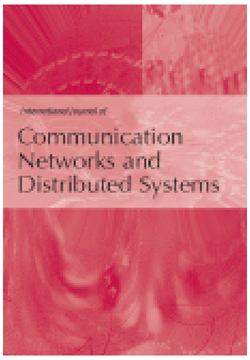


International Journal of Communication Networks and Distributed Systems

ISSN online: 1754-3924 - ISSN print: 1754-3916
https://www.inderscience.com/ijcnds

An enhanced priority-based multi-hop clustering algorithm for vehicular ad hoc networks

S. Rakhi, Kanakanahalli Ramanna Shobha

DOI: 10.1504/IJCNDS.2024.10055893

Article History:

Received:	26 October 2022
Last revised:	03 December 2022
Accepted:	06 December 2022
Published online:	30 November 2023

An enhanced priority-based multi-hop clustering algorithm for vehicular ad hoc networks

S. Rakhi*

Department of Electronics and Communication Engineering,
Atria Institute of Technology,
Bangalore, India
Email: rakhi.s@atria.edu
*Corresponding author

Kanakanahalli Ramanna Shobha

Department of Electronics and Telecommunication Engineering,
Ramaiah Institute of Technology,
Bangalore, India
Email: shobha_shankar@msrit.edu

Abstract: The importance of intelligence transportation system is increasing as it improves road safety and efficacy by means of vehicular ad hoc networks (VANET). The nodes in VANET are intelligent machines that can communicate with each other. Due to high mobility and frequent network fragmentation, stability is always a challenge in VANET. Even though traditional clustering methods address this issue, they exhibit less stability in highly dynamic scenarios. To improve the stability of the clusters, a new multi-hop clustering method named enhanced priority-based multi-hop clustering algorithm (EPMCA) is proposed. The best neighbours are chosen using neighbour following method. Then, stable clusters are established based on the average velocity of the cluster and the association lifetime between the nodes by the cluster head. The proposed algorithm shows significant improvement in average cluster head and cluster member duration, average cluster head changes, and number of clusters for varying communication ranges compared to existing techniques.

Keywords: vehicles; stability; cluster; priority; multi-hop; vehicular ad hoc networks; VANET; enhanced priority-based multi-hop clustering algorithm; EPMCA.

Reference to this paper should be made as follows: Rakhi, S. and Shobha, K.R. (2024) 'An enhanced priority-based multi-hop clustering algorithm for vehicular ad hoc networks', *Int. J. Communication Networks and Distributed Systems*, Vol. 30, No. 1, pp.42–69.

Biographical notes: S. Rakhi has completed her MTech in VLSI and Embedded Systems in 2010 from the Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India securing the state first rank with a gold medal. She is currently pursuing her PhD at VTU, Karnataka, India. She serves as an Assistant Professor in the Department of Electronics and Communication Engineering, Atria Institute of Technology, Bangalore, Karnataka. Her research interests include security in wireless networks and vehicular networks. She became a member (M) of IEEE in 2021.The author is a recipient of AWSAR

Award for best research article in the year 2019–2020 by the Department of Science and Technology, Government of India. She has presented research papers in the areas of VLSI, vehicular ad hoc networks and wireless networks.

Kanakanahalli Ramanna Shobha received her ME in Digital Communication Engineering from the Bengaluru University, Karnataka, India, and PhD from the Visvesvaraya Technological University. She is currently working as an Associate Professor in the Department of Electronics and Telecommunication Engineering, Ramaiah Institute of Technology, Bengaluru. Her research areas include mobile ad hoc networks, IoT and cloud computing. She has more than 25 papers publications to her credit. She is a senior IEEE member serving as a chair of IEEE Bangalore Section Sensors Council Chapter, ExCom member of WIE and IEEE Communication Society, Bangalore. She is also an active member of IETE, ISoc and IAENG.

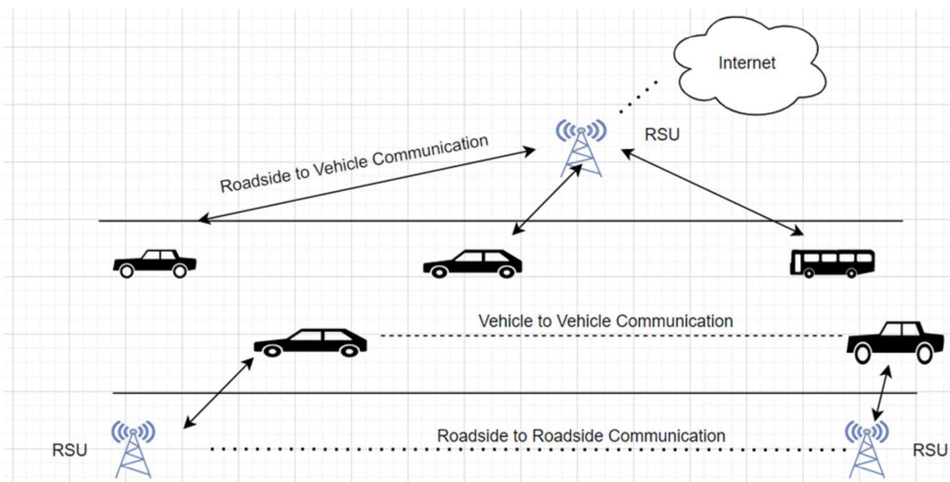
1 Introduction

As a significant element of intelligent transportation systems (ITS), vehicular ad hoc network (VANET) has fascinated many researchers from various fields and enormous contributions have been developed in this area. The major areas of research in VANET are routing, stability and security of the network. It needs VANET to disseminate information to road users. VANET architecture is divided into vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communication. VANET incorporates safety and security features with the help of V2V and V2I communication. The self-organisation of vehicles, information sharing and communication between vehicles are established using V2V communication (Chen et al., 2020). V2I communication is used mainly for accessing the internet through infrastructure roadside unit (RSU). In the architecture shown in Figure 1, V2V or intra-vehicular communication uses IEEE 802.11p/dedicated short range communication (DSRC) standard. V2I or inter-vehicular communication takes the help of existing wireless technologies like 4G/LTE or 5G in near future. The current VANET architecture combines the advantages of DSRC and the 4G/LTE. A rapid establishment of an ad hoc network is possible through DSRC. DSRC provides higher data transfer rates with the lowest latency. 4G/LTE network has higher bandwidth and wide propagation range. These features of DSRC and 4G/LTE are employed in VANET (Xu et al., 2017). Thus, the vehicular network architecture exploits the advantages of lower latency of DSRC and a wider propagation range of 4G/LTE or 5G (You et al., 2020).

In VANET, fast-moving nodes exhibits frequent topology changes, link reliability issues and overhead in message dissemination. To overcome these issues and to enhance communication, the clustering technique has been widely used. A virtual communication environment is formed by clustering that improves the stability, reliability and scalability of the network. Clustering is a common and practical approach in VANET management (Wang et al., 2015). It is a technique of grouping vehicular nodes based on predefined metrics like density, velocity, and geographical locations of the vehicular nodes with a centralised node (Bali et al., 2014). Centralised node is the cluster head (CH) and it can form its members called cluster members (CM).

Clustering in VANET unveils better scalability (Ren et al., 2021). Clustering allows efficient data delivery and less bandwidth consumption in the network. It can carry efficient load balancing. Since VANET has a highly dynamic scenario, it is tough to find an optimal solution to split vehicles to form stable clusters. The inadequate transmission range of radio modules and the high mobility of vehicles lead to frequent link fragmentation which limits a stable cluster formation in VANET. As compared to traditional mobile ad hoc networks (MANET), VANET exhibit rapid movement of vehicles with limited boundary conditions with no energy scarcity. So, we cannot adopt traditional MANET clustering algorithms in VANET (Verma et al., 2020). In a cluster-based framework, vehicles are connected to form clusters. The maximum range of communication between the vehicle nodes in a cluster is based on the DSRC standard. Hence, the nodes in the same cluster can exchange messages without any infrastructure support in a decentralised approach. Each cluster consists of a CH that communicates to the members and roadside infrastructure unit. Use of clustering technique to exchange the packets decreases the number of handshakes between the nodes and the infrastructure. But at the same time, unbalanced clusters are likely to generate more control packets that increase the network overload. Formation of clusters and its maintenance invariably involve extra overhead associated with control packets. Thus, a lucid clustering technique always emphasis on setting up least number of clusters and maintaining the clusters without increasing overhead in the network.

Figure 1 VANET architecture (see online version for colours)



To incorporate the above-mentioned advantages there are methods adopted for clustering, based on the number of hops a CH can connect. They are called single-hop and multi-hop clustering. In single-hop clustering, the data is exchanged between CH and next-hop neighbours. But the rapid change in topology reduces the performance of the VANET using single-hop. Multi-hop clustering helps N hop neighbours to communicate to a CH that are not direct neighbours. This improves the communication range between the nodes, thus improving efficiency. There are still challenges in clustering which have to be resolved related to the total number of clusters formed, transmission range between the nodes and intercluster interference. Based on the research and study of multi-hop clustering available in recent years, a novel multi-hop clustering with enhanced neighbour following strategy is proposed in this work to resolve these issues to certain extent. The main contributions of this work are as follows.

- In the unclustered phase of the network, a novel technique based on the destination is introduced to choose the most stable neighbour chasers.
- A novel cluster formation algorithm with improved stability is incorporated next.
- A cluster maintenance technique is incorporated to improve the reliability of the clusters.

The rest of the paper is organised as follows. Section 2 describes the related work of various clustering techniques in VANET. Section 3 details the system architecture. The overview of the clustering process and the proposed enhanced priority-based multi-hop clustering algorithm (EPMCA) is explained in Section 4. Section 5 explains the performance evaluation and experimental results. Section 6 draws conclusions of the work done.

2 Related work

Clustering techniques have been addressed in all ad-hoc networks (Cooper et al., 2017; Kadhim, 2021; Jabbar and Trabelsi, 2022; Srivastava et al., 2020). As a method of forming groups, clustering algorithm has been extensively studied since it improves routing scalability and stability. Stabilising neighbouring nodes in the fast-moving network is always challenging. Compared to traditional clustering, the reactive clustering algorithm is well suited in VANET since each vehicle owns a state. Whenever any cluster strategy changes, the node simply changes its state and act accordingly rather than undergoing re-clustering as in traditional methods (Gomathy et al., 2020). Reactive clustering can be done in one-hop and multi-hop mechanisms. Most of the earlier studies are mainly based on one-hop clustering, with one-hop communication between CH and cluster member. The main disadvantage of one-hop clustering is limited coverage area which leads to frequent re-clustering. In a multi-hop clustering method, the CH can communicate to its one-hop neighbour directly and multi-hop neighbours indirectly (Katiyar et al., 2020; Gu et al., 2022 Liu et al., 2021). The transmission range can be extended this way even though the distant neighbour is not in the direct communication range of the CH. Various multi-hop clustering techniques based on neighbours, topology, mobility, energy and weight have been proposed over the years (Zhang et al., 2011, 2019; Liu et al., 2018). The lowest and highest id clustering are the simplest and the earliest clustering algorithms proposed which are based only on the id of the nodes (Nguyen et al., 2015). But these algorithms mainly considered the unique ids of the mobile nodes for CH selection without giving any weightage to the mobility of nodes. Various one hop and multi-hop clustering techniques are discussed here in detail.

In unified clustering framework (UCF) (Zhang et al., 2021) relative position, velocity and link lifetime of the vehicles are considered at different traffic scenarios. Stable neighbours are selected based on neighbour sampling method. In this approach, vehicles moving in the same direction and that have a velocity less than the threshold value is selected as neighbours. Compared to traditional method in which CH is selected based on mobility metrics and broadcasting, this technique use back off timer that depends on position, velocity and link life time of nodes. Two methods are used to select CH, a metric based and random selection based. Frequent reclustering is a major concern in dynamic scenario. This is reduced by cashing CH scheme. The scheme shows better stability compared to other few algorithms. Wang and Lin (2013) proposed a pass passive clustering-aided routing algorithm (CAR) that uses a single-hop method in one-way multilane highway platoon scenario (Wang and Lin, 2013; Alsarhan et al., 2020; Zhang and El-Sayed, 2012). In this algorithm construction of stable clusters is done in the route discovery phase. The algorithm uses metrics such as node degree, expected transmission count, link stability to find CM. Less overhead is the advantage of this passive clustering. However, this passive algorithm did not consider broadcasting of the messages which is common in VANET (Lin et al., 2020; Elhoseny and Shankar, 2019).

AODV-MEC clustering algorithm (Ren et al., 2018) in IOV network that is based on edge computing strategy. It uses a reward function that depends on speed, energy consumption and link holding time of the nodes while cluster formation. Intermediate nodes in a route are calculated based on Q-learning algorithm. A RSU is selected as an edge server node that helps in maintaining lesser delay and overhead in a cluster. Mobile edge computing which is used here is mainly reduces bandwidth consumption and latency. This algorithm is purely cloud based. The results discussed in this work do not talk about varying speed scenarios. A link reliability-based clustering algorithm (LRCA) (Ji et al., 2018) provides efficient and consistent data transmission in VANET. A link lifetime-based neighbour sampling scheme is used here to filter out redundant unsteady neighbours. The CH is selected based on link reliability. This method is a single-hop clustering technique. But the delay in neighbour sampling is still a concern in the algorithm.

In Alsarhan et al. (2020), a neighbour discovery phase based on RSSI, speed and location are used to find neighbours. Any node can become CH satisfying least velocity condition and broadcast to other nodes. If multiple CHs are formed, based on a game theory stable CH is selected. This method increases the life time of the cluster. AMONET, is a clustering algorithm based on c (MFO), that work effectively in the highly mobile scenario in VANETs based on bio inspired techniques (Shah et al., 2022; Qureshi et al., 2021; Joshua et al., 2019; Hamdi et al., 2022). AMONET is a bio-inspired procedure that generates optimised clusters for dependable and effifcient communication. It is based on the navigation method used by moth which keeps a special angle with respect to moon. It is an effective mechanism for travelling for long distance. This method covers the entire network, and generates the fewest clusters and lowering vehicular routing costs. Cheng et al. proposed a clustering algorithm for LTE network in which eNodeB is the RSU. CH is elected based on mobility metric. eNodeB forms the cluster based on density. This method checks similarity in any activities in the nodes and alerts the other nodes (Cheng and Huang, 2019).

OCSR is a hierarchical clustering approach (Pandey et al., 2022; Zhao et al., 2022). A fixed node is selected as a CH. K-means clustering is used to form members. The method improves QoS parameters of the network (Fatemidokht and Rafsanjani, 2020). However, selecting CH is a challenge in dynamic scenario. In Kalaivani and Mouli (2020), CH is selected based on resource availability, degree of connection and relative velocity. CM are added to CH based on density in the same region my measuring the link lifetime. A distributed dispatching information table is maintained by each node to analyse the road conditions. This method improves cluster stability and reduces the delay in network. Flooding of messages is unavoidable in this case. Double head clustering (Alsuhli et al.,

2020) is a mobility-based clustering technique. Two CHs are available in same cluster named primary CH and secondary CH. CH selection is based on an eligibility criterion based on number of neighbours, relative position and speed. CM is selected by CH based on link expiration time and signal to noise ratio (SNR). CM connects to most eligible CH. It is single hop method. Maintaining two different CH is difficult in dynamic VANET scenario. All of the above mentioned algorithms are based on one-hop clustering (Elira et al., 2021; Sellami and Alaya, Alaya; Hamdi et al., 2022). Less coverage and frequent re-clustering are the main issues here. Subsequently, multiple clusters are formed in the network and a node has to select different CHs in its journey which decreases the overall cluster stability. A clustering algorithm should be able to reduce frequent CH changes and overhead. Frequent CH changes reduce routing efficiency and throughput. So, researchers understood the importance of multi-hop clustering and are moving towards developing multi-hop clustering techniques. However not many contributions are published in this area.

VMaSc (Ucar et al., 2013) is a multi-hop algorithm where CH is elected by calculating relative mobility which is a function of the difference in speed between all of the neighbours to itself. The node with the minimum relative mobility is elected as CH. It is suited if the nodes are moving in a particular lane with similar speeds. The neighbour strategy is a new method that can improve the stability of the clusters which reduces CH changes. Chen et al. (2015) propose a distributed multi-hop clustering algorithm (DMCNF) for VANET based on neighbourhood theory to elect CH. In this algorithm, neighbour strategy is used to find the best stable neighbours based on relative mobility, number of followers and past cluster related information. A node with least mobility and a greater number of followers is elected as a CH. The multi-hop neighbours are joined through members. The range of the CH can be extended which reduces CH changes. However, it does not take into account of the internode linkability which decreases the cluster stability. The DMCNF is modified in PMC (Zhang et al., 2019) to form the stable neighbours. A priority metric is calculated based on the number of followers, expected transmission count and link lifetime between the nodes. A node can choose a neighbour based on this metric. Once CH is elected based on the number of followers, the members join directly or indirectly. Since multi-hop neighbours are present in the clusters the range is increased. Frequent re-clustering is a drawback here because CH does not check the linkability with CMs.

A multi-hop clustering approach over vehicle-to-internet communication, MCA-V2I (Senouci et al., 2019) is another multi-hop clustering to improve the performance of the V2I communication. Since a node can communicate to roadside units, it shares the multi-hop neighbour information to RSU which performs the clustering process (Saidi et al., 2020). A mobility rate metric and breadth-first search (BFS) algorithm is introduced for clustering and connecting to RSU respectively. It uses a slave CH other than a master CH. Since RSUs are having a longer range, a greater number of mobile vehicles can be connected to them. Even though the authors claim that the performance is increased since two CHs are present, the overhead increases proportionally in this work. Considering all the above techniques we can conclude that multihop clustering provides improved stability, more coverage area and lesser number of clusters in a network.

3 System architecture

In this section, the multi-hop clustering architecture model and destination based best neighbour following strategy to choose a stable neighbour is explained. Analytic Hierarchical Process is explained in detail that is used to find the weightage of different metric to select the best neighbour. An enhanced priority based multi-hop clustering technique for VANET is discussed next. To implement the architecture assumptions are made.

- The highway road is divided into segments without interconnections.
- Radio modules in each vehicle in the network should follow DSRC standard.
- The nodes broadcast the beacons periodically using the 802.11p protocol standard within a range R.
- The nodes which receive the information from neighbours stores the details in the neighbour table called NBTable.
- The multi-hop distance is set in the range between 1 to 3 in this work for comparison purpose.

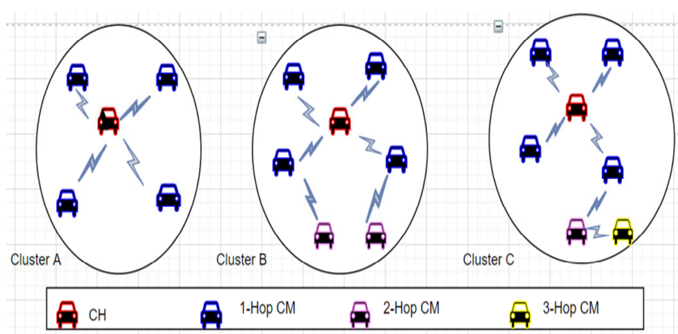
3.1 Multi-hop cluster architecture model

Multi-hop clustering is the best method to form a stabilised network in VANET. To understand the method the N hop cluster architecture model is shown in Figure 2. N stands for the number of hops. In the diagram Cluster A shows a 1-hop or single-hop clustering in which a CH connects directly to all members. The network performance is less in this case since the communication range is limited. Multi-hop clustering is the best option to improve communication range. The network performance can be further improved in multi-hop clustering by calculating a priority metric to choose the best neighbour. In cluster B and cluster C multi-hop connections are introduced in Figure 2. In cluster B, initially, each node connects to its direct one-hop neighbours using the priority neighbour following method. The priority metric is calculated based on a few parameters which are explained later. Once a CH is selected and a node becomes its CM, CH can connect to indirect multi-hop neighbours through CM and hence the range of CH is extended. CH in cluster B has 2-hop members. Cluster C has 3-hop connections. In a highly dynamic scenario, the multi-hop feature is of great use. Hence by incorporating destination and link lifetime in priority multihop algorithm (Wang et al., 2015) improves the performance of the proposed multi-hop algorithm. Before learning about the improved neighbour following strategy, the knowledge of features of a multi-hop cluster in a VANET scenario is important. They are mentioned below.

- Multiple hops: a cluster has CH and CM. A CH connects to CM through direct or indirect connections. A CH connects to a member using a one hop direct connectivity. If a node is connected to CH through an intermediate node called the parent node then the connection becomes indirect. This is possible through multiple hops. The multiple hops of a particular node and the number of nodes a CH can serve are configurable. Figure 2 shows the CHs communicating directly to CMs in cluster A. But in cluster B, a cluster is formed from direct and indirect neighbours with the help of multi-hop connections as explained previously. It is to be noted that the indirect neighbours will not be available in the communication range of CH. The indirect node has to be connected through an intermediate node that becomes its parent during communication.

- Hierarchical distribution: The HIERARCHY in multi-hop clustering is CH → CM → multi-hop members (MMs). MM is also referred to as a child node. CH communicates and manages CM and this member takes care of the MMs. The CM will become the parent node for the MMs. So perfect load balancing is maintained throughout the network. A distributed architecture is maintained in this fashion throughout the network. But care must be taken in selecting the maximum number of nodes a particular CH or CM can connect and manage.
- CH sharing: with the help of multi-hop clustering, the stability of the cluster can be improved since the CH is shared not only with the direct members but also with indirect members. Thus, the number of clusters formed in the entire network can be reduced. In multi-hop clustering, every node chooses the most stable neighbouring node to follow. Both these nodes share the same CH. Thus, the formation of multiple clusters can be reduced in the entire network.
- CH changes: since the clusters are using a multi-hop strategy, the cluster fragmentation is highly reduced. The child will try to maintain the CH through the parent until the parent has the connectivity to the CH. Subsequently, the duration for which a node remains a member of a particular CH increases. The selection of most stable parent is very important in multi-hop clustering. Cluster fragmentation can be reduced with this selection.

Figure 2 N-hop cluster architecture (see online version for colours)



3.2 Destination based best neighbour following strategy

The selection of the most stable neighbour node as a parent to follow is very important for the child node in multi-hop connections. The child will be following the parent throughout the journey only if the child is satisfying the neighbour following strategy. In a N-hop cluster, the CH node is at N hop distance with members. Thus, the direct member nodes can connect with CH in an ad-hoc mode. But selecting the indirect nodes that are not in the communication range R of the CH becomes critical for forming a stable cluster. During the initial phase of the network when no clusters are formed, each node can select a stable neighbour as a parent in its range to follow. Destination based priority neighbourhood metric is used in selecting the best parent node as shown in Algorithm 1. When the CH announcement is done, nodes can connect to the CHs through their parent. Therefore, the selection of the best parent is most important in any multi-hop clustering.

In large VANET since the connections are dynamic, node fragmentation is a frequent phenomenon. Hence the selection of a stable neighbour is a great challenge. A node cannot choose a parent or child based on one particular criterion. Accordingly, few cluster connectivity metrics are included in this work which helps to find the most stable parent-child connection. Multiple neighbour metrics that are used in the literature is considered in this work (Shah et al., 2022; Liu et al., 2018). Two parameters, relative destination and final destination are introduced during the calculation of the metric. Consider two nodes i, j with final destinations (x_i, y_i) and (x_j, y_j). The final destination is obtained from the global positioning system (GPS) in the vehicle. The vehicle's GPS manages the x and y coordinates of the destinations. The relative destination is the closest destination of a vehicle in a road segment. The destination of a node is broadcasted by them in the beacon messages using x and y positions.

Algorithm 1 EDP calculation

```
V = (V1, V2 ··· Vn), Set of n vehicles in the network
for all nodes n do
    Received Beacons
    Calculate the metric Δd_{i,j} for destination check
    Calculate the metric Δrd_{i,j} for destination check
        If Δd_{i,j} = sqrt(((xi − xj)^2 + (yi − yj)^2)) = 0 then
            Same destination
            Select as neighbours
    else if
            Check for relative destination Δrd_{i,j}
                if Δrd_{i,j} = sqrt(((rd_{xi} − rd_{xj})^2 + (rd_{yi} − rd_{yj})^2)) = 0 then
                    Same destination
                    Select as neighbours
                else
                    No selection
                End if
    End if
        Calculation of enhanced destination-based PRI(EDP_{i,j}) by each node
                EDP_{i,j} = A.overline(S_{Gi,j}) + B.Q_{Lij} + C.overline(A_T)
                Select least EDP_{i,j} to select most stable parent
end for
```

The final destination check of the two nodes i and j with positions x and y can be found from the Euclidian distance which is given in (1).

$$d_{i,j} = \sqrt{\left(x_i - x_j\right)^2 + \left(\left(y_i - y_j\right)^2\right)} \tag{1}$$

If $\Delta d_{i,j} = 0$ then both nodes i and j's final destinations are the same and the node j can be made as a neighbour of node i anytime in the journey. If $\Delta d_{i,j} \neq 0$ then check for the final destination of j which becomes the relative destination for i. Since the road is divided into segments, each segment provides a relative destination. Node i checks its relative destination coordinates and compares them with neighbouring node j's final destination. If it matches then both can become the neighbour's up to that particular road segment and updates their neighbour table. The relative destination check can be calculated using (2).

$$\sqrt{\left(rd_{xi} - x_j\right)^2 + \left(\left(rd_{yi} - y_j\right)^2\right)} \tag{2}$$

Once the neighbour table of a node is updated with the neighbours, three parameters are used to get the best parent node to follow which are described below.

3.2.1 Node degree, N_i and succeeding grade S_G

In VANET the positional relationship between the nodes is obtained through beacon messages broadcasted periodically. The neighbours are chosen based on the distance between the nodes and the transmission range of a particular node. A vehicle broadcasts the beacons to all the neighbouring nodes within the transmission range R and updates neighbours whose destinations match. Two vehicles i and j are considered as 1-hop neighbours only if the distance between the nodes is less than R. This is called node degree. The neighbours of a node $i(N_i)$ is found using (3).

$$N_i = \{j, \text{such that distance } D_{i,j} < R\} \tag{3}$$

During cluster formation the vehicular nodes that are moving in the same direction considered since we need a stable cluster throughout the journey from source to destination. The total node degree is calculated by considering all nodes in the neighbourhood that are communicating to the node. There are situations where nodes will be within the transmission range but not participating in the communication. Such nodes are not considered here. The total node degree is considered as the cardinality number of a set in set theory. They are the direct members of node i as in (4).

$$Degree\ of\ i,\ degree_i = |N_i| \tag{4}$$

In multi-hop architecture, the total neighbours are direct and indirect neighbours of a node i. Initially, consider that the network is in an un-clustered state and clusters are yet to be formed. At that time every vehicle will try to find its direct and indirect neighbours. The CM is formed using these direct and indirect neighbours. A vehicle i exhibits a direct relation with j provided j is present in N_i. This relation is represented by, a function f: f: $i \rightarrow j \wedge j \in N_i$.

At the beginning of the formation of the network, vehicles try to choose direct one-hop neighbours. In indirect neighbour selection, the neighbour need not be in N_i but should be a follower node of j, which can exist between i and j which is represented by, $i \rightarrow x \rightarrow j$. f_c is a function that represents the total number of nodes following directly and indirectly: $f_c : \{i \mid i \rightarrow j \vee i \mapsto j\}$. N_c is connected multi-hop neighbours of i. The total neighbours can be calculated from N_i and N_c by a metric called succeeding grade given in (5).

$$\text{Succeeding Grade, } SG = degree + N_c \tag{5}$$

This metric is used while forming the cluster. The larger the value of S_G better stability is observed in the clusters. For example, if the maximum hop is 2 and a vehicle has the largest S_G, it is the most stable node in 2 hop distance.

If C_c denotes a cluster with CH c which is directly and indirectly connected to communicating vehicles. This can be represented by: $C_c : \{i \mid i \rightarrow c \vee i \mapsto c \vee i = c\}$.

3.2.2 Quality link metric, Q_L

The quality link metric, Q_L is the number of expected transmissions of a packet to be received without error at the destination. It purely shows the reliability between two nodes that are communicating (Aydin et al., 2021). Since Q_L calculates the quality of a link, in multi-hop vehicular networks this parameter has to be considered. However, Q_L do not take care of the interference issues. Q_L is designed for single radio channel environment which is suitable for VANET. It gives the link throughput. This metric calculates the number of required transmissions of a packet, f_t in the forward direction and r_t in the reverse direction between wireless links. To compute these values, all nodes broadcast a probe packet every second only to the nodes travelling towards the same relative destination. Each node stores the number of probe packets received earlier from each neighbour in a beacon time. So, each vehicle can find the transmission and reception rates of probes in both directions through a particular wireless link.

The true value of the Q_L can be calculated from (6) taking reciprocal of the product of f_t and r_t

$$Q_L = \frac{1}{f_t * r_t} \tag{6}$$

Smaller Q_L will give better throughput or link quality since f_t and r_t are high.

3.2.3 Association lifetime A_T

Association lifetime, A_T between two vehicles is the time duration of connectivity of the vehicles within a fixed communication range. Since network fragmentation is more in VANET, A_T plays an important role to understand how long the nodes are in contact. Using the neighbour following concept, if the A_T between the nodes is high, the longer will be the connection between the nodes. This improves the stability of the cluster. Hence, this metric is used to find the best neighbour node to follow. The bond between the node is higher if A_T is more. The calculation of A_T is based on the accuracy of data received from the radio modules in vehicles whose destinations are the same. Consider a cluster C which has a CH c and n members. The number of hops can be configured between 1 to 3. Consider two nodes, vehicle n_0 with a velocity $v_0(t)$ at a position $p_0(t)$ and n_1 with a velocity $v_1(t)$ at a position $p_1(t)$ at time t. Both vehicles are in the same transmission range R. At time t, assume that n_0 is broadcasting beacons to n_1. The relative distance between the nodes can be represented using (7).

$$|p_0(t) - p_1(t)| < R \tag{7}$$

At time t + Ø, the latest locations of the two vehicular nodes can be obtained using (8) and (9).

$$p_0(t) = p_0(t) + \varnothing v_0(t) \tag{8}$$

$$p_1(t) = p_1(t) + \varnothing v_1(t) \tag{9}$$

The link between the two vehicles will expire when the transmission range becomes just equal to R. So, the equality condition to be satisfied considering is given in (10).

$$\begin{aligned} (t + A_T) - p_1(t)(t + A_T)| &= |p_0(t) - p_1(t)| \\ &= A_T\left[v_0(t) - v_1(t)\right| = R \end{aligned} \tag{10}$$

The exact locations and speed of the vehicles in 2D coordinates are represented in (11) and (12).

$$\begin{pmatrix} (0) \\ (1) \end{pmatrix} \equiv \left(p_{0x}(t),\ p_{0y}(t)\right) \tag{11}$$

$$v(t)\begin{cases} (0) \\ (1) \end{cases} \equiv \begin{matrix} \left(v_{0x}(t),\ v_{0y}(t)\right) \\ \left(v_{1x}(t),\ v_{1y}(t)\right) \end{matrix} \tag{12}$$

Using these equations, A_T at time t can be calculated using (13).

$$A_T(t) = \frac{A(t) - B(t)}{C(t)} \tag{13}$$

where

$$\begin{aligned} A(t) &= \left[R^2\left(\Delta_{vx}^2(t) + \Delta_{vy}^2(t)\right) - \left(\Delta_{px}(t)\Delta_{vy}(t) - \Delta_{py}(t) - \Delta_{vx}(t)\right)^2\right]^{1/2} \\ B(t) &= \Delta_{px}(t)\Delta_{vx}(t) - \Delta_{py}(t)\Delta_{vy}(t) \\ C(t) &= \Delta_{vx}^2(t) + \Delta_{vy}^2(t) \end{aligned} \tag{14}$$

and

$$\begin{aligned} \Delta_{px}(t) &= p_{0x}(t) - p_{1x}(t) \\ \Delta_{py}(t) &= p_{0y}(t) - p_{1y}(t) \\ \Delta_{vx}(t) &= v_{0x}(t) - v_{1x}(t) \\ \Delta_{vy}(t) &= v_{0y}(t) - v_{1y}(t) \end{aligned} \tag{15}$$

Equation (15) can be simplified as follows. If $\Delta v_{i,j}$ is the difference between the velocities and $\Delta d_{i,j}$ is the difference between the distances then association lifetime can also be written as:

$$A_T = \left(\left|\Delta v_{i,j}\right| R - \Delta d_{i,j}\right) \big/ \Delta v_{i,j}{}^2 \tag{16}$$

The exact positions and velocities of all vehicles in a range R can be received from the beacon messages. Then, the average association time AT_{Avg} can be calculated using (17) for *n* neighbouring vehicles connected to a particular vehicle at time *t*.

$$AT_{Avg} = \frac{\sum_{1}^{n} A_T(t)}{n} \tag{17}$$

By considering all the above metric the best parent to follow can be found before the formation of the cluster. Any child node can follow a parent node using the best of these values. A metric called enhanced destination-based priority (EDP) is formulated here that takes care of the nodes that are travelling in the same direction and destination. $EDP_{i,j}$ is the priority between the vehicles *i* and *j* travelling in the same direction towards the same destination and is calculated using (18).

$$EDP_{i,j} = \mathcal{A}.\overline{S_{Gi,J}} + \mathcal{B}.Q_{Lij} + \mathcal{C}.\overline{A_T} \tag{18}$$

Each parameter is given a particular weightage considering the effectiveness of parameters in the network. The sum of weights is always 1. The $\overline{S_{Gi,J}}$ and $\overline{A_T}$ are normalised values. The smaller the value of $EDP_{i,j}$, higher is the priority in choosing vehicles as a parent. In PMC algorithm (Wang et al., 2015) the authors have not taken care of the relative destination of motion while calculating the priority and forming the clusters. To calculate the weights of S_G, Q_L and A_T analytic hierarchy process (AHP) is used.

3.3 Selection of weights based on AHP

AHP is a powerful and simple tool that solves problems in multiple criteria decisions making while calculating weight (Klutho, 2013). Decisions are made based on a scale of intensity of importance. The steps involved in calculating the weights are explained below.

Step 1 To construct a set of pairwise comparison matrix called the criteria matrix (*X*). The 3 decision making parameters are succeeding grade S_G, link quality Q_L and association lifetime A_T represented in (19).

$$X = \begin{bmatrix} S_G \\ Q_L \\ A_T \end{bmatrix} \begin{bmatrix} S_G & Q_L & A_T \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} \begin{bmatrix} 1 & 3 & 6 \end{bmatrix} \tag{19}$$

Scale 1 to 9 in Table 1 shows the relative importance of one parameter over the other. In the above equation (19) the values for succeeding grade S_G, link quality L_Q and association lifetime A_T is taken as 1, 3 and 6 respectively. The values in (19) are given as per the importance of the parameters chosen in the algorithm. Here, the association lifetime is given highest significance compared to the other two parameters.

Step 2 Construct a reciprocal matrix through pairwise comparison using Table 1. The value x_{ij} denotes the strength of preference of i^{th} criteria over j^{th} criteria shown in (20).

$$X^R = [x_{ij}] = \begin{bmatrix} 1 & S_G Q_L & S_G A_T \\ \dfrac{1}{S_G Q_L} & 1 & Q_L A_T \\ \dfrac{1}{S_G A_T} & \dfrac{1}{Q_L A_T} & 1 \end{bmatrix} \tag{20}$$

Construct a normalised vector of X^N using (21).

$$X^N = \left[\frac{1}{k} \frac{x_{ij}}{\sum_{i=1}^{k} x_{ij}} \right] \tag{21}$$

where k denotes the number of criteria to be weighted, here $k = 3$. Calculate the weightage of each parameter in (18) using the equation (22).

$$W_i^T = \left[\frac{1}{k} \sum_{j=1}^{k} \frac{x_{ij}}{\sum_{i=1}^{k} x_{ij}} \right] = [\mathcal{A}\mathcal{B}\ \mathcal{C}] \tag{22}$$

Step 3 Consistency check. A consistency check is used to confirm the judgement errors if any, in the weight calculation. Initially, the consistency index, CI is calculated using the equation (23).

$$CI = \left[\frac{\lambda - n}{n - 1} \right] \tag{23}$$

where n denotes the number of elements that can be compared in criteria matrix X. $n = 3$ is used for our calculations. λ is calculated by the equation (24).

$$\lambda = \left[\frac{\sum_{i=1}^{n} \mu_i}{n} \right] \tag{24}$$

where μ_i is the consistency vector which can be calculated using the equation (25).

$$\mu_i = \left[\frac{\sum_{j=1}^{n} w_j x_{ij}}{w_i} \right] \tag{25}$$

Finally, the consistency ratio (CR) which is the ratio of the consistency index (CI) and the random index (RI) is calculated. The consistency of the AHP algorithm mainly depends on the value of CI in the calculation of CR using the equation (26).

Table 1 Scale of relative importance

Intensity of importance	*Definition*
1	Equal
2	Weak
3	Moderate
4	Moderate plus
5	Strong
6	Strong plus
7	Very strong
8	Very very strong
9	Extreme

The value of *RI* can be referred from the standard Table 2.

$$CR = \frac{CI}{RI} \tag{26}$$

If $CR < 0.1$ then the judgement is correct else repeat the algorithm for new judgement values. Table 3 shows different combinations of weight values using the AHP algorithm. The best results for $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ in (18) are obtained for the 3rd-row values in Table 3.

Table 2 Random index

Exponent number	*1*	*2*	*3*	*4*	*5*	*6*
RI	0	0	0.58	0.90	1.12	1.24

Figure 3 EPMCA flowchart

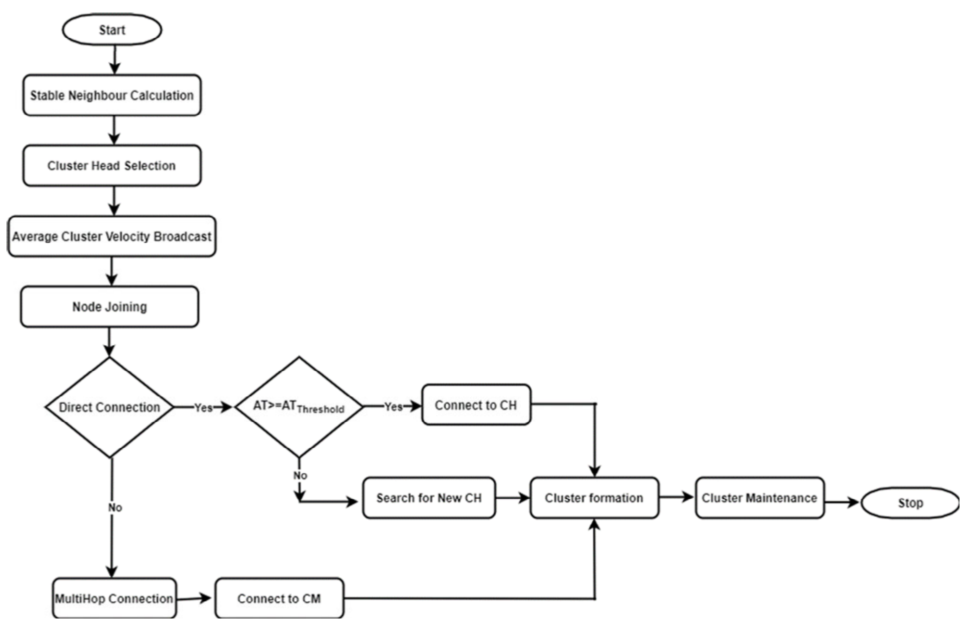


Table 3 Values of A, B, C

S_G	Q_L	A_T	$\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$
1	2	6	58	34	8
1	3	7	64	28	8
1	3	6	60	30	10
1	4	8	69	24	7
1	4	9	70	24	6

4 Enhanced priority based multi-hop

Once the network is deployed, nodes move independently and choose the best parent to follow as per the method explained in Section 3. The clustering process starts in the network. Every vehicle maintains a state in the cluster during this period. As time elapses these states tend to change as per the role taken by each node. There are 5 states in this clustering as shown in Table 4. S_0 is the initial state (IS). This state shows that the vehicle is just entered in the network and the nodes are ready to receive beacon messages from other nodes. S_1 is the primary state (PS) after connecting to a network from which nodes can change their states further. S_2 is the cluster member state (CMS) when a node becomes a part of a cluster. S_3 is the cluster head state (CHS) taken by a node when it announces itself as a CH. S_4 is the orphan state (OS) when nodes are no more a part of any cluster. The state transitions of a node with the help of a finite state machine (FSM) diagram are shown in Figure 4.

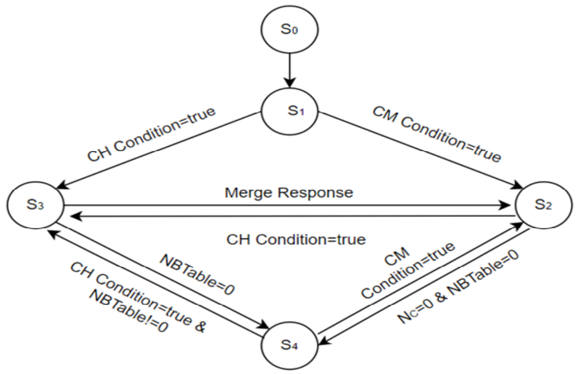
Table 4 Vehicle states

State	*Name*	*Abbreviation*
S_0	Initial state	IS
S_1	Primary state	PS
S_2	CM state	CMS
S_3	CH state	CHS
S_4	Orphan state	OS

In VANET once the network is deployed, node remains in IS for a time period and each vehicle broadcast beacon periodically during this interval. The beacon packets also called hello packets are transmitted with periodicity 10 hz. It contains the node id, velocity, position, direction of motion, CH id, number of hops to CH, number of followers and final destination. The neighbouring vehicles participating in the communication receive the beacons from neighbours. Each node updates its neighbour table NBTable. NBTable mainly includes vehicle id, state info, position, number of followers, parent id, CH id, timestamp, etc.

Using destination based best neighbour following strategy the priority metric is calculated. Node chooses the most stable neighbour with the lowest priority metric, updates NBTable and broadcasts these details in beacon packets. This stable neighbour node acts as a parent during clustering. Each time a beacon is received, a node updates the NBTable with the neighbours who are travelling towards the same destination. If a node is not receiving any hello packet for a specific duration from a neighbour, its details are removed from the node's NBTable. All these activities are performed when a node is in S0 state. The node remains in S0 state for a timer period.

Figure 4 FSM diagram of nodes



When this timer period expires, it will move to S_1 where it can change the states as per Figure 4. In S_1 state the vehicle can now participate in clustering. It can change the state to CH or member. If the node is in S_1 state and if it satisfies the CH selection condition it can move to S_3 state. It can become a member of an existing cluster by sending a request to a CH if it does not satisfy the head selection criteria. After receiving the response from the CH, a node becomes a part of an existing cluster with S_2 state. Once it is in CH state, based on the requests from other nodes it will form a new cluster. A CH node in S_3 changes its state to S_2 by sending a merge request to CM based on cluster merging criteria explained later in this section. In certain cases, the nodes become isolated from other nodes with $N_i = 0$ or when their NBTable is empty. Then they cannot join any other clusters. They are called orphan nodes with OS S_4. In that case either it can wait for a timer to expire and become the CH again after satisfying CH condition with NBTable $\neq 0$ or it can become a cluster member.

CH selection is an important criterion in clustering since a stable CH improves the performance of the entire clustered network. CH changes should be always kept minimum in a network to establish a reliable and stable network. Here, two parameters are taken for CH selection. The number of followers and average relative velocity of nodes. In every time period t, each node has all the data about the vehicles in its communication range. Every node calculates the average velocity difference, A_{vi} from all neighbours using (27). If i and j are two nodes, where $i \neq j$.

$$A_{vi} = \frac{1}{n_i} \sum_{11}^{n_i} v_i - v_j \tag{27}$$

A lower value of the relative velocity of a particular node means that the neighbours have spent a longer duration within its transmission range. This shows a more stable node that can be elected as a CH. The relative velocity, RA_{vi} can be calculated using (28).

$$RA_{vi} = \frac{A_{vi}}{v_i} \tag{28}$$

The overall average relative velocity, ARA_{vi} of vehicles that are connected directly to a particular node, *i* is given in (29).

$$ARA_{vi} = \frac{1}{n_i} \sum_{1}^{n_i} RA_{vi} \tag{29}$$

Lesser the value of ARA_{vi}, lesser the mobility of the vehicle related to its neighbour vehicles in its communication range provided all nodes are active. Hence, low value of ARA_{vi} indicates that the node is moving closer to its neighbours for a longer duration of time. From the above metrics, a criterion for the CH selection is formulated. If a particular node *i* wants to become a CH node then the criteria in (30) has to be satisfied. A node that satisfies both criteria in this equation advertises itself as a CH and is ready to form a cluster in the network. Once the node becomes a CH, it broadcasts the average velocity of its cluster in the CH announcement periodically.

Since VANET is highly dynamic, checking the average velocity of a particular cluster is important before joining that cluster. When a CH receives a request from a node it calculates its association lifetime, AT_{ij} with that node. If the AT_{ij} of a vehicular node is larger than the specified threshold value, only then CH sends a response to that node and accepts it as a member. The threshold value is calculated by considering the average value of the association time of all the members in the cluster as in (31). The association lifetime is given high importance in calculating the weight in the AHP algorithm as discussed in Section 3. Because, it defines how long a particular node is in contact with a CH during its journey.

$$CH\ criteria = \begin{cases} Succeeding\ grade\ is\ more\ than\ neighbours \\ ARA_{vi}\ is\ less\ than\ neighbours \end{cases} \tag{30}$$

$$\begin{aligned} AT_{ij} &= \left(\left| \Delta v_{ij} \right| R - \Delta d_{ij} \right) \big/ \Delta v_{ij}^2 \\ A_{Tij} &\geq A_{Thresholdij} \end{aligned} \tag{31}$$

Algorithm 2 Clustering

```
for all nodes n do
CH selection criteria = true
CH announcement with average cluster velocity
    if CH flag = true then
        if CH announcement received then
            send join req by nodes
                if CH_NB < Max_Capacity_CH && A_Tij ≥ A_Thresholdij then
                    Send join response b CH
                    Node_state >> CM
                else
                    Try for other CH
                end if
        end if
else
```

```
            CH flag = false
    end if
end for
for i = 1 to n do
    if CH id ≠ null in Parent ∈ CM then
            Check for CH flag = true
            Send join request by MH nodes to Parent
                If ParentNB = Max_Capacity_CM && NHop < max_hop
                    Send join resp by Parent
                        Node >> CM
                else
                        CH flag = false
                end if
            end if
end for
for i = 1 to n do
    if NBTable = null then
        n state = orph state
    else CH condition = true
        n State = CH
        reinitiate clustering
    end if
end for
```

A new strategy is adopted in this work while forming the CM compared to the work in Wang et al. (2015). The complete flowchart of the EPMCA is shown in Figure 3. Here, the destination of each node is considered while choosing the best parent to follow. At each time interval, the final destinations of the vehicles are checked while forming the neighbours. Once the beacons are received and the nodes update their neighbour list, each node calculates $EDP_{i,j}$ of their neighbours within the priority timer interval. The node having the least $EDP_{i,j}$ is designated as a stable parent to follow. Once the parent node is selected as the node to follow, the node checks if the parent is already a member of the cluster or not. If yes, then the node will make the CH of the parent as its CH and become the CM through multi-hop connection. The advantage of the multi-hop connection is that even if the range of the CH is not available to the child node still it can become a member of the same CH. This forms an indirect linkage between the CH and the node. If the parent is not a member of any cluster, then the child follows the parent until the parent gets connected to a new CH or it can connect to any other CH in its range. If the CH reaches its maximum capacity, then no nodes are added to the cluster, else the requested node receives a response from CH to join. During the vehicle movement, if a node cannot follow any parent and it does not have any neighbours, such nodes become orphan nodes. When this node satisfies CH condition or if it finds new neighbours then it can prepare for a CH announcement. The orphan nodes which come in the same transmission range can make a new cluster. The complete steps are shown in Algorithm 2.

During the transit of the vehicles when two CHs *i* and *j* that are in two different clusters become neighbours, then there will be a possibility of inter-cluster interference. To avoid this issue cluster merging should be initiated. The CH who wants to merge sends a request to the other CH. During this time CHs exchanges information about average velocity and the number of followers. The relative destination of the CHs is also checked here. Consider two CHs *i* and *j* which are coming into the range R. If the relative mobility of *i* is greater than j and both are travelling in the same direction having same destination then merging takes place. The CH *i* joins as a member of CH *j*. If multiple merge responses are received by a vehicle, it checks own maximum capacity and allows the members to join accordingly. Multi-hop clustering has an advantage in merging also. When a CH is merged with a new cluster the CMs and the followers will automatically become the members of the new CH satisfying the maximum capacity criteria. CMs and followers update the neighbours list with new details. If maximum criteria are not satisfied then the followers are omitted from becoming members. In a clustered network a node can hear from two CHs during its journey. In that case, the node acts as a gateway between the two clusters.

5 Performance evaluation and results

In this section, the performance of EPMCA in comparison with other multi-hop clustering algorithms which are based on neighbour theory is presented in detail. EPMCA is compared with new multi-hop clustering algorithm for VANETs (Wang et al., 2015), Vehicular multi-hop algorithm for stable clustering in Vehicular ad-hoc networks (Aissa et al., 2015) and distributed clustering algorithm for VANET based on neighbourhood follow (Liu et al., 2018). NS3 (release 3.0) network simulator is used here. The vehicle traces are obtained from MOVE-SUMO traffic simulator.

Table 5 Simulation parameters

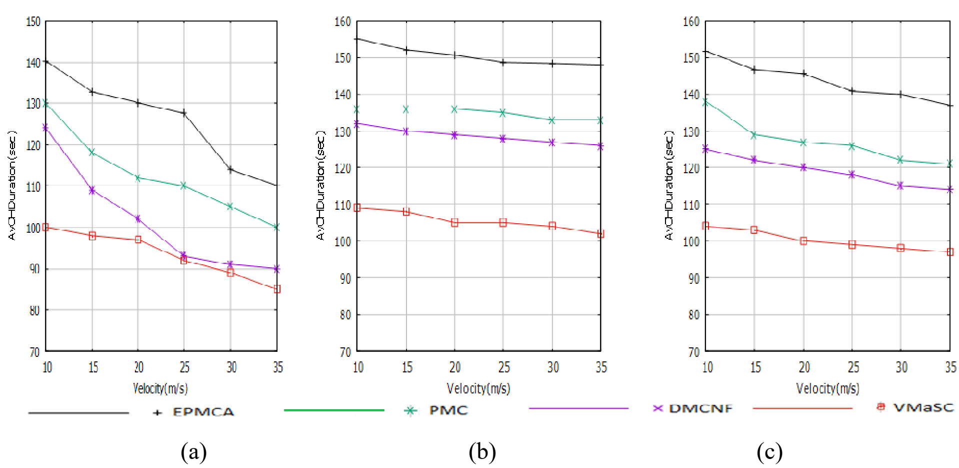
Parameters	*Values*
Simulation time	300 s
Road length/topology	3 km, 2-way highway
Max speed	10–35m/s
No. of vehicles	100
Propagation model	Two ray ground
Transmission range	100 m–300 m
Protocol	802.11p
Frequency/BW	5.9G Hz/10 MHz
Multihop	1–3
Beacon packet size	64 bytes
No. of trials	40

The simulation parameters used are as shown in Table 5. The simulation time is 300 s and multiple runs are performed to obtain better results. The performance of the algorithm is evaluated using average CH duration time, average CM duration time, the average number of CH changes, the number of clusters formed and cluster overhead. The results are obtained for various speeds from 10m/s (36 km/h) to 35 m/s (144 km/h) and the transmission range between 100 m to 300 m based on DSRC standard (Xu et al., 2017). The number of hops N = 1, 2, 3 is taken while comparing with other algorithms. The beacon size is 100 bytes maximum. In priority multi-hop clustering (PMC) (Wang et al., 2015) algorithm, priority metrics are considered first to select the most stable neighbour. CH is elected based on average relative mobility and number of followers. VMacSC (Aissa et al., 2015) is another multi-hop algorithm where the least velocity as a function of velocity difference is used to form multi-hop clusters. DMCNF (Liu et al., 2018) is the first algorithm that proposed distributed clustering in which CHs are selected based on neighbourhood follow relationship. To evaluate the stability of the clusters and the performance of the EPMCA algorithm, five main parameters listed above are considered. The proposed algorithm improves the network performance parameters in the range between 5% to 10% compared to other algorithms. The comparison between these algorithms is explained in the results.

5.1 *Average CH duration*

CH duration is the time a node in the CH state changes to a non-CH state. The non-CH states are CM state and OS. The average CH duration time is the ratio of the total CH duration time to the number of CHs (Wang et al., 2015). Figure 5 shows the average CH duration with velocity. Figures 5(a), 5(b) and 5(c) represents three communication radius R = 100 m, 200 m and 300 m respectively. Average CH duration shows a decreasing tendency when the velocity increases for all ranges in all algorithms. As velocity increases, CH will have difficulty in maintaining the CH state since retaining stable neighbour vehicles will become challenging. At high velocities, network fragmentation is a common issue in VANET. This results in re-clustering and cluster merging.

Figure 5 Average CH duration vs. velocity, (a) R = 100 m, (b) R = 200 m, (c) R = 300 m (see online version for colours)



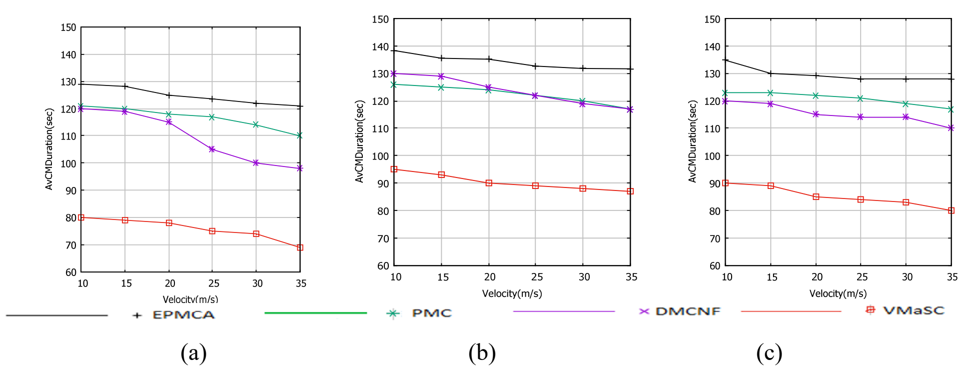
From the results in Figure 5, it is clear that EPMCA proposed in this paper and PMC have higher CH duration time since both use stable neighbour strategy. The stability of each cluster can be improved by this strategy. CH duration of VMaSC is relatively less.

In EPMCA, using enhanced priority neighbourhood strategy the most stable neighbours are selected. The transmission range has a major influence on stability. All the graphs show a marginal increase in CH duration time for R = 300 m. So, it is clear that the CH is not losing the connection with neighbours for wide ranges. From the readings, it is observed that EPMCA has got more CH duration time compared to PMC even though both use neighbourhood theory. However, in EPMCA the destination-based neighbourhood selection improves the best neighbour selection. Also, each CH announces its average cluster velocity so that the nodes who can maintain that velocity only chooses to become a member. CH does a proper verification check before selecting a member by checking its association lifetime, which predicts how long those nodes can be in the vicinity. Hence nodes with longer association time are selected as members. These parameters make EPMCA more stable than other multi-hop algorithms. The stability of the cluster is increased at higher transmission ranges because of higher linkability.

5.2 Average cluster member duration

Cluster member duration is the time period between a node joining a cluster and leaving the cluster. This condition is obvious when a node changes state to the CM state or CM changes to other states. On average, in most of the multi-hop algorithms, the average CM duration decreases as the speed is increased. But the vehicle velocity moderately changes in EPMCA. This is because the CH broadcasts the average velocity of the cluster. So, the nodes which are planning to continue with the same speed only will join the CH. A vehicle can change its parent during the journey, but the possibility is more for the new parent to be a member of its old CH itself. This is possible because the parents are selected based on relative destinations. So, the vehicles still follow the same CH indirectly which increases the stability of the cluster. From Figures 6(a), 6(b) and 6(c), it is clear that the average CM duration is high in EPMCA and manages to be steady with increased transmission radius also.

Figure 6 Average cluster member duration vs. velocity, (a) R = 100 m, (b) R = 200 m, (c) R = 300 m (see online version for colours)



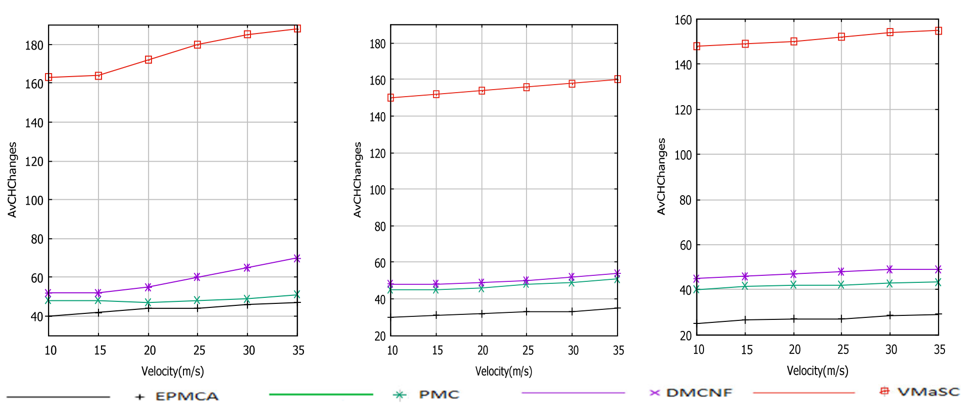
Compared to other algorithms EPMCA has a higher average CM duration. Before making a node its member, CH checks the association time which is a prediction of the lifetime between those particular nodes with CH. This parameter should be greater than the prescribed threshold value for a node to become member. Hence most of the members will be travelling with the CH for a longer duration of time which increases the CM duration in the cluster. Even though PMC and DMCNF use the neighbour strategy, both does not validate the member selection. This strategy makes the new EPMCA technique to perform better than other algorithms. So even at higher speeds, cluster has stable members.

5.3 Number of CH changes

The Number of CH Changes is the total number of nodes whose state changes from CH to CM or OCH. A low count gives higher stability. In Figure 7 for all R values this parameter moderately increase with an increase in velocity. The communication radius also affects count as shown in Figures 7(a), 7(b) and 7(c). CH count mainly considers the status of neighbour nodes. In a multi-hop cluster, a CH has direct and indirect members. Direct members are the one-hop neighbours and indirect are the multi-hop neighbours in a cluster.

In the usual scenario, CH is re-elected when there is a drastic reduction in the one-hop neighbours. One-hop neighbours usually satisfy the condition of velocity and association lifetime. The node state changes in multi-hop neighbours will have a minimum effect on CH selection. Due to all the above reasons compared to PMC and DMCNF, EPMCA shows lesser values for CH change count. However, these algorithms show stable CH changes at larger R since all of them use neighbour strategy. The results prove that as the transmission range increases, the link reliability of the members will become more so the CH count will become stable. VaMaSC shows a higher value since the CH selection is based only on the relative speeds of the vehicles.

Figure 7 Average CH changes vs. velocity, (a) R = 100 m, (b) R = 200 m, (c) R = 300 m (see online version for colours)



(a) (b) (c)

5.4 Number of clusters

In a network, if the number of clusters formed is less, then the network efficiency increases. Figure 8 shows the average number of clusters with varying velocities. The numbers of clusters are calculated for each simulation run and the average is taken. The numbers of cluster changes are less with increasing velocity in EPMCA. The cluster size tends to decrease as the number of hops between the nodes increases. In a multi-hop architecture, the number of clusters formed depends on mobility of nodes also.

Figure 8 Average cluster number vs. velocity (see online version for colours)

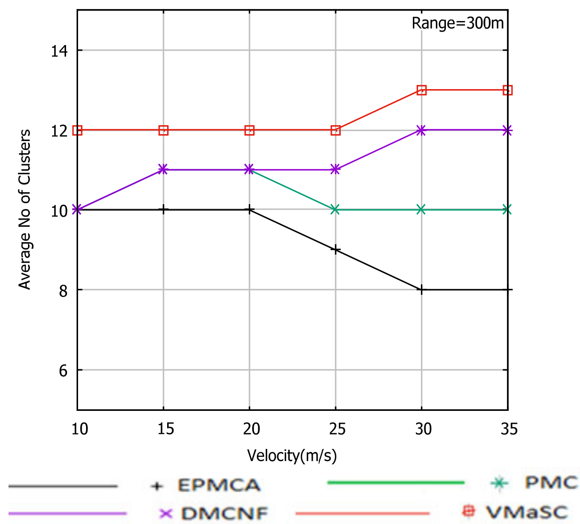
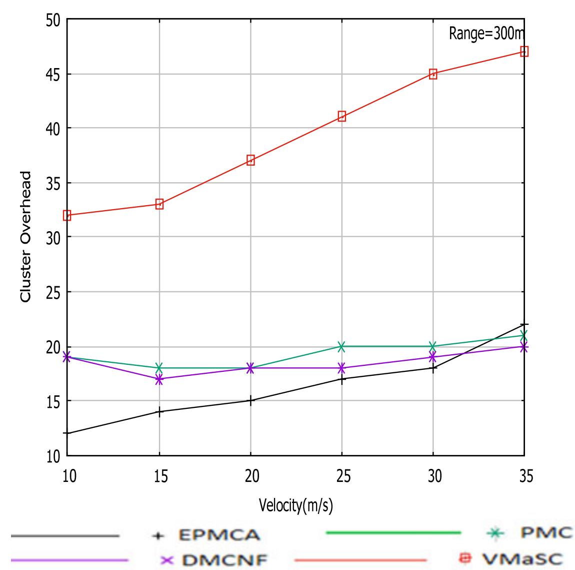


Figure 9 Cluster overhead vs. velocity (see online version for colours)



Since in EPMCA the CH changes are minimal, the number of clusters is reduced compared to other multi-hop algorithms. A CH preserves its state for a longer period of time and the number of cluster member changes is also minimal in this algorithm. Hence the number of clusters formed during a time interval is less. Even though a parent node with a child node changes its CH, the parent vehicle can still choose a member of the same cluster as a new CH if any in range R. It maintains to be in a CMS in a cluster. Since the member can have all the details of the new CH, it is easy to form a new cluster. The number of clusters has a direct dependency on the transmission range. If the range of a CH is more, then the members it can collaborate with also will be more considering the maximum capacity. A cluster is stable only if the coherence of the cluster structure formed is high.

5.5 Clustering overhead

It is calculated from the total number of control packets acquired by the vehicle from its neighbours during clustering. Overhead is the ratio of control packets to the total packets exchanged between the nodes. It can be calculated by considering the total control packets exchanged between the nodes throughout the cluster formation phase and maintenance phase. The clustering overhead increases in EPMCA compared to the other two priority based algorithms only at higher speeds as shown in Figure 9. It gives a reasonable result at lower speeds. At higher speeds, re-clustering occurs frequently and the number of packets exchanged is higher. During the cluster formation phase and maintenance phase, the control packets are disseminated among the nodes to improve the cluster health as velocity increases. The size of the beacon message in EPMCA is larger since the destination details are also included compared to PMC. Due to these reasons the overhead increases at high velocities. In Figure 9, the transmission range R = 300 m with $N = 3$ and a maximum ten member nodes in a cluster is considered.

6 Conclusions

In this manuscript, a novel multi-hop clustering mechanism is proposed for VANET. A destination based optimal neighbour following strategy is used for neighbour selection by considering three parameters, i.e., node degree, quality link metric and association lifetime. The enhanced clustering algorithm, i.e., EPMCA chooses the most stable multi-hop neighbours while forming the cluster. An improvement in the stability of the cluster throughout the journey of the nodes is possible due to this approach. The selection of CH is done based on the average cluster velocity and the association lifetime of nodes. This approach can reduce the number of clusters in the network effectively. Cluster coverage and load balancing are improved by the cluster merging process. Experimental results show that the cluster health monitoring parameters are improved in the enhanced algorithm, i.e., EPMCA as compared to the other multi-hop clustering algorithms. This approach is fully distributed and lightweight since infrastructure support is minimal. Even though the results show an improvement in stability and reliability in clustering, at higher velocities, overhead is slightly increased. This occurs due to the increase in control packets which are exchanged between the nodes at different interval of time. The overall results reveal that the optimal performance of the proposed algorithm is superior to the other existing algorithms. This work can be extended further in the direction to reduce the overhead and also in the security aspects in VANET.

References

Aissa, M., Belghith, A. and Bouhdid, B. (2015) 'Cluster connectivity assurance metrics in vehicular ad hoc networks', *Procedia Comput. Sci.*, Vol. 52, No. 1, pp.294–301.

Alsarhan, Y.K., Al-Dubai, A., Zomaya, A.Y. and Hussain, A. (2020) 'Novel fuzzy and game theory based clustering and decision making for VANETs', in *IEEE Transactions on Vehicular Technology*, February, Vol. 69, No. 2, pp.1568–1581, DOI: 10.1109/TVT.2019.2956228.

Alsuhli, G.H., Khattab, A. and Fahmy, Y.A. (2020) 'Double-head clustering for resilient VANETs', *Wireless Communications and Mobile Computing*, Article ID: 2917238, Vol. 2020, 17p.

Aydin, M.A., Karabekir, B. and Zaim, A.H. (2021) 'Energy efficient clustering-based mobile routing algorithm on WSNs', in *IEEE Access*, Vol. 9, pp.89593–89601, DOI: 10.1109/ACCESS.2021.3090979.

Bali, R.S., Kumar, N. and Rodrigues, J.J.P.C. (2014) 'Clustering in vehicular ad hoc networks: taxonomy, challenges and solutions', *Veh. Commun.*, Vol. 1, No. 3, pp.134–152.

Chen, J., Mao, G., Li, C. and Zhang, D. (2020) 'A topological approach to secure message dissemination in vehicular networks', in *IEEE Transactions on Intelligent Transportation Systems*, January, Vol. 21, No. 1, pp.135–148, DOI: 10.1109/TITS.2018.2889746.

Chen, M.F., Shi, S., Guo, W. and Zheng, X. (2015) 'Distributed multi-hop clustering algorithm for VANETs based on neighbourhood follow', *Eurasip J. Wirel. Commun. Netw.*, Vol. 2015, No. 1, pp.1–12.

Cheng, X. and Huang, B. (2019) 'A center-based secure and stable clustering algorithm for VANETs on highways', *Wireless Communications and Mobile Computing*, pp.1–10, DOI: 10.1155/2019/8415234.

Cooper, C., Franklin, D., Ros, M., Safaei, F. and Abolhasan, M. (2017) 'A comparative survey of VANET clustering techniques', in *IEEE Communications Surveys & Tutorials*, Vol. 19, No. 1, pp.657–681, firstquarter, DOI: 10.1109/COMST.2016.2611524.

Elhoseny, M. and Shankar, K. (2019) 'Energy efficient optimal routing for communication in VANETs via clustering model', in *Emerging Technologies for Connected Internet of Vehicles and Intelligent Transportation System Networks*, Vol. 242, pp.1–14, Springer: Cham, Switzerland.

Elira, B., Keerthana, K.P. and Balaji, K. (2021) 'Clustering scheme and destination aware context-based routing protocol for VANET', *International Journal of Intelligent Networks*, Vol. 2, pp.148–155, DOI: 10.1016/j.ijin.2021.09.006.

Fatemidokht, H. and Rafsanjani, M.K. (2020) 'QMM-VANET: an efficient clustering algorithm based on QoS and monitoring of malicious vehicles in vehicular ad hoc networks', *J. Syst. Software*, July, Vol. 165, No. 2, p.110561.

Gomathy, V. et al. (2020) 'Malicious node detection using heterogeneous cluster based secure routing protocol (HCBS) in wireless ad hoc sensor networks', *J. Ambient Intell. Humanized Comput.*, Vol. 11, No. 11, pp.4995–5001.

Gu, K., Dong, X., Li, X. and Jia, W. (2022) 'Cluster-based malicious node detection for false downstream data in fog computing-based VANETs', in *IEEE Transactions on Network Science and Engineering*, 1 May-June, Vol. 9, No. 3, pp.1245–1263, DOI: 10.1109/TNSE.2021.3139005.

Hamdi, M.M., Audah, L. and Rashid, S.A. (2022) 'Data dissemination in VANETs using clustering and probabilistic forwarding based on adaptive jumping multi-objective firefly optimization', in *IEEE Access*, Vol. 10, pp.14624–14642, DOI: 10.1109/ACCESS.2022.3147498.01.

Jabbar, M.K. and Trabelsi, H. (2022) 'Clustering review in vehicular ad hoc networks: algorithms, comparisons, challenges and solutions', *International Journal of Interactive Mobile Technologies*, Vol. 16, No. 10, pp.1–10.

Ji, X., Yu, H., Fan, G., Sun, H. and Chen, L. (2018) 'Efficient and reliable cluster-based data transmission for vehicular ad hoc networks', *Mob. Inf. Syst.*, Article ID: 9826782, Vol. 2018, pp.1–15.

Joshua, C.J., Duraisamy, R. and Varadarajan, V. (2019) 'A reputation based weighted clustering protocol in VANET: a multi-objective firefly approach', *Mob. Netw. Appl.*, Vol. 24, No. 4, pp.1199–1209.

Kadhim, A. (2021) *Parked Vehicles as Cluster Heads and Gateways for Efficient Routing in VANET*, DOI: 10.21203/rs.3.rs-397977/v1.

Kalaivani, D. and Mouli, P.C. (2020) 'Link survivability rate-based clustering for QoS maximisation in VANET', *Int. J. Grid Util. Comput.*, Vol. 11, No. 4, p.457.

Katiyar, A., Singh, D. and Yadav, R.S. (2020) 'State-of-the-art approach to clustering protocols in VANET: a survey', *Wireless Networks*, Vol. 26, No. 7, pp.5307–5336.

Klutho, S. (2013) *Mathematical Decision Making: An Overview of the Analytic Hierarchy Process*, Tutorial.

Lin, Y., Zhang, Z., Huang, Y., Li, J., Shu, F. and Hanzo, L. (2020) 'Heterogeneous user-centric cluster migration improves the connectivity-handover trade-off in vehicular networks', in *IEEE Transactions on Vehicular Technology*, Vol. 69, No. 12, pp.16027–16043, December, DOI: 10.1109/TVT.2020.3041521.

Liu, L., Chen, C., Qiu, T., Zhang, M., Li, S. and Zhou, B. (2018) 'A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs', *Veh. Commun.*, Vol. 13, pp.78–88.

Liu, Q., Liu, R., Wang, Z., Han, L. and Thompson, J.S. (2021) 'A V2XIntegrated positioning methodology in ultradense networks', in *IEEE Internet of Things Journal*, Vol. 8, No. 23, pp.17014–17028, 1 December, DOI: 10.1109/JIOT.2021.3075532.

Nguyen, V.D., Kim, O.T.T., Dang, D.N.M., Kim, S.S. and Hong, C.S. (2015) 'Application of the lowest-ID algorithm in cluster-based TDMA system for VANETs', *2015 International Conference on Information Networking (ICOIN)*, pp.25–30, DOI: 10.1109/ICOIN.2015.7057851.

Pandey, P.K., Kansal, V. and Swaroop, A. (2022) 'OCSR: overlapped cluster-based scalable routing approach for vehicular ad hoc networks (VANETs)', *Wireless Communications and Mobile Computing*, Article ID: 1426815, Vol. 2022, 15p, https://doi.org/10.1155/2022/1426815.

Qureshi, K.N. et al., (2021) 'Nature-inspired algorithm-based secure data dissemination framework for smart city networks', *Neural Comput. Appl.*, Vol. 33, pp.10637–10656, DOI: 10.1007/s00521-020-04900-z.

Ren, M.J., Zhang, M., Khoukhi, L., Labiod, H. and Vèque, V. (2018) 'A unified framework of clustering approach in vehicular ad hoc networks', in *IEEE Transactions on Intelligent Transportation Systems*, Vol. 19, No. 5, pp.1401–1414, May, DOI: 10.1109/TITS.2017.2727226.

Ren, M., Zhang, J., Khoukhi, L., Labiod, H. and Vèque, V. (2021) 'A review of clustering algorithms in VANETs', *Annals of Telecommunications*, pp.66–76, DOI: 10.1007/s12243-020-00831-x.

Saidi, A., Benahmed, K. and Seddiki, N. (2020) 'Secure cluster head election algorithm and misbehavior detection approach based on trust management technique for clustered wireless sensor networks', *Ad Hoc Networks*, 1 September 2020, Article ID: 102215, Vol. 106, No. 2, p.102215.

Sellami, L. and B. Alaya, (2021) 'SAMNET: self-adaptative multi-kernel clustering algorithm for urban VANETs', *Veh. Commun.*, June, Vol. 29, No. 3, p.100332.

Senouci, O., Aliouat, Z. and Harous, S. (2019) 'MCA-V2I: a multi-hop clustering approach over vehicle-to-internet communication for improving VANETs performances', *Futur. Gener. Comput. Syst.*, Vol. 96, No. 7, pp.309–323.

Shah, Y.A. et al. (2022) 'An evolutionary algorithm-based vehicular clustering technique for VANETs', in *IEEE Access*, Vol. 10, pp.14368–14385, DOI: 10.1109/ACCESS.2022.3145905.

Srivastava, A., Prakash, A. and Tripathi, R. (2020) 'Fuzzy-based beaconless probabilistic broadcasting for information dissemination in urban VANET', *Ad Hoc Networks*, Article ID: 102285, Vol. 108, No. 10, p.102285.

Ucar, S., Ergen, S.C. and Ozkasap, O. (2013) 'VMaSC: vehicular multi-hop algorithm for stable clustering in vehicular ad hoc networks', *IEEE Wirel. Commun. Netw. Conf. WCNC*, pp.2381–2386.

Verma, A., Kumar, S., Gautam, P.R., Rashid, T. and Kumar, A. (2020) 'Fuzzy logic based effective clustering of homogeneous wireless sensor networks for mobile sink', *IEEE Sensors Journal*. p.1, DOI: 10.1109/JSEN.2020.2969697.

Wang, H., Liu, R.P., Ni, W., Chen, W. and Collings, I.B. (2015) 'VANET modeling and clustering design under practical traffic, channel and mobility conditions', in *IEEE Transactions on Communications*, March, Vol. 63, No. 3, pp.870–881, DOI: 10.1109/TCOMM.2015.2388575.

Wang, S.S. and Lin, Y.S. (2013) 'PassCAR: a passive clustering aided routing protocol for vehicular ad hoc networks', *Comput. Commun.*, Vol. 36, No. 2, pp.170–179.

Xu, Z., Li, X., Zhao, X. and Zhang, M.H. and Wang, Z. (2017) 'DSRC versus 4G-LTE for connected vehicle applications: a study on field experiments of vehicular communication performance', *Journal of Advanced Transportation*, Volume 2017, Article ID: 2750452, pp.1–10, http://doi.org/10.1155/2017/2750452.

You, X., Wang, C-X. and Liang, Y-C. (2020) 'Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts', *Science China Information Sciences*, Vol. 64, No. 1, pp.1–74.

Zhang, D., Ge, H., Zhang, T., Cui, Y.Y., Liu, X. and Mao, G. (2019) 'New multi-hop clustering algorithm for vehicular ad hoc networks', *IEEE Trans. Intell. Transp. Syst.*, Vol. 20, No. 4, pp.1517–1530.

Zhang, D., Gong, C., Zhang, T. et al. (2021) 'A new algorithm of clustering AODV based on edge computing strategy in IOV', *Wireless Netw*., Vol. 27, pp.2891–2908, https://doi.org/10.1007/s11276-021-02624-z.

Zhang, L. and El-Sayed, H. (2012) 'A novel cluster-based protocol for topology discovery in vehicular ad hoc network', *Procedia Comput. Sci.*, Vol. 10, pp.525–534.

Zhang, Z., Boukerche, A. and Pazzi, R.W. (2011) 'A novel multi-hop clustering scheme for vehicular ad-hoc networks', *MobiWac'11 – Proc. 9th ACM Int. Symp. Mobil. Manag. Wirel. Access, Co-located with MSWiM'11*, pp.19–26.

Zhao, H., Tang, J., Adebisi, B., Ohtsuki, T., Gui, G. and Zhu, H. (2022) 'An adaptive vehicle clustering algorithm based on power minimization in vehicular ad-hoc networks', in *IEEE Transactions on Vehicular Technology*, March, Vol. 71, No. 3, pp.2939–2948, DOI: 10.1109/TVT.2021.3140085.