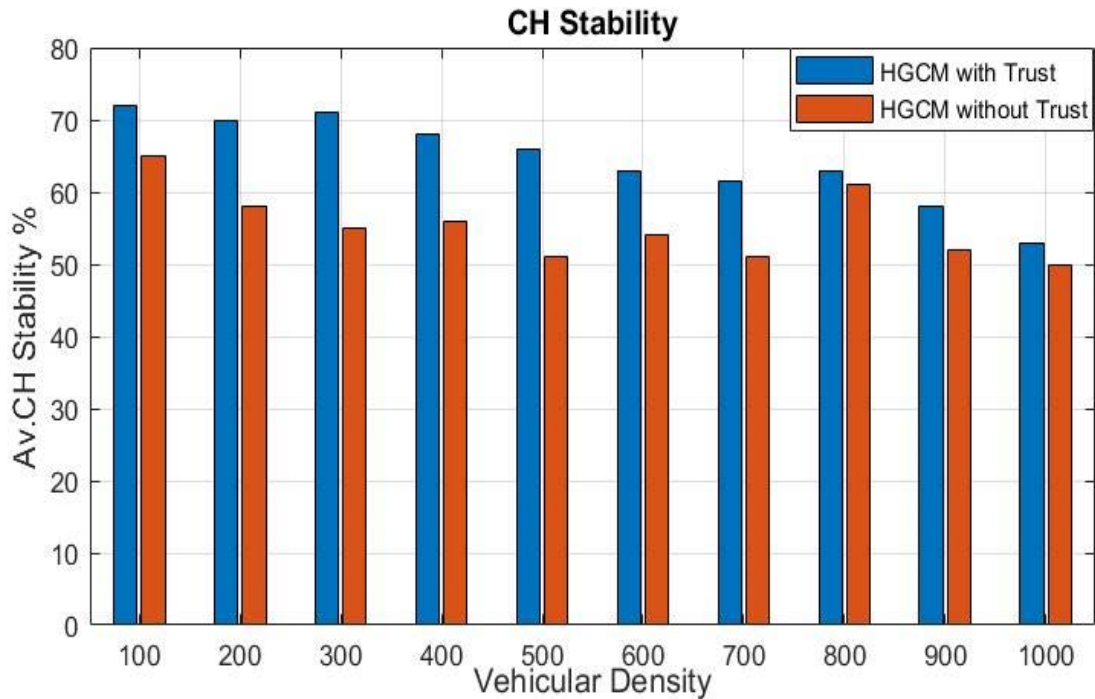


FIGURE 9. Average stability of CH for HGCM with and without trust factor.

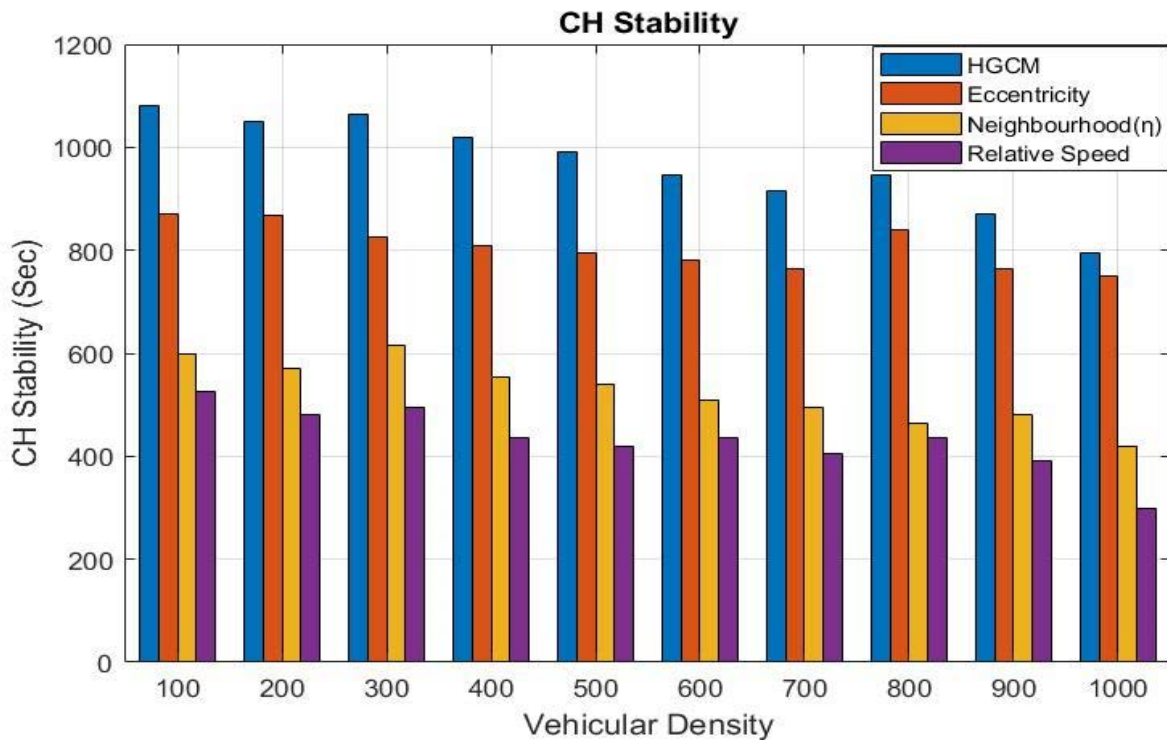


FIGURE 10. Stability of CH at different vehicle densities.

found to be 76% with a formation of 16 clusters. By contrast, the HGCM system suggested in our work procures 72% with 4 clusters.

In the method proposed by Arkian et al. [21], a high number of dynamic clusters are projected with a low

variance of 90 vehicles only for a highway length of 3000 m with two-lane analysis. This method incorporates neighbourhood analysis; thus, when only 90 vehicles exist, to cover all vehicles in a sparse area, the number of clusters is bound to be high. for the low traffic

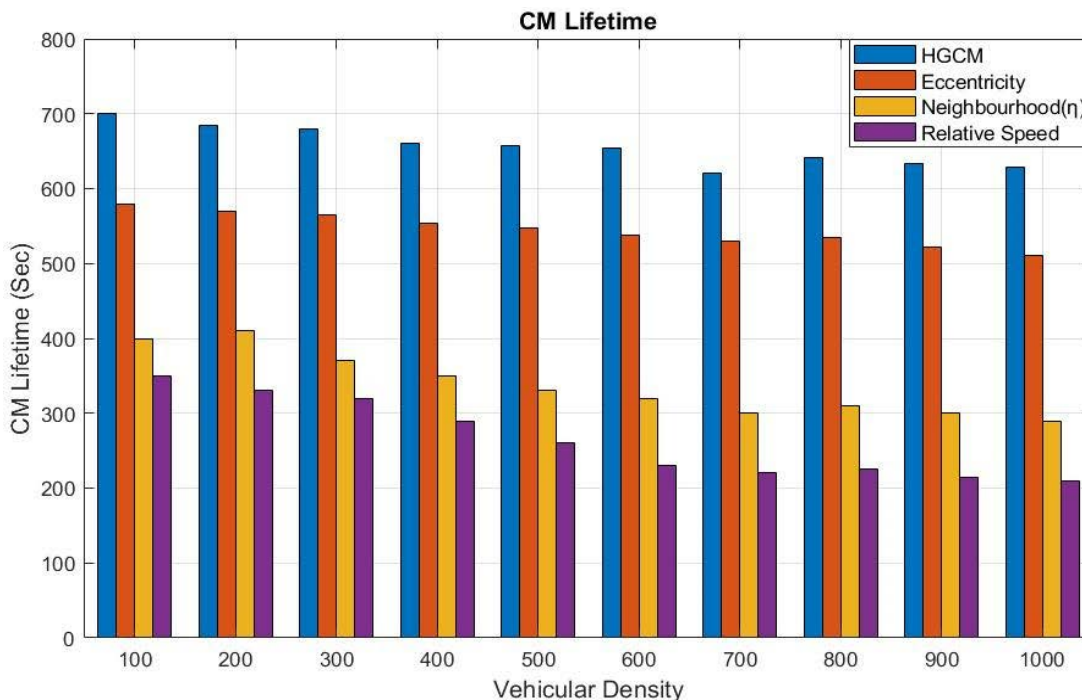


FIGURE 11. CM lifetime at different vehicle densities.

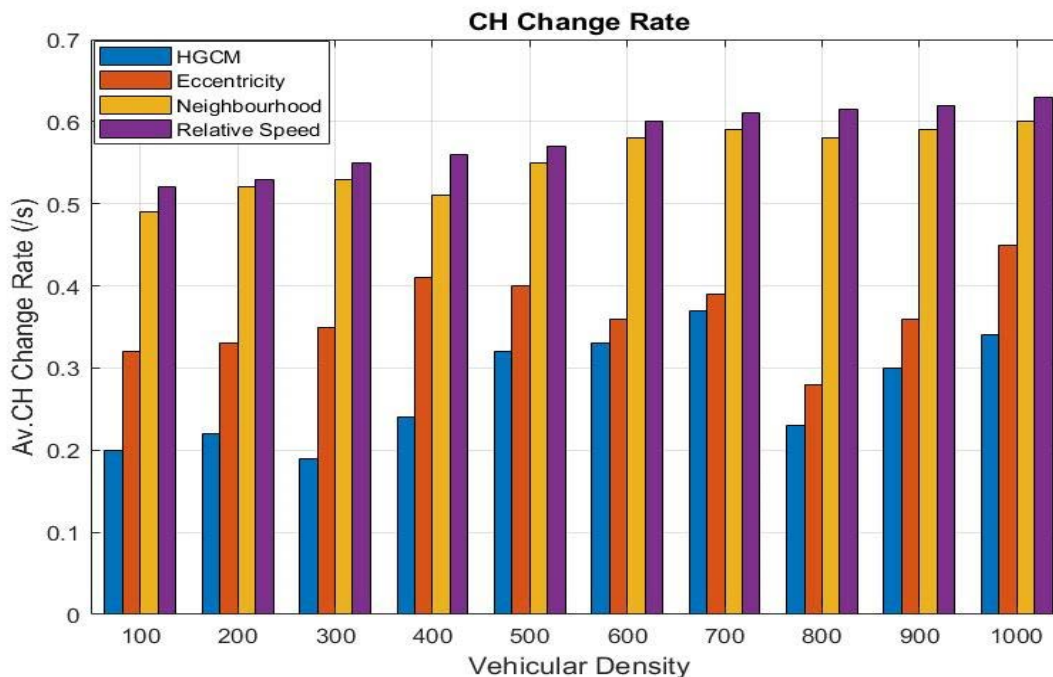


FIGURE 12. CH change rate at different vehicle densities.

flow with a high number of clusters, the CH stability is 58%.

Although the CH stability of the method proposed by Mukhtaruzzaman et al. [55] is 76%, we can conclude that our HGCM using hypergraph theory improves the clustering efficiency compared with other algorithms in terms of the

number of clusters constructed with 72% of CH stability. The CH stability for the maximum number of clusters formed is reported in Table 5.

The graph in Figure 13 is plotted for the CH stability for various vehicle densities. The CH stability decreases with the increase of traffic density. On the same network

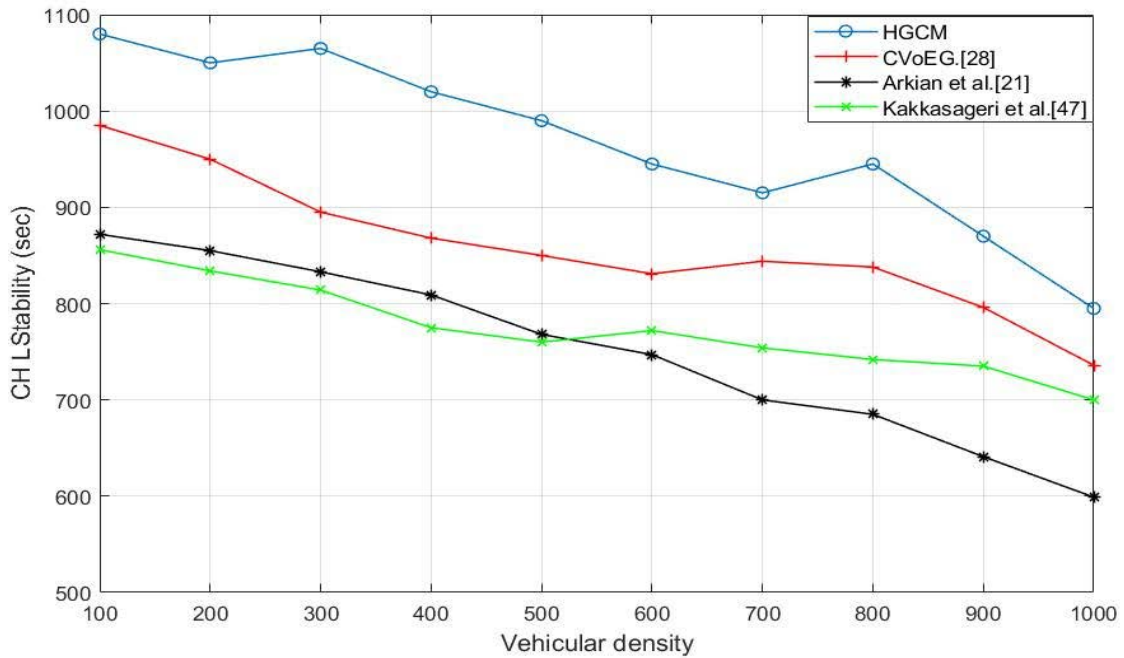


FIGURE 13. Stability of CH at different vehicle densities.

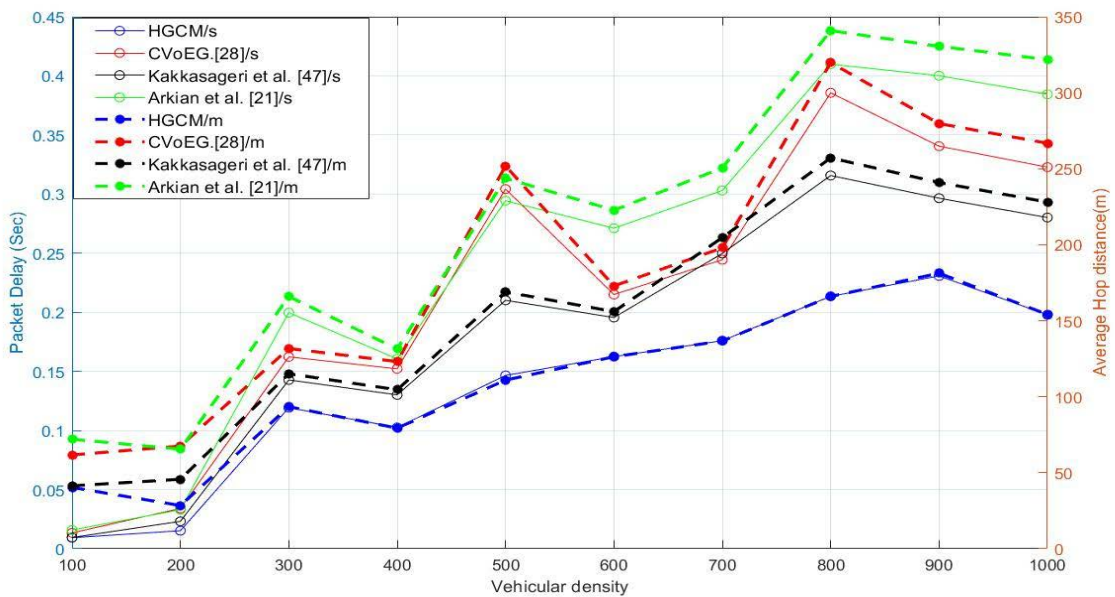


FIGURE 14. Analysis of PD and Hop distance at different vehicle densities.

TABLE 5. Comparison with different algorithms at low traffic density.

Algorithm	No.of Clusters	CH stability
HGCM (proposed)	4	72%
CVoEG.[28]	20	65.5%
Khan et al.[33]	20	-
JCV.[55]	16	76%
Arkian et al.[21]	55	58%

conditions listed in Tables 3 and 4, the CH stability is also evaluated using the algorithms in articles [28], [21], and [47], with the same vehicle properties recorded from SUMO as for the proposed work. In Figure 13, CVoEG [28] seems

to select a lesser durable CH than the proposed HGCM, followed by the method proposed by Arkian et al. [21] at low traffic (100–500 vehicles) and by Kakkasageri et al. [47] at high traffic (600–1000 vehicles). The reason is that the ch selection in Arkian et al.’s method [21] is based on vehicle speed. As we have mentioned previously, the speed metric is lost in an urban scenario when there is immense congestion. hence, this proposed method achieves the lowest stability. HGCM achieves good CH stability in comparison with other algorithms due to the effectiveness of hypergraph theory and the novel set of CH selection parameters. our hgcm

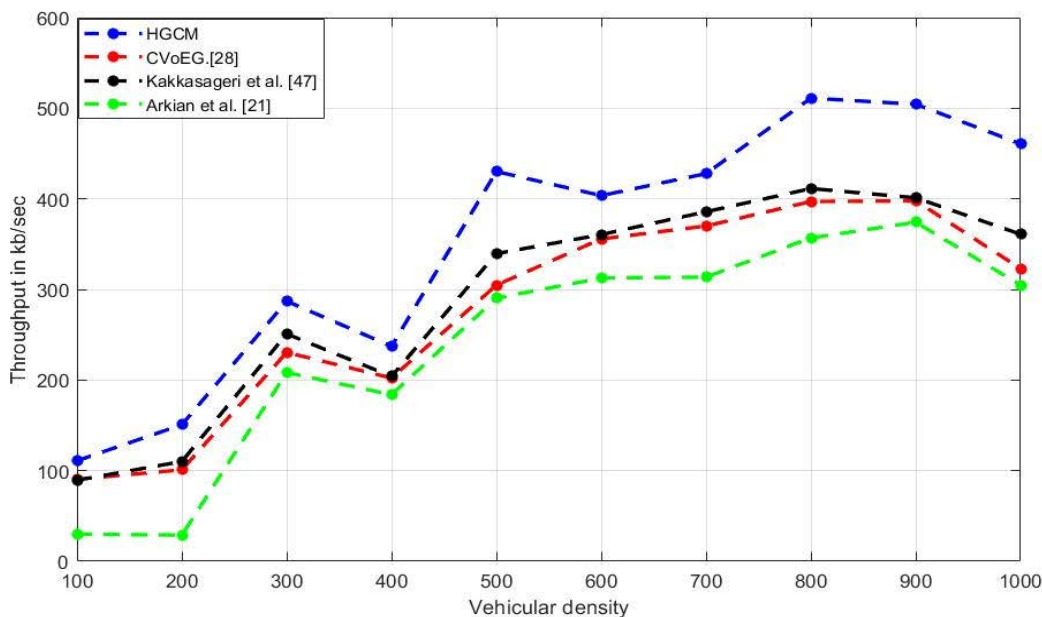


FIGURE 15. Throughput at different vehicle densities.

succeeds in achieving more than 53% of CH stability from the total time at all vehicle densities. Thus, the presentation of VANET as a hypergraph with its eigenvalues improves the CH stability.

3) EFFECT OF DIFFERENT TRAFFIC DENSITIES ON ROUTING PERFORMANCE

The stable CH improves the routing parameters, such as PD and throughput. These parameters are distance dependent. The minimum distance travelled by the packet leads to low PD and high throughput. All CMs should be one hop away from the CH. In an efficient cluster, the hop distance would be minimal. In the work presented in this paper, HGCM divides the network into 12 efficient clusters, which results in an average hop distance of 150 m for 1000 vehicles. By contrast, it is 240, 260 and 330 m for Kakkasageri et al [47], CVoEG [28], and the method proposed by Arkian et al. [21], respectively. Figure 14 shows the hop distance versus vehicle density curves on the right-hand y-axis and PD versus vehicle density on the left-hand y-axis. the maximum delay is witnessed in the method proposed by Arkian et al. [21] because it has a maximum hop distance. The CH location in the method proposed by Arkian et al. [21] is random, and it does not guarantee the centrality of CH while the CH location in CVoEG [28] is chosen based on the graph centrality. Nonetheless, the work presented by Kakkasageri et al. [47] guarantees less delay than the proposed work in [21] and [28] because the selection of Ch is based on the neighbouring degree; this provides the minimum hop distance. PD is low for a small average hop distance. With the increase in vehicle densities, the average hop distance increases and so is the PD. This finding validates that HGCM clustering shows better performance for a sparse network, which aligns with the general convention that a

crowded area increases PD. In sum, our HGCM reduces the PD by approximately 25%, 41% and 48% compared with the methods of Kakkasageri et al. [47], CVoEG [28] and Arkian et al. [21], respectively, at high traffic.

Throughput depends on the number of packets received in a small span. The minimum PD increases the throughput for the proposed HGCM scheme irrespective of the number of vehicles. Figure 15 shows the throughput curves. The hypergraph network presentation and novel set of CH selection help achieve 460 kb/s throughput compared with 350, 330, and 310 kb/s in other works at a density of 1000 vehicles. The proposed scheme helps achieve consistently improved throughput performance by approximately 31%, 39% and 48% compared with the methods of Kakkasageri *et al.* [47], CVoEG [28] and Arkian *et al.* [21], respectively, for high traffic.

VI. CONCLUSION AND FUTURE WORK

We have developed a novel cluster generation and maintenance strategy in this study. The CH is chosen based on a combination of four indicators that help maintain the stability of the dynamic network. A changing structure and the frequent connection and disconnection of communication links amongst vehicles are modelled in a directed evolving hypergraph formulation of VANET. Spectral clustering creates the ideal number of groups on the basis of the density of vehicles. Each cluster has a single RSU at its centre. Relative velocity score, eccentricity, neighbourhood degree and trust score are all recommended in this study for finding the most stable CH in each cluster. The proposed HGCM is tested for various vehicle densities in a real area in Iraq's capital, Baghdad. Compared with individual measures and other techniques,

our cumulative approach significantly improves CH stability. The addition of the trust element results in 20% gain in average CH stability over the combined performance of three existing measures (i.e. relative speed, eccentricity and neighbourhood). A one-hop network configuration is used to evaluate the approach for different integrated network metrics, including packet latency and throughput. The average packet distance travelled by the proposed method is 150 m with a delay of 0.2 s, whereas the other comparative algorithms under the same network conditions report a PD of 0.4 s for approximately 330 m according to the PD analysis for the worst-case scenario (i.e. 1000 vehicles). Therefore, HGCM has the lowest PD whilst still allowing for the shortest possible hop distance. In addition, PD directly influences throughput; hence, HGCM has the maximum throughput compared with other methods.

In future work, we will attempt to generate more efficient clustering by using a modularity matrix instead of an adjacency matrix. In addition, we intend to explore more metrics for analysis, through which the proposed methodology can be understood. The analysis of the algorithm's computational complexity will also be a benchmark of study in the next part of this work. CH stability decreases with the increase in vehicle density in the network. The solution can be verified by increasing the number of clusters in the network in the next part.

REFERENCES

- [1] F. Li and Y. Wang, "Routing in vehicular ad hoc networks: A survey," *IEEE Veh. Technol. Mag.*, vol. 2, no. 2, pp. 12–22, Jun. 2007.
- [2] M. M. C. Morales, C. S. Hong, and Y.-C. Bang, "An adaptable mobility-aware clustering algorithm in vehicular networks," in *Proc. 13th Asia-Pacific Netw. Oper. Manage. Symp.*, Sep. 2011, pp. 1–6.
- [3] J. P. Singh and R. S. Bali, "A hybrid backbone based clustering algorithm for vehicular ad-hoc networks," *Proc. Comput. Sci.*, vol. 46, pp. 1005–1013, Jan. 2015.
- [4] S. A. Mohammad and C. W. Michele, "Using traffic flow for cluster formation in vehicular ad-hoc networks," in *Proc. IEEE Local Comput. Netw. Conf.*, Oct. 2010, pp. 631–636.
- [5] S. Vodopivec, J. Bester, and A. Kos, "A survey on clustering algorithms for vehicular ad-hoc networks," in *Proc. 35th Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2012, pp. 52–56.
- [6] P. Basu, N. Khan, and T. D. C. Little, "A mobility based metric for clustering in mobile ad hoc networks," in *Proc. 21st Int. Conf. Distrib. Comput. Syst. Workshops*, Apr. 2001, pp. 413–418.
- [7] O. Kayis and T. Acarman, "Clustering formation for inter-vehicle communication," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Sep. 2007, pp. 636–641.
- [8] H. Su and X. Zhang, "Clustering-based multichannel MAC protocols for QoS provisionings over vehicular ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3309–3323, Jun. 2007.
- [9] Z. Y. Rawshdeh and S. M. Mahmud, "Toward strongly connected clustering structure in vehicular ad hoc networks," in *Proc. IEEE 70th Veh. Technol. Conf. Fall*, Sep. 2009, pp. 1–5.
- [10] N. Maslekar, M. Boussedjra, J. Mouzna, and L. Houada, "Direction based clustering algorithm for data dissemination in vehicular networks," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Oct. 2009, pp. 1–6.
- [11] R. T. Goonewardene, F. H. Ali, and E. Stipidis, "Robust mobility adaptive clustering scheme with support for geographic routing for vehicular ad hoc networks," *IET Intell. Transp. Syst.*, vol. 3, no. 2, pp. 148–158, Jun. 2009.
- [12] C. Shea, B. Hassanabadi, and S. Valaee, "Mobility-based clustering in VANETs using affinity propagation," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–6.
- [13] S.-S. Wang and Y.-S. Lin, "Performance evaluation of passive clustering based techniques for inter-vehicle communications," in *Proc. 19th Annu. Wireless Opt. Commun. Conf. (WOCC)*, May 2010, pp. 1–5.
- [14] E. Souza, I. Nikolaidis, and P. Gburzynski, "A new aggregate local mobility (ALM) clustering algorithm for VANETs," in *Proc. IEEE Int. Conf. Commun.*, May 2010, pp. 1–5.
- [15] A. Daeinabi, A. G. P. Rahbar, and A. Khademzadeh, "VWCA: An efficient clustering algorithm in vehicular ad hoc networks," *J. Netw. Comput. Appl.*, vol. 34, no. 1, pp. 207–222, 2011.
- [16] G. Wolny, "Modified DMAC clustering algorithm for VANETs," in *Proc. 3rd Int. Conf. Syst. Netw. Commun.*, 2008, pp. 268–273.
- [17] A. Koulakezian, "ASPIRE: Adaptive service provider infrastructure for VANETs," Univ. Toronto, Toronto, ON, Canada, Tech. Rep., 2011.
- [18] Z. Zhang, A. Boukerche, and R. Pazzi, "A novel multi-hop clustering scheme for vehicular ad-hoc networks," in *Proc. 9th ACM Int. Symp. Mobility Manage. Wireless Access (MobiWac)*, 2011, pp. 19–26.
- [19] N. Maslekar, M. Boussedjra, J. Mouzna, and H. Labiod, "A stable clustering algorithm for efficiency applications in VANETs," in *Proc. 7th Int. Wireless Commun. Mobile Comput. Conf.*, Jul. 2011, pp. 1188–1193.
- [20] Z. Y. Rawshdeh 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, pp. 1–13, Dec. 2012.
- [21] H. R. Arkian, R. E. Atani, A. Pourkhalili, and S. Kamali, "Cluster-based traffic information generalization in vehicular ad-hoc networks," *Veh. Commun.*, vol. 1, no. 4, pp. 197–207, Oct. 2014.
- [22] Y. Chen, M. Fang, S. Shi, W. Guo, and X. Zheng, "Distributed multi-hop clustering algorithm for VANETs based on neighborhood follow," *J. Wireless Commun. Netw.*, vol. 2015, no. 98, pp. 1–12, 2015.
- [23] D. Zhang, H. Ge, T. Zhang, Y.-Y. Cui, X. Liu, and G. Mao, "New multi-hop clustering algorithm for vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 4, pp. 1517–1530, Apr. 2019.
- [24] X. Ji, H. Yu, G. Fan, H. Sun, and L. Chen, "Efficient and reliable cluster-based data transmission for vehicular ad hoc networks," *Mobile Inf. Syst.*, vol. 2018, pp. 1–15, Jul. 2018.
- [25] G. H. Alsuhli, A. Khattab, and Y. A. Fahmy, "Double-head clustering for resilient VANETs," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–17, Mar. 2019.
- [26] 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, Apr. 2019, Art. no. e0214664.
- [27] X. Cheng and B. Huang, "A center-based secure and stable clustering algorithm for VANETs on highways," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–10, Jan. 2019.
- [28] Z. Khan, P. Fan, S. Fang, and F. Abbas, "An unsupervised cluster-based VANET-oriented evolving graph (CVoEG) model and associated reliable routing scheme," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3844–3859, Oct. 2019.
- [29] K. Bylykbashi, D. Elmazi, K. Matsuo, M. Ikeda, and L. Barolli, "Effect of security and trustworthiness for a fuzzy cluster management system in VANETs," *Cogn. Syst. Res.*, vol. 55, pp. 153–163, Jun. 2019.
- [30] W. Ahsan, M. F. Khan, F. Aadi, M. Maqsood, S. Ashraf, Y. Nam, and S. Rho, "Optimized node clustering in VANETs by using meta-heuristic algorithms," *Electronics*, vol. 9, no. 3, p. 394, Feb. 2020.
- [31] H. Fatemidokht and M. K. Rafsanjani, "QMM-VANET: An efficient clustering algorithm based on QoS and monitoring of malicious vehicles in vehicular ad hoc networks," *J. Syst. Softw.*, vol. 165, Jul. 2020, Art. no. 110561.
- [32] U. Maan and Y. Chaba, "Accurate cluster head selection technique for software defined network in 5G VANET," *Wireless Pers. Commun.*, vol. 118, no. 2, pp. 1271–1293, May 2021.
- [33] Z. Khan, A. Koubaa, S. Fang, M. Y. Lee, and K. Muhammad, "A connectivity-based clustering scheme for intelligent vehicles," *Appl. Sci.*, vol. 11, no. 5, p. 2413, Mar. 2021.
- [34] O. Senouci, S. Harous, and Z. Aliouat, "A new heuristic clustering algorithm based on RSU for Internet of Vehicles," *Arabian J. Sci. Eng.*, vol. 44, no. 11, pp. 9735–9753, Nov. 2019.
- [35] S. A. Sert and A. Yazici, "Increasing energy efficiency of rule-based fuzzy clustering algorithms using CLONALG-M for wireless sensor networks," *Appl. Soft Comput.*, vol. 109, Sep. 2021, Art. no. 107510.
- [36] S. Elaraby and S. M. Abuelenin, "Connectivity analysis of directed highway vehicular ad hoc networks using graph theory," *Int. J. Commun. Syst.*, vol. 34, no. 5, p. e4745, Mar. 2021.

- [37] M. Ran and X. Bai, "Vehicle cooperative network model based on hypergraph in vehicular fog computing," *Sensors*, vol. 20, no. 8, p. 2269, Apr. 2020.
- [38] G. Khayat, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, H. Maalouf, and E. Pallis, "VANET clustering based on weighted trusted cluster head selection," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2020, pp. 623–628.
- [39] F. Sanhaji, H. Satori, and K. Satori, "Cluster head selection based on neural networks in wireless sensor networks," in *Proc. Int. Conf. Wireless Technol., Embedded Intell. Syst. (WITS)*, Apr. 2019, pp. 1–5.
- [40] M. Hadded, P. Muhlethaler, A. Laouiti, and L. A. Saidane, "A novel angle-based clustering algorithm for vehicular ad hoc networks," in *Vehicular Ad-Hoc Networks for Smart Cities*. Singapore: Springer, 2017, pp. 27–38.
- [41] D. Ghoshdastidar and A. Dukkipati, "Uniform hypergraph partitioning: Provable tensor methods and sampling techniques," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 1638–1678, 2017.
- [42] D. Ghoshdastidar and A. Dukkipati, "A provable generalized tensor spectral method for uniform hypergraph partitioning," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 400–409.
- [43] X. Wang and Y. Xu, "An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 569, no. 5, p. 52024, 2019.
- [44] P. Ahrendt, "The multivariate Gaussian probability distribution," Tech. Univ. Denmark, Lyngby, Denmark, Tech. Rep., 2005.
- [45] D. R. White and S. P. Borgatti, "Betweenness centrality measures for directed graphs," *Soc. Netw.*, vol. 16, no. 4, pp. 335–346, 1994, doi: [10.1016/0378-8733\(94\)90015-9](https://doi.org/10.1016/0378-8733(94)90015-9).
- [46] J. Cheng, G. Yuan, M. Zhou, S. Gao, Z. Huang, and C. Liu, "A connectivity-prediction-based dynamic clustering model for VANET in an urban scene," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8410–8418, Sep. 2020.
- [47] M. S. Kakkasageri and S. S. Manvi, "Multiagent driven dynamic clustering of vehicles in VANETs," *J. Netw. Comput. Appl.*, vol. 35, no. 6, pp. 1771–1780, Nov. 2012, doi: [10.1016/j.jnca.2012.07.002](https://doi.org/10.1016/j.jnca.2012.07.002).
- [48] B. Soni, D. K. Patel, and M. López-Benítez, "Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics," *IEEE Access*, vol. 8, pp. 97437–97451, 2020, doi: [10.1109/ACCESS.2020.2995633](https://doi.org/10.1109/ACCESS.2020.2995633).
- [49] X. Qian and L. Hao, "On the performance of spectrum sensing in cognitive vehicular networks," in *Proc. IEEE 26th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Aug. 2015, pp. 1002–1006.
- [50] H. A. Shah and I. Koo, "Reliable machine learning based spectrum sensing in cognitive radio networks," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–17, Sep. 2018, doi: [10.1155/2018/5906097](https://doi.org/10.1155/2018/5906097).
- [51] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, "SUMO—Simulation of urban mobility: An overview," in *Proc. 3rd Int. Conf. Adv. Syst. Simulation*, May 2014, pp. 1–6.
- [52] A. Wegener, M. Piórkowski, M. Raya, H. Hellbrück, S. Fischer, and J.-P. Hubaux, "TraCI: An interface for coupling road traffic and network simulators," in *Proc. 11th Commun. Netw. Simulation Symp. (CNS)*, May 2008, pp. 155–163, doi: [10.1145/1400713.1400740](https://doi.org/10.1145/1400713.1400740).
- [53] F. Arena, G. Pau, and A. Severino, "A review on IEEE 802.11 p for intelligent transportation systems," *J. Sensor Actuator Netw.*, vol. 9, no. 2, p. 22, Apr. 2020, doi: [10.3390/jsan9020022](https://doi.org/10.3390/jsan9020022).
- [54] M. K. J. Alsabah, H. Trabelsi, and W. Jerbi, "Survey on clustering in VANET networks," in *Proc. 18th Int. Multi-Conf. Syst., Signals Devices (SSD)*, Mar. 2021, pp. 493–502.
- [55] M. Mukhtaruzzaman and M. Atiquzzaman, "Junction-based stable clustering algorithm for vehicular ad hoc network," *Ann. Telecommun.*, vol. 76, no. 9, pp. 777–786, 2021.



MAYS KAREEM JABBAR was born in Amarah, Misan, Iraq. She received the B.Sc. degree in computer engineering from the Computer Engineering and Information Technology Department, University of Technology, Baghdad, Iraq, in 2007, and the master's degree in wireless communications networks from the Computer Engineering Department, Eastern Mediterranean University (EMU), Famagusta, Northern Cyprus, in 2014. She is currently pursuing the Ph.D. degree with the Computer System Engineering Department, École Nationale d'Ingénieurs de Sfax (ENIS), University of Sfax, Tunisia.

She is working as a Lecturer with the College of Engineering, University of Misan, Misan. She is a member of the Smart Electric Vehicle Group at the Computer Embedded System Laboratory, ENIS. Her current research interests include VANET, V2V, and clustering algorithms.



HAFEDH TRABELSI studied at the École Nationale d'Ingénieurs de Sfax (ENIS), from 1983 to 1989. He received the M.S. degree from the École Centrale de Lyon, France, in 1990, the Ph.D. degree from the University of Paris—XI, Orsay, France, in 1994, and the Research Management Ability degree from ENIS, all in electrical engineering, in 2008.

He joined Tunisian University, in 1995. He is currently a Professor of electric power engineering with ENIS. Since 2013, he has been a Full Professor with the University of Sfax, Tunisia, holding the Chair for Smart Electric Vehicle Group at the Computer Embedded System Laboratory, ENIS. His research interests include the design of new electric machines by using finite element method and the implementation of advanced control systems for electric vehicle. His on-going research is focusing on smart city and secure systems.

• • •