IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 June 2022, accepted 7 July 2022, date of publication 14 July 2022, date of current version 20 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3190964

RESEARCH ARTICLE

An Intelligent Machine Learning Based Routing Scheme for VANET

KHALID KANDALI[1], LAMYAE BENNIS[2], OMAR EL BANNAY[3], AND HAMID BENNIS[1]
[1]IMAGE Laboratory, SCIAM Team, Graduate School of Technology, Moulay Ismail University, Meknes 50040, Morocco
[2]Faculty of Sciences of Meknes, Moulay Ismail University, Meknes 11201, Morocco
[3]Laboratory of Science and Technology for the Engineer, LaSTI-ENSA, Sultan Moulay Slimane University, Khouribga 25000, Morocco

Corresponding author: Khalid Kandali (kandali.khalid@gmail.com)

ABSTRACT Today, Vehicular Ad-hoc Networks (VANET) have become an interesting research topic for developing Intelligent Transport Systems. In urban environments, vehicles move continuously and at different speeds, which leads to frequent changes in the network topology. The main issue faced in an urban scenario is the performance of routing protocols when delivering data from one vehicle to another. This paper introduces ECRDP, an Efficient Clustering Routing approach using a new clustering algorithm based on Density Peaks Clustering (DPC) and Particle Swarm Optimization (PSO). First, the PSO algorithm is applied to determine the cluster heads, or a new fitness function for finding the best solutions is formulated using the DPC algorithm. Next, clustering is performed based on the reliability of links parameter between vehicles. Then, a maintenance phase is proposed to update the cluster heads and redistribute the vehicles in the clusters. Finally, the effectiveness of the suggested scheme is evaluated by a simulation operated by MATLAB on a real urban scenario. The results achieved show an overall increase in stability, proven by a reduction in change rate by 74%, and an improvement in performance indicated by an increase in intra-cluster throughput by 34% and inter-cluster by 47%, as well as an overall reduction of average delay by 16%.

INDEX TERMS Clustering, density peak clustering, machine learning, particle swarm optimization, routing, vehicular ad-hoc NETwork.

I. INTRODUCTION

In recent years, Vehicular Ad-hoc Networks have received increased interest from scientists and engineers worldwide [1], [2]. This consideration is due to the importance of VANET in solving traffic and safety problems and in improving entertainment facilities in Intelligent Transport Systems (ITS) [3]–[5]. VANET have made rapid progress due to the development of communication, electronic and computer technologies. VANET aim to make vehicles an intelligent means of transport that collect, transfer and present information through Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communication [6]–[8].

However, due to the large increase in the number of vehicles and random mobility in urban environments, VANET suffer from various problems in terms of availability, scalability and overall network stability [9], [10]. This still affects the efficiency of many services such as routing.

According to several studies [11]–[13], clustering remains one of the best methods to overcome the difficulties encountered and improve the reliability and scalability of VANET in urban scenarios. The clustering technique is a virtual distribution of vehicles into groups called clusters [14]. This offers through Cluster Heads (CHs) and Cluster Members (CMs) a better use of the network resources, and provide a more reliable and secure routing [15].

This paper proposes a routing approach using a new clustering method based on the Particle Swarm Optimization and Density Peaks Clustering algorithm. First, the PSO algorithm is applied to determine the cluster heads, where a new fitness function for finding the best solutions is formulated using the DPC algorithm. Second, clustering is performed based on the reliability of links parameter between vehicles. Third, a maintenance phase is proposed to update the cluster heads

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

and redistribute the vehicles in the clusters. The main contributions of this paper are summarized as follows:

- The proposed approach uses a new clustering method that combines the two algorithms PSO and DPC to guarantee the stability of clusters in an urban environment;
- To select the cluster heads, the PSO algorithm is introduced. PSO provides a global search capability to find K optimal solutions. These solutions are automatically considered as CHs in our approach. In this sense, we reconstruct the fitness function that is the basis of the PSO algorithm by introducing density and distance parameters (ρ_i and δ_i) into the original form. These two parameters are obtained according to the parameter d_c;
- The determination of the cut-off parameter value d_c is a challenge in the execution of the DPC algorithm. This value influences the clustering results. In this work, we introduced a criterion factor T to calculate the value of d_c. This factor takes into consideration the position, speed and direction of the nodes;
- The classification of vehicles in the clustering phase is performed on the basis of link reliability (*REL*), and not on the distance between vehicles as in classical clustering algorithms;
- A maintenance phase is presented due to the high density and mobility of the vehicles. This phase deals with the re-selection of CHs and the clustering of vehicles.

Thus, the rest of the paper is organized in the following paragraphs. In Section 2, a literature review is given discussing the main works related to our topic. Section 3 gives the theoretical background of the methods exploited in this work. In Section 4, we detail the main steps developed in our approach. The evaluation of the effectiveness and the results of comparison with GAPC [16], and NMDP-APC [17] are discussed in Section 5. Finally, the conclusion of this paper is provided in Section 6.

II. LITERATURE REVIEW

Several research works attempt to solve the impact of density in an urban environment on the routing process in VANET by using clustering technique. There are works that have proposed solutions based on intelligent methods, while other works are focused on clustering processes and algorithms.

A. INTELLIGENT ROUTING

Several intelligent techniques (such as Machine Learning (ML), Artificial Intelligence (SI), Deep Learning (DL), Fuzzy Logic (FL), etc.) have recently been proposed for designing intelligent routing systems [18]. The motivation behind such approach lies in the ability and high potential of these methods to solve various challenges in ad hoc networks (MANET, VANET, FANET, ...) such as end-to-end delay, packet delivery rate, route stability, energy consumption and routing overhead [19]–[21]. With all this in mind, paper [22] presents a detailed survey of the main advantages and areas of use of the above techniques in VANET.

Different intelligent methods are used in [16], [17], [23]–[25] in order to provide intelligent clustering for VANET, such as Whale-inspired optimization, PSO algorithm, Grasshopper Optimization, Bio-inspired metaheuristic framework and Global Affinity Propagation Clustering.

The Affinity propagation clustering has been the principal idea operated in [16], [17] to perform clustering in V2V communication. In the paper [16], the algorithm relies on mobility parameters and communication-related parameters to overcome the difficulties caused by high vehicle mobility and road topology. The main objective of [17] is to represent the homogeneity of VANET traffic dynamics based on multidimensional parameters. This work has introduced a similarity function based on motion dynamics such as vehicle position and speed. Although the two papers [16], [17] aim to reduce communication latency, they provide algorithms that require more computation and memory.

The paper [23] has introduced a new clustering algorithm based on grasshopper optimization for VANET. This approach selects appropriate cluster heads in order to reduce network overhead in high density scenarios. Nevertheless, this work did not evaluate the performance of the algorithm in other metrics such as clustering efficiency or the number of isolated vehicles. The paper [24] has suggested a cluster routing scheme based on the PSO algorithm for V2V communication. The objective is to ensure the stability of the transmission links and to improve the routing efficiency within the cluster itself and between other clusters. But, the parameter of density, which is very important to introduce an effective clustering algorithm, has not been evoked in this scheme [26]. In [25], Husnain and Anwar have proposed a new whale-inspired optimization approach for clustering in VANET. This method integrates several parameters such as transmission area, density, speed and direction in the clustering process. However, other performance measures are needed to evaluate the effectiveness of this approach.

B. CLUSTERING ROUTING

The papers [27]–[31] have introduced several clustering algorithms for Urban environment in VANET. These works aim to provide effective clustering using different parameters such as velocity, position and direction.

In [27], the authors have introduced a new clustering algorithm that combines the spectral clustering algorithm and the force-directed algorithm. This approach aims to optimize the lifetime of all clusters and to ensure the stability of the VANET. Anyhow, the authors did not assess the effectiveness of the proposed approach in terms of other routing parameters. In [28], the authors have proposed a Unified Framework of Clustering (UFC) to improve the performance of clusters. Indeed, this work aims to increase the efficiency of cluster formation, adjust the rate of change of clusters and ensure cluster stability. Nevertheless, this method requires more computation and memory. The paper [29] has suggested a routing approach adapted to the current era of autonomous vehicles. The main goal is to decrease the network overhead during the cluster construction process and to reduce the loss of data broadcast messages. However, a maintenance step is not provided in the clustering algorithm. A stable and scalable clustering algorithm for VANET is proposed in [30]. This approach is center-based with grid partitioning and utilizes vehicle mobility characteristics by exploiting the increasing range of V2I communication. The main goal is to design a clustering based on a global view, which ensures the stability and predictivity of the decision. After all, the metrics often used to evaluate the performance in the simulation are either insufficient. A simulation with real scenario using for example SUMO returns more realistic results. In [31], a stable clustering algorithm is proposed for an urban scenario in VANET. This algorithm uses the regularity of bus traffic and vehicle mobility parameters in the clustering process. The CHs are selected based on the increment, circumcenter and centroid of an equilateral triangle. This approach employs the fixed routes of bus nodes in urban areas as a reference index, taking into account vehicle mobility such as speed, position and direction, but it ignores the density parameter, which is an important parameter in the urban environment.

Table 1 provides a qualitative comparison between the different works cited in this section.

In contrast, the peculiarity of our approach compared to the works discussed in this section is: (*i*) the presence of the notion of vehicle density in the clustering process. (*ii*) The determination of the d_c value based on a criterion value T. (*iii*) The use of the PSO algorithm to select the CHs, due to its strong global search capability. (*iv*) The integration of the notion of link reliability model to reformulate the objective function for the purpose of distributing the vehicles into clusters.

III. THEORETICAL BACKGROUNDS

In this section, we outline the theoretical basis of the proposed solution. Firstly, we discuss the PSO algorithm and its structure. Then, we move to the DPC algorithm to highlight its main aspects and strengths.

A. PARTICLE SWARM OPTIMIZATION METHOD

The PSO method belongs to the family of swarm intelligence methods. This method is often used to solve optimization problems. The PSO algorithm is a population-based search algorithm, inspired by the social behavior of birds in a flock [32], [33]. Each individual represents a particle and is considered as an expected solution. The particles fly in a hyperdimensional search space with an adaptable speed. The position and velocity of each particle change dynamically within the swarm according to its own flight experience and according to the socio-psychological tendency of the individuals. Indeed, each particle is able to memorize the best position in the search space thanks to the provision of a memory. This change is influenced by experience or knowledge. By applying this method, the search solution is obtained so as to return to the best positions in the search space.

The PSO method has several advantages, such as:

- The understanding and implementation of the method is simple.
- There are not many parameters to adjust.
- The method does not require enough computation time.
- The process is based on velocity and not only on position.
- It can converge quickly.

1) PARTICLE VELOCITY AND POSITION

The PSO algorithm considers individuals as particles in the D-dimensional search space, where the position of each particle refers to an expected solution of the problem. Each particle adjusts its position x and velocity v to find a new solution.

Let m be the total number of particles. Each particle changes its position and velocity according to the following formulas:

$$v_{id}(t+1) = w \times v_{id}(t) + c_1 r_1 [P_{ij}(t) - x_{ij}(t)] + c_2 r_2 [P_{gj}(t) - x_{ij}(t)] \tag{1}$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \tag{2}$$

where $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, D$; w denotes the inertia weight; $v_{id}(t)$ is the velocity of particle i at iteration j; $x_{ij}(t)$ is the position of particle i at iteration j; c_1 and c_2 are the learning factors; r_1 and r_2 are random values between 0 and 1; $P_{ij}(t)$ is the current optimal position that particle i has been searching for; $P_{gj}(t)$ is the current optimal position of the particle swarm as whole.

The velocity update formula contains three parts: The first part refers to the degree of confidence of the particle in its current movement. It represents the product of the inertial weight and the current velocity of the particle. In other words, this part is calculated based on the inertial motion of the original velocity. The second part shows the situation of personal influence, i.e., the judgment of the particle on its own history. The third part shows the social influence. It deals with the mutual cooperation and information sharing of the particles.

2) ALGORITHM STEPS

The main steps of the PSO algorithm are described in the flowchart in Fig. 1.

The PSO algorithm creates the population of particles uniformly distributed in the search space, and finds the optimal position of each particle. Then, the particles update their positions and velocities based on their individual endpoints and the previous optimal position. Finally, the algorithm finds the optimal solution by exchanging information between the particles. The execution of the algorithm is iterative, i.e., it repeats until the stopping criteria are satisfied. This reinforces the overall convergence of the algorithm.

B. DENSITY PEAK CLUSTERING ALGORITHM

Density-based clustering algorithms have attracted more attention recently and have been widely applied in several

TABLE 1. A qualitative comparison of clustering routing schemes.

Work	Algorithms used	Parameters used	Cluster maintenance	Real scenario	Simulators	Performances
[16]	-Affinity Propagation Clustering.	- Velocity. - Position. - Owned communication. - Required communication rate. - Normal neighbour list.	- Yes	- Yes	- MATLAB	- Average throughput of CHs. - Average throughput. - Average packet loss rate. - Average packet delay. - Average duration of CHs. - Average duration of CMs. - Change rate of CHs. - Number of CHs.
[17]	- Affinity Propagation Clustering.	- Position. - Speed.	- No	- Yes	- MATLAB	- Minimum number of iterations. - Number of clusters.
[23]	- Grasshoppers' optimization.	- Position. - Direction. - Velocity.	- No	- No	- MATLAB	- Number of clusters.
[24]	- Particle swarm optimization.	- Route particle. - Velocity coding rules. - Iteration rules. - Fitness function.	- Yes	- Yes	- NS3. - SUMO.	- PDR. - Delay.
[25]	- Whale Optimization Algorithm.	- Transmission range. - Node density. - Speed. - Direction.	- No	- No	- Not mentioned	- Load Balance Factor. - Number of clusters.
[27]	- Spectral clustering algorithm. - Force-directed algorithm. - Whole cluster stability.	- Velocity. - Position	- No	- No	- SUMO. - MOVE.	- Average cluster lifetime. - Cluster size.
[28]	- Unified framework of clustering.	- Relative position. - Relative velocity. - Link lifetime.	- Yes	- No	- NS2. - SUMO.	- CH duration. - CM duration. - Number of clusters. - Clustering efficiency. - Number of initial CHs. - CM disconnection rate. - Average role change rate. - CM re-clustering delay. - CM re-clustering success ratio.
[29]	- Path-based Clustering.	- Path information.	- No	- No	- OMNeT++. - SUMO.	- Cumulative Survived Cluster Member Rate. - Total Overhead. - Cumulative Data . Dissemination Rate.
[30]	- Center-based stable evolving clustering.	- Vehicles' mobility. - Grid partitioning.	- Yes	- No	- MATLAB	- Clustering efficiency. - Average cluster head (CH) duration. - Average cluster member duration. - Number of clusters.
[31]	-Stable clustering algorithm.	- Velocity. - Position. - Direction. - Traffic regularity of buses.	-Yes	- Yes	- NS3. - SUMO.	- Cluster head changes.
Proposed approach	- Particle swarm optimization. - Density peaks clustering.	- Node density. - Position. - Velocity. - Link reliability.	- Yes	- Yes	- MATLAB.	- CH lifetime. - CM lifetime. - Average number of clusters. - Clustering efficiency. - PDR. - Delay.

domains [34]. These algorithms detect high density regions in the data space to form clusters. A set of points is treated as a cluster if the data points are densely connected. Among the main density-based clustering algorithms, we cite the Density Peak Clustering algorithm proposed by [35]. It is a clear and easy to implement algorithm that features new ideas. It is characterized by a relatively low computational complexity, and does not require a large number of control parameters. The DPC selects centers of varying sizes manually on the basis of a generated decision graph. The points with the highest density are predicted to be centers. Each selected centroid is the farthest point from the other high-density points. For this purpose, the DPC algorithm uses the cut-off distance parameter d_c to generate the decision graph,

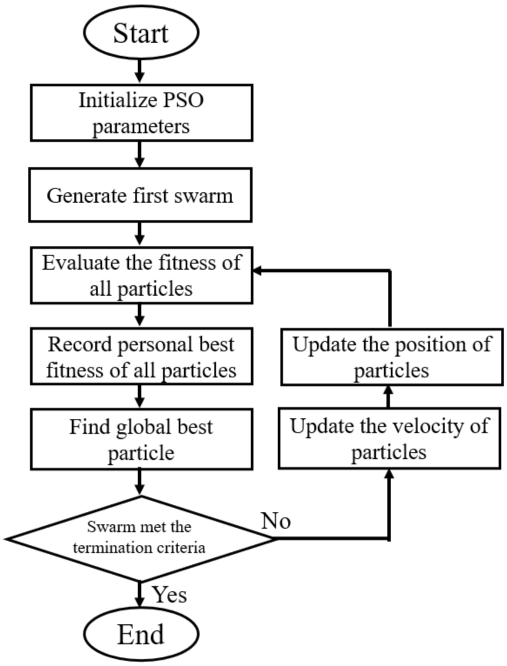


FIGURE 1. **Flow chart of the PSO algorithm.**

with the aim of distinguishing the density level in terms of the distance between the data points. Thus, the performance of DPC necessarily depends on the way the d_c parameter is assigned a value. Each data point in the decision graph is identified by two parameters: the local density ρ_i, and the minimum distance to other high-density data points δ_i. The DPC algorithm contains three phases: (*i*) the density and distance calculation, (*ii*) the decision graph creation, (*iii*) the clustering process.

1) DENSITY AND DISTANCE CALCULATION

The density ρ_i for each data point i is given based on the distance between the data points. It is calculated by the following formula:

$$\rho_i = \sum \chi(d_{ij} - d_c) \tag{3}$$

With

$$\chi(d_{ij} - d_c) = \begin{cases} 1 & (d_{ij} - d_c) < 0 \\ 0 & (d_{ij} - d_c) \geq 0 \end{cases}$$

D_{ij} indicates the Euclidean distance between two data points i and j; d_c represents the cut-off distance. In general, the value of d_c represents 1% or 2% of the total number of data points. The minimum distance δ_i from each data point i to a higher local density point is calculated using the following formula:

$$\delta_i = \min_{j:\rho_j > \rho_i} (d_{ij}) \tag{4}$$

The value δ_i for each data point i indicates the shortest distance from any other data point with a higher density value than i.

2) DECISION GRAPH

The second step in the DPC algorithm is to generate the decision graph in an autonomous way. This is a new method to identify the centroid of each data cluster [35]. This graph is based on the values of ρ_i and δ_i of each data point already calculated in the previous step. As shown in Fig. 2 (b), the diagram consists of two axes: the horizontal axis is drawn with the density value ρ. The vertical axis represents the distance values δ. According to these two parameters, the DPC algorithm selects the centers of the clusters. Therefore, points with large density and distance values will be considered as next centers. Fig. 2 shows the centroid selection process using the decision diagram method. Fig. 2 (a) shows the distribution of the data points. As it is remarkable in Fig. 2 (b), points 1 and 10 are located in the upper right corner of the decision diagram, i.e. They have the largest local density and distance value compared to the other points. Therefore, the two points 1 and 10 will be considered as centers of clusters.

3) CLUSTERING PROCESS

The DPC algorithm can be implemented easily. It works on simple principles, and also it can form clusters in a short time. It is based on the decision diagram in the selection of cluster centers. It is a method based on qualitative selection instead of using quantitative analysis. As such, each point in the diagram has a local density and a minimum distance. Points with a relatively large local density value and distance are selected as cluster centers. A large value of local density and distance means that the points concerned are intrinsically dense and they are surrounded by neighbors with relatively high density and they are located at a relatively large distance from other points with higher density. Once the cluster centers are selected, the DPC algorithm assigns the remaining points to the cluster of the highest density point closest to them. The process continues until the final clustering result is acquired.

IV. PROPOSED SCHEME (ECRDP)

In this section, a new clustering algorithm for the routing process is proposed for an urban scenario. This approach exploits the advantages of both PSO and PDC algorithms.

First, the PSO algorithm is applied to determine the initial cluster heads. A new Fitness function based on the DPC algorithm is introduced. Then, we propose the clustering phase and finally the maintenance phase.

A. NETWORK DESCRIPTION

An urban road traffic environment with V2V and V2I communication is considered. To perform clustering, several types of information are needed such as topological information, or mobility information such as position, speed and direction. This information can be obtained locally using the V2V architecture or globally using the V2I architecture, as shown in Fig. 3. In this network, vehicles are densely populated and aligned on the roads considering the following:

We consider N vehicles in the network.

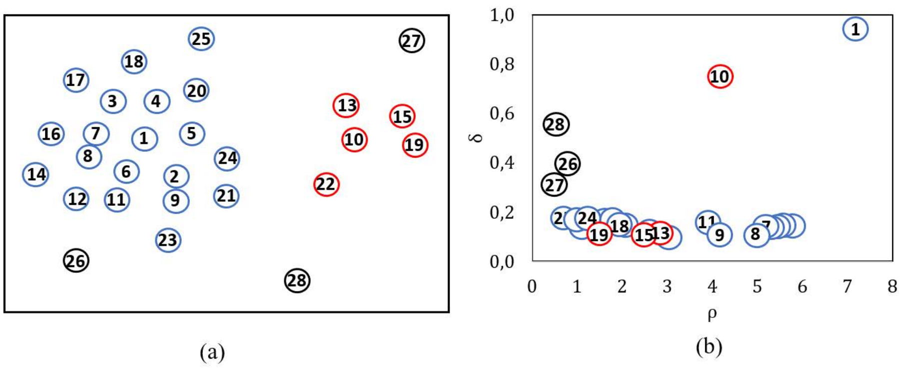


FIGURE 2. An example of DPC's decision graph: (a) Data point distribution; (b) Decision graph.

Each vehicle defines a transmission area called R thanks to the DSRC (Dedicated short-range communication) technology.

Each vehicle is equipped with a GPS system to obtain mobility parameters such as position, speed and direction, and stores these parameters in its routing table. The vehicle periodically exchanges the details of its table with other adjacent vehicles or with the RSU, through HELLO packets.

B. PROPOSED SCHEME

In VANET, the constant movement of vehicles in the field influences the efficiency and performance of routing protocols, as well as the stability of links. Therefore, the number of neighbors, link quality and speed are considered as factors that impact the segmentation of vehicles into subsets. Thus, we were motivated to overcome these difficulties by integrating Link Reliability model, and the PSO algorithm, an iterative particle movement inspired.

The major contribution of this work lies in the use of the advantages of the DPC algorithm and the PSO algorithm to design a clustering mechanism adapted to VANETs in an urban scenario. As shown in Fig. 4, the proposed scheme addresses the following needs:

First, the routing efficiency depends on the quality and stability of the clusters and the selection of the cluster heads. The cluster head (CH) is responsible for data transmission and cluster management, and if it leaves a cluster, the routing process does not work properly. So, the proper selection of cluster heads remains the first challenge solved by this paper.

Then, when creating clusters, distance is often the parameter to consider. However, in an urban environment characterized by a high density of vehicles, the notion of link reliability influences the stability of clusters and becomes a critical parameter. Therefore, this paper shows how to introduce the link reliability parameter in the cluster formation process.

Finally, due to the high mobility of the vehicles in the VANET network, it is possible that the clusters lose their organization. For this, a maintenance phase is highly recommended to update the clusters.

1) CLUSTER HEAD ELECTION

The density of vehicles in the urban communication environment has an impact on the stability of clusters and the performance of routing protocols. Several parameters such as direction, number of neighbors, speed and position characterize the distribution of vehicles. Therefore, the selection of stable cluster heads is correlated with these factors. The transmission of data between clusters that have appropriate cluster heads ensures good connectivity and guarantees the stability of the links between vehicles. Consequently, the selection of cluster heads is among the most important challenges in the clustering process in VANETs.

To maintain cluster stability, three factors are considered in the cluster head (CH) election process. First, each selected cluster head must have a similar direction to the other vehicles that belong to the cluster. Second, the value of speed and position must be similar to the values of other vehicles. Finally, the cluster head must have an adequate density and position with respect to the other cluster members. Hence, the PSO algorithm is applied to solve this challenge.

In the next section, using the DPC algorithm, we explain the creation of the new fitness function for the PSO algorithm, which will help us select the CHs. This function is based on the three factors discussed above.

First, for each vehicle, we start by calculating a criterion factor T based on which we obtained the value of the cut-off distance parameter d_c used by DPC algorithm. Next, we calculated the density and the distance values ρ_j and δ_i for each vehicle. Then, the fitness function F is obtained. Finally, the position and velocity of each vehicle are updated. Conequently, the vehicles with optimal position are selected as CHs.

a: PARAMETERS SETTING

According to several uses of this algorithm, the value of d_c is almost 1% or 2% of the total size of the set of nodes. According to several experiments, the choice of this value has a great impact on the clustering results.

In [26], a new method for calculating the d_c value is introduced in order to improve the clustering results. This method contains the following steps:

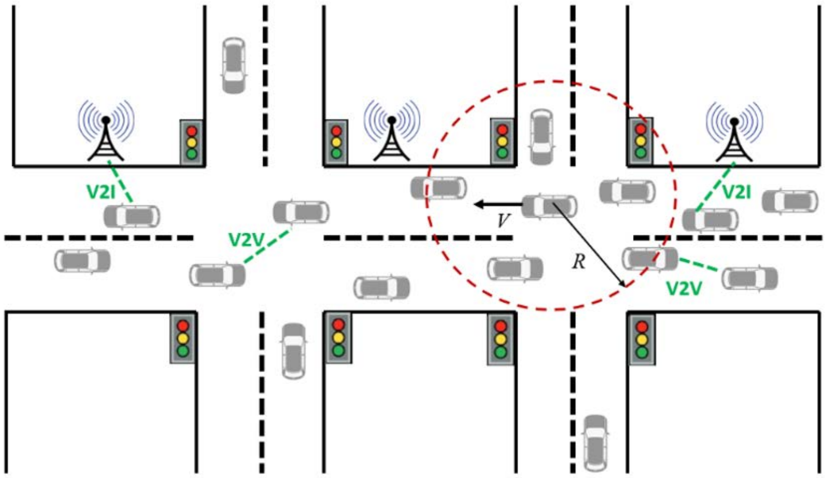


FIGURE 3. Network Model.

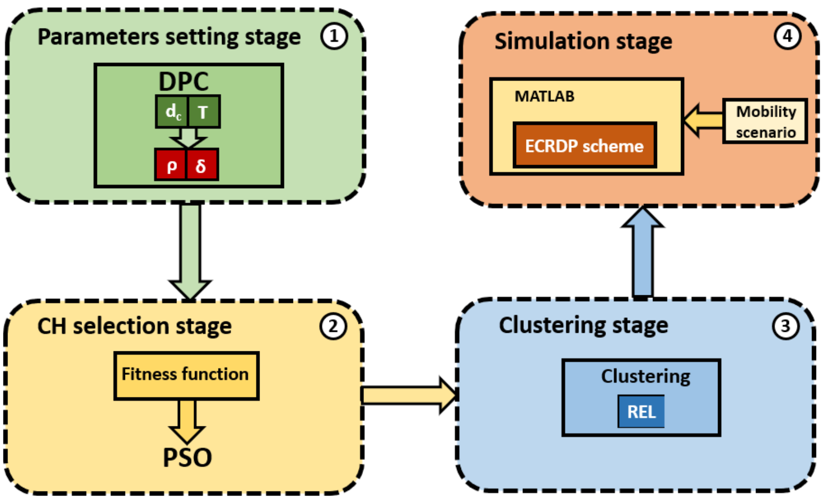


FIGURE 4. The stages of the proposed scheme.

- Calculate the Gaussian distance between each two nodes:

$$Distance = 1 - e^{-\frac{d_{ij}^2}{2}} \tag{5}$$

where d_{ij} is the Euclidean distance between node i and node j.

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{6}$$

- Obtain the minimum Gaussian distance $Distance_{min}$ and the maximum Gaussian distance $Distance_{max}$.
- Calculate the d_c value using the following formula:

$$d_c = \frac{Distance_{min} + Distance_{max}}{2} \tag{7}$$

However, in this paper, we propose a new method to calculate the d_c value. Our formula is based on finding by [26], where we exchanged Gaussian distance with a criterion factor T which takes into consideration the velocity, the average Euclidean distance, and the total number of node i neighbors. The motivation behind our formula improvement stems from the fact that the VANET environment is characterized by high mobility of vehicles, and the quality of the link between vehicles which in turn depends not only on the distance, but also on other mobility parameters such as velocity and number of neighbors. The following formula is used for factor T calculation:

$$T_{ij} = 1 - e^{-\frac{d_{ij}^2}{2}} \tag{8}$$

To calculate the value of d_{ij}, the following formula is used:

$$d_{ij} = \frac{1}{V_{ij} + 0.1} \times \frac{1}{S_{ij} + 0.1} \times \frac{N_i}{N} \tag{9}$$

Or V_{ij} is the standard deviation of the velocity of node i and its neighbor j. It is calculated by formula (10); S_{ij} is the average Euclidean distance between node i and its neighbors, as mentioned by formula (11); N_i the total number of neighbors of node i and N is the total number of nodes.

$$V_{ij} = \sqrt{\frac{\sum_{j=1}^{N_i}(|v_i| - |v_{ij}|\cos\theta_{ij})^2}{N_i + 1}} \tag{10}$$

$$S_{ij} = \frac{\sum_{j=1}^{N_i}\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{N_i + 1} \tag{11}$$

$|V_{ij}|$ represents the absolute velocity of node j neighboring node i, with $i \in [1, N_i]$. θ_{ij} is the direction angle between node i and node j. (x_i, x_j) and (y_i, y_j) are the updated positions of nodes i and j respectively.

From formula (7), we can calculate the value d_c:

$$d_c = \frac{T_{min} + T_{max}}{2} \tag{12}$$

Thus, we express the density for each node i by the following formula:

$$\rho_i = \sum_j e^{-(\frac{T_i}{d_c})^2} \tag{13}$$

b: FITNESS FUNCTION

In this paper, the d_c value is not obtained according to an empirical value, but it is calculated according to the formula (12), which takes into account the speed and the position of the nodes. Also, it is remarkable that the formula for calculating the d_c value in the classical DPC algorithm returns a discrete value, while the new formula proposed in this paper gives a continuous value. This new calculation method takes into consideration the VANET specifications on the one hand, and on the other hand, it decreases the probability of several nodes having the same density value, i.e., the probability of having a conflict is very low. From all the above, the fitness function is designed as follows:

$$F = \frac{1}{\rho \times \delta} = \frac{1}{\sum_j e^{-(\frac{T_i}{d_c})^2} \times \lim_{j:\rho_j > \rho_i}(T_{ij})} \tag{14}$$

where T_{ij} is the criterion factor of node i, d_c is the cut-off distance calculated by formula (12); Generally, in DPC, $\delta = min(T_{ij})$ when it is a general node, but for a node with a large density we take $\delta = max(T_{ij})$. A node has a high probability of being cluster head when it has a small value of F.

To guarantee high performance of the clustering algorithm, a convergence condition is introduced to determine the number of iterations of execution of PSO during the cluster head selection step. This convergence is defined in the following formula:

$$|F_n - F_{n-1}| \leq \epsilon, \quad n \geq 2 \tag{15}$$

where ϵ represents the convergence parameter and n indicates the number of iterations. The PSO algorithm reaches convergence after a certain number of iterations if the difference between F_n and F_{n-1} is very small.

c: POSITION AND VELOCITY UPDATE

In the PSO algorithm, each particle is defined by a position and a velocity. It updates these values periodically. The best positions obtained at the end of the algorithm will be considered as initial cluster heads. The update of the position and velocity is done according to the formulas (1) and (2) mentioned in Section 3. But, the values (w, c_1, c_2, r_1, r_2) are defined according to the method proposed in the paper [26]. This method contains the following steps:

1. Calculate the inertia weight w. This value affects the exploration and exploitation of the search space by ensuring the gradual slowing down of the particles [36]. The higher w is, the lower the slow-down effect is for a given particle. According to the work [36], the linear decrease of the value$w$ returns good results. That is, the particles exploit the good region without overshooting the positions too much if the inertia weights are low. Based on the work [26], the decrease in the value of w is given by the following function:

$$w = w_{max} - \frac{t \times (w_{max} - w_{min})}{t_{max}} \tag{16}$$

w_{max} indicates the final value of the inertia weight, w_{min} is the initial value of the inertia weight, t represents the number of iterations, t_{max} indicates the maximum number of iterations.

2. The second step defines the values of c_1 and c_2. These two factors are called acceleration coefficients. The paper [37] proposed a new version of the time-varying acceleration coefficient PSO algorithm (TVAC-PSO). In this work, the authors suggested a method to determine the two acceleration factors. It starts with c_1 and c_2 such as $c_1 > c_2$, then it gradually and linearly decreases and increases respectively. Also, according to the paper [37], the ranges of values of c_1 and c_2 are as follows: c_1 decreases linearly from 2.5 to 0, while c_2 increases linearly from 0 to 2.5. This linear change is given by the following two formulas:

$$c_1(t+1) = (c_{1,final} - c_{1,initial}) \times \frac{t}{t_{max}} + c_{1,final} \tag{17}$$

$$c_2(t+1) = (c_{2,final} - c_{2,initial}) \times \frac{t}{t_{max}} + c_{2,final} \tag{18}$$

Note that $c_{initial}$ is the initial value of the acceleration coefficient, c_{final} is the final value of the acceleration coefficient, t is the number of iterations, t_{max} is the maximum number of iterations.

3. r_1 and r_2 are random values between 0 and 1.

d: CLUSTER HEAD SELECTION

In this section, a new strategy is proposed for CH selection combining PSO and DPC algorithms, the workflow of our proposition is outlined in Fig.5.

The CH selection process proposed in this approach is as follows:

1) Consider dimension of the network.
2) Split the road into two lanes.
3) Place vehicles and RSU in the network.
4) Apply PSO using DPC for selecting CH

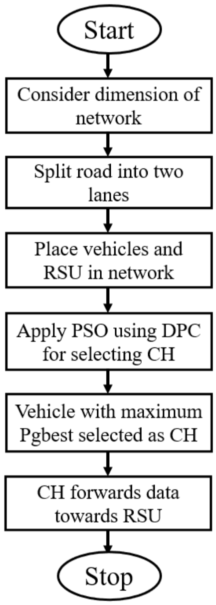


FIGURE 5. Flowchart of CH selection process.

a. Initialize vehicles

b. For each vehicle i

i. Calculate criterion factor T.

ii. Calculate the cut-off distance parameter d_c.

iii. Calculate the density and distance values ρ_j and δ_i.

iv. Compute fitness value based on formula (14).

v. If present fitness value superior than (pbest) then

1. Pbest = present fitness.

2. Pgbest is the greatest between all Pbest.

c. For each vehicle

i. Update velocity

ii. Update position

d. Verify that the maximum iterations are reached. Otherwise, Step 4.b is performed.

5) Vehicle of maximum gbest selected as CH.

6) CH forwards the data towards the RSU.

2) CLUSTER CREATION

After determining the cluster heads using both the PSO algorithm and the DPC algorithm, the second step aims to create and group the vehicles into several clusters. The clustering is done based on the link reliability (REL) as a metric, instead of the usual distance-based clustering. The reason to consider this metric is that the value of REL shows the transferability of information packets with the minimum of link failure [38]. It is a parameter that evaluates the performance and stability of the system. The Euclidean distance does not show the quality of the links between vehicles.

A CH broadcasts ELECTED Messages to all neighboring vehicles. The ELECTED Message includes the cluster head ID, position, velocity and direction of the cluster head. The Life Time indicates the time that the cluster head broadcasts the ELECTED Message. The format of this message is described in Fig.6 (a). Each node k that receives an ELECTED Message sends a JOIN Message to the CH sender. This message includes the cluster head ID, node ID, position, velocity and direction of the node. The format of the JOIN Message is described in Fig.6 (b). Otherwise, if node k receives multiple ELECTED Messages, it runs a test. Node k calculates the link reliability with all accessible CHs. This node decides to join the CH with a higher REL value by sending the JOIN Message. The reliability link between a node k and a CH is calculated using the following formula:

$$REL(link_{k,CH_i}) = Erf\left(\frac{(\frac{2R}{t} - \mu_{\Delta v})}{\sigma_{\Delta v}\sqrt{2}}\right) - Erf\left(\frac{(\frac{2R}{t} + T_p - \mu_{\Delta v})}{\sigma_{\Delta v}\sqrt{2}}\right), \quad when\ T_p > 0 \tag{19}$$

where Erf is the GAUSS error function. It is a special integer function generally used in probability and statistics and in diffusion problems. It is defined as follows:

$$Erf(x) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}\,dt \tag{20}$$

μ and σ represent the mean value and variance of the speed respectively. Δv indicates the difference in speed between two vehicles.

T_p indicates the expected time that a link L between two vehicles remains available. It is calculated according to two scenarios:

Scenario 1: Both vehicles move in the same direction (Fig.7)

$$T_p = \begin{cases} \dfrac{2R - L_{k,CH_i}}{|V_k - V_{CH_i}|} & if\ V_{CH_i} > V_k \quad with\ V_k \neq V_{CH_i} \\ \dfrac{R - L_{k,CH_i}}{|V_k - V_{CH_i}|} & if\ V_{CH_i} < V_k \end{cases} \tag{21}$$

Scenario 2: The two vehicles move in different directions (Fig.8).

$$T_p = \begin{cases} \dfrac{R + L_{k,CH_i}}{|V_k + V_{CH_i}|} & if\ the\ vehicles\ move\ towards\ each\ other \\ \dfrac{R - L_{k,CH_i}}{|V_k + V_{CH_i}|} & if\ the\ vehicles\ move\ towards\ each\ other \end{cases} \tag{22}$$

R is the transmission area, L is the Euclidean distance between vehicle k and vehicle CH_i, V_k and V_{CH_i} indicate the speed of the two vehicles respectively.

After the initialization of the network, each vehicle broadcasts HELLO packets to all the neighboring vehicles, to share and retrieve the necessary information in the formula (19). The format of the HELLO packet contains mainly the preamble, Life Time, Node ID, Node position, Node velocity, Node direction, as it is shown in Fig.6 (c).

Preamble	Life Time	Cluster head ID	Position	Velocity	Direction

(a)

Preamble	Cluster head ID	Node ID	Position	Velocity	Direction

(b)

Preamble	Life Time	Node ID	Node Position	Node Velocity	Node Direction

(c)

FIGURE 6. Format of exchanged messages: (a) Format of ELECTED Message; (b) Format of JOIN Message; (c) Format of HELLO packet.

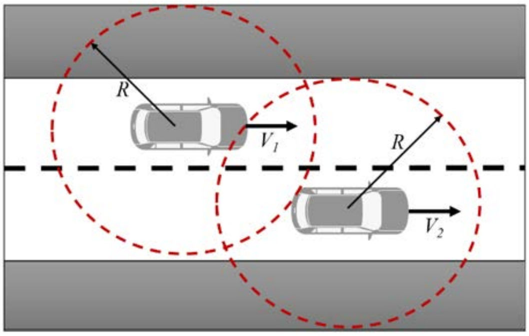


FIGURE 7. Scenario 1 to calculate T_P.

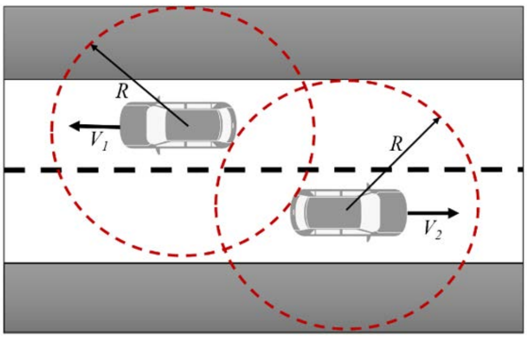


FIGURE 8. Scenario 2 to calculate T_P.

The main steps in creating clusters are:

1) Each vehicle k broadcasts HELLO messages to share mobility information (the position, speed, and direction).

2) The cluster head CH_i broadcasts ELECTED messages to all neighboring vehicles.

3) If vehicle k receives a single ELECTED message, it immediately sends a JOIN message back to CH_i to join the cluster.

4) If vehicle k receives multiple ELECTED messages, it calculates the *REL* value with all accessible cluster heads. Vehicle k sends the JOIN message to the cluster head with the largest *REL* value to join the cluster.

3) CLUSTER MAINTENANCE

The VANET network topology changes frequently due to the high speed of vehicles, which leads to several changes in the cluster structure. Therefore, we propose in this section a cluster maintenance phase. This step aims to keep the cluster structure and thus ensure the performance of the algorithm in the routing process.

The cluster maintenance phase considers three possible cases:

1- Cluster heads leave the cluster.
2- New vehicles access the clusters.
3- Vehicles leave the clusters.

If new vehicles access the clusters or vehicles leave the clusters, the cluster heads periodically broadcast ELECTED Messages to the accessible vehicles to update the cluster structure. In this way, the new vehicle can join the cluster by receiving ELECTED Message. Each vehicle that receives this type of message calculates the *REL* value with the cluster heads senders. The vehicle joins the cluster with a large *REL* value.

If one or more cluster heads leave the cluster, the cluster head election process will be executed again to select new cluster heads.

4) DESCRIPTION OF THE ECRDP

This paper proposes a new routing approach (ECRDP) based on the clustering process. This approach combines the two algorithms PSO and DPC. First, the fitness function is given by the DPC algorithm. Next, the K cluster heads are selected based on the optimal solutions found by the PSO method. Then, the clustering process is performed to create clusters. Finally, a maintenance step is proposed to update the clusters.

The clustering process proposed in this approach is as follows:

1) Initialize network.

2) Apply PSO and DPC for selecting CH.

3) Consider the K optimal vehicles given by the PSO algorithm as the initial cluster heads.

4) The Reliability Link between each vehicle to each cluster heads are calculated.

5) Each vehicle is assigned to a cluster head according to the principle of maximum *REL*.

6) Perform the cluster maintenance phase and update the cluster heads.

7) Stop the execution when the clustering results remain the same or when the end of iteration condition is reached.

In summary, steps 2 describe the PSO optimization process. The fitness function used is given by using the DPC algorithm. Steps 3 to 5 describe the clustering process. The maintenance phase is given by step 6, where the clusters undergo an update. The execution of this process continues until the clusters are stable, or the end of iteration condition is reached.

The pseudo code of ECRDP is described in Algorithm 1:

Algorithm 1 - Pseudo Code of ECRDP

Begin
1: **If** the event has happened **then**
2: Clustering using PSO and DPC.
3: Send data from source to destination.
5: **Else**
6: Do nothing.
7: **End if**
End

V. SIMULATION AND PERFORMANCE

In the following section, we describe simulation environment, as well as some key performance indicators based on which we qualify our work and compare it to the existing literature. Then, we enumerate some important results and discuss their significance.

A. SIMULATION ENVIRONMENT

In order to evaluate the performance of the proposed approach, we performed the simulation using MATLAB and comparing the effectiveness of ECRDP with two works in the literature: the first work called GAPC is a clustering scheme based on traffic regularity and mobility parameters [16], and the second work NMDP-APC is an Affinity Propagation Clustering based approach [17].

The vehicle mobility is generated on a real scenario, integrating an open-source map. The traffic area of the simulation is the city of Cologne as shown in Fig.9. This simulation scene has an area of $1Km \times Km$ and it contains a set of boulevards, streets, and intersections. Then, to study the effect of traffic volume on the network, we imported the traffic pattern into MATLAB, based on the parameters depicted in Table 2.

B. PERFORMANCE METRICS

According to the literature review presented in the second part of this paper, the performance evaluation of clustering-based routing protocols in VANETs includes the stability of the formed clusters and the routing efficiency. Thus, this simulation takes into consideration several performance metrics, in order to make it more complete and objective. These metrics are divided into two categories:

TABLE 2. Simulation parameters.

Parameter	Value
Simulation time	200 s
Square of simulation zone	1000 $m \times m$
Number of vehicles	40, 80, 120, 160, 200
Lane Number	2
Maximum speed	60 km/h
Transmission range	250 m
Packet size	64 bytes
Transmission rate	54 Mbps
Examined protocols	ECRDP, GAPC, NMDP-APC
Number of iterations of PSO	40
r1, r2	0.4, 0.6

FIGURE 9. Simulation scene.

1) CLUSTER STABILITY PERFORMANCE METRICS

Several metrics indicate the stability of clusters in VANETs. These metrics include the stability of cluster heads and the stability of cluster members. Thus, they are further divided into two sets: cluster lifetime and cluster scale:

a: CLUSTER LIFETIME

- *Cluster head lifetime.*

Cluster heads lifetime indicates the period of time a vehicle remains as a cluster head. A long lifetime reflects a high stability of the clusters. It is calculated by dividing the cluster head role assignment cumulative time by cluster head role assignment count.

- *Cluster member lifetime.*

Cluster members lifetime indicates the average period of being a cluster member. The longer this metric is, the more clusters are stable. The cluster member lifetime is obtained by dividing cluster member role assignment cumulative time by cluster role assignment count.

- *Change rate of cluster head.*

This metric represents the number of times a vehicle becomes a cluster head during a specific time period. It is obtained by dividing cluster head role assignment count by a period of time which is usually the simulation time. The smaller this value, the higher the efficiency of clustering is.

b: CLUSTER SCALE

- *Number of clusters.*

The number of clusters defines the total number of clusters formed after the execution of the clustering algorithm during

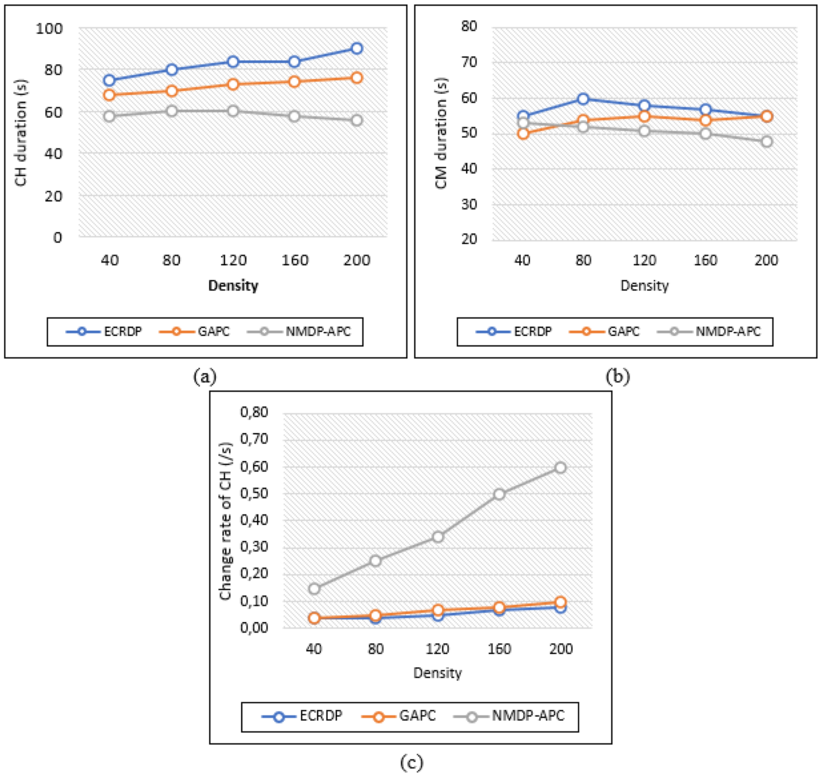


FIGURE 10. Stability performance on duration: (a) CH duration Vs Density; (b) CM duration Vs Density; (c) Change rate of CH Vs Density.

a given time period. A low number of clusters increases the clustering performance.

- *Number of isolated vehicles.*

This metric indicates the stability of clusters in terms of scale. It represents the number of vehicles that were not assigned to any cluster during the clustering process.

- *Clustering efficiency.*

The other aspect of performance is the clustering efficiency, i.e. the percentage of vehicles that participate in the clustering. It is calculated by dividing the number of vehicles participating in the clustering process by the total number of vehicles.

2) PERFORMANCE METRICS OF COMMUNICATION

The evaluation of communication performance in VANETs includes several criteria, namely throughput, delay and packet transmission success. Thus, in order to provide a thorough simulation, four metrics are considered:

- *Packet Delivery Ratio (PDR).*

PDR indicates the percentage of packet transmission in the network. It is obtained by dividing the number of packets received by the destination nodes by the total number of packets sent by the source nodes.

- *Average delay.*

Delay shows the time it takes in average for a packet to be transmitted from a source node to a destination node.

- *Throughput of clusters.*

It is the average packet transmission speed of all clusters that make up the network. It is calculated by dividing the sum of the amount of data transferred intra-clusters from all vehicles by the transfer time.

- *Throughput of cluster head.*

Throughput of CH is a metric that indicates the communication performance for CH vehicles. It is obtained by dividing the amount of data transferred inter-clusters during the transfer time.

C. RESULTS AND ANALYSIS

In this simulation, two other approaches (GAPC [16], and NMDP-APC [17]) are compared with ECRDP based on traffic density. The reason for choosing these two approaches is that GAPC and NMDP-APC are clustering-based routing solutions for VANETs, and they have the same objectives as our proposed approach.

1) STABILITY PERFORMANCE ANALYSIS

Fig.10 (a) compares the CH lifetime of three the algorithms considering different traffic densities. The CH lifetime depends necessarily on the traffic density. The lifetime increases when the number of vehicles increases. This is due to the frequent change of the network topology when the number of vehicles becomes bigger. From the figure, it can

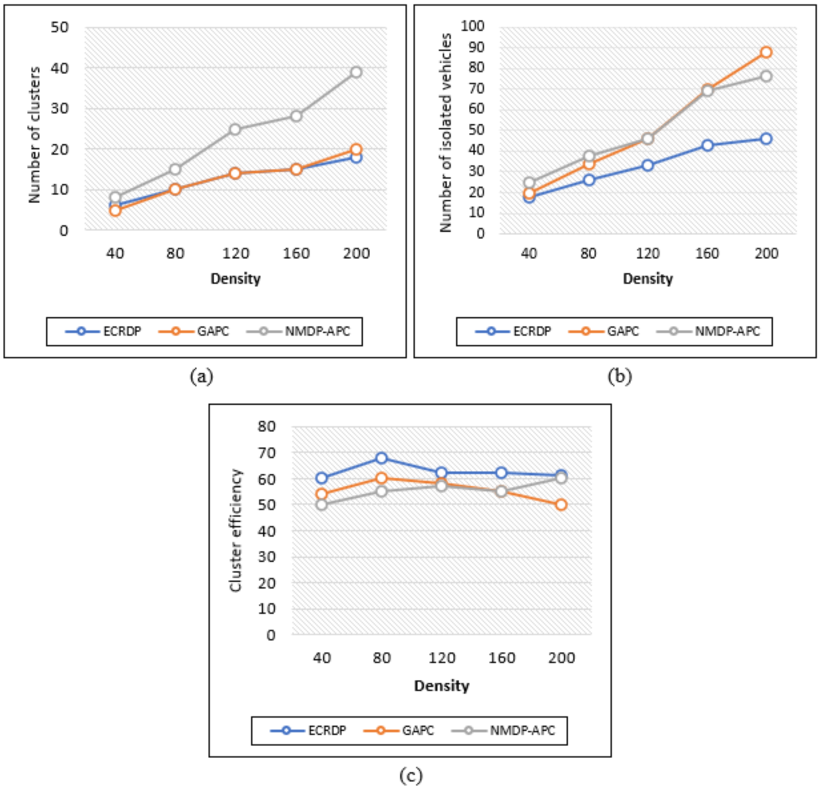


FIGURE 11. Stability performance on scale: (a) Number of clusters Vs Density; (b) Number of isolated vehicles Vs Density; (c) Cluster efficiency Vs Density.

be seen that the CH lifetime in ECRDP is high compared to GAPC and NMDP-APC in different densities. This means that the CHs remain active much longer in ECRDP than in the other approaches. A high duration of CHs guarantees the stability of clustering and improves the routing process. Therefore, in our approach, vehicles that communicate with a large number of neighbours (i.e. have a high density value) and have a suitable speed and position are selected as cluster heads using the PSO algorithm. This shows, that unstable nodes will not be considered as CHs.

According to the results described in Fig. 10 (b), the CM lifetime for ECRDP is higher compared to the other two approaches. Together with the high lifetime of the CHs, this ensures stable clustering. The CM lifetime for ECRDP keeps a value above 55*s* when the traffic level increases. In contrast, this duration does not exceed 55*s* for GAPC and NMDP-APC in different traffic densities. This is due to the fact that with the increase in the number of vehicles and with high mobility, cluster members may leave some clusters. This periodically triggers the cluster maintenance phase which allows vehicles to join other clusters. To solve this challenge caused by the permanent modification of the cluster structure, we introduced a metric called "Link reliability" which indicates the link status between vehicles during a certain time. Based on this metric, vehicles will be assigned to clusters. Therefore, the probability of having stable clusters during longer periods of time is higher.

In Fig. 10 (c), the rate of CHs update is too high for NMDP-APC, while this metric is reduced in the ECRDP and GAPC approaches. This rate increases when the number of vehicles becomes large. Consequently, the long lifetime of CHs and CMs has an impact on the rate of CHs update. This decrease in the rate indicates the stability of the clusters in a very dense traffic.

Fig.11 (a) shows statistics on the number of clusters generated during the simulation period. As observed, this number is more volatile in GAPC and NMDP-APC than in ECRDP when the number of vehicles increases. ECRDP shows a high stability of the network considering a lower number of clusters during the clustering phase. Therefore, ECRDP has a high clustering capacity with a minimum number of clusters, and this is considered an advantage over the other approaches.

According to Fig.11 (b), the number of isolated vehicles is lower for the ECRDP. Whereas the compared approaches show a larger number when the traffic is dense. This means that the clustering algorithm used in ECRDP covers a maximum number of vehicles. A large number of isolated vehicles shows the inefficiency of the clustering algorithm.

Fig.11 (c) describes the impact of traffic density on the clustering efficiency for the three approaches. The results

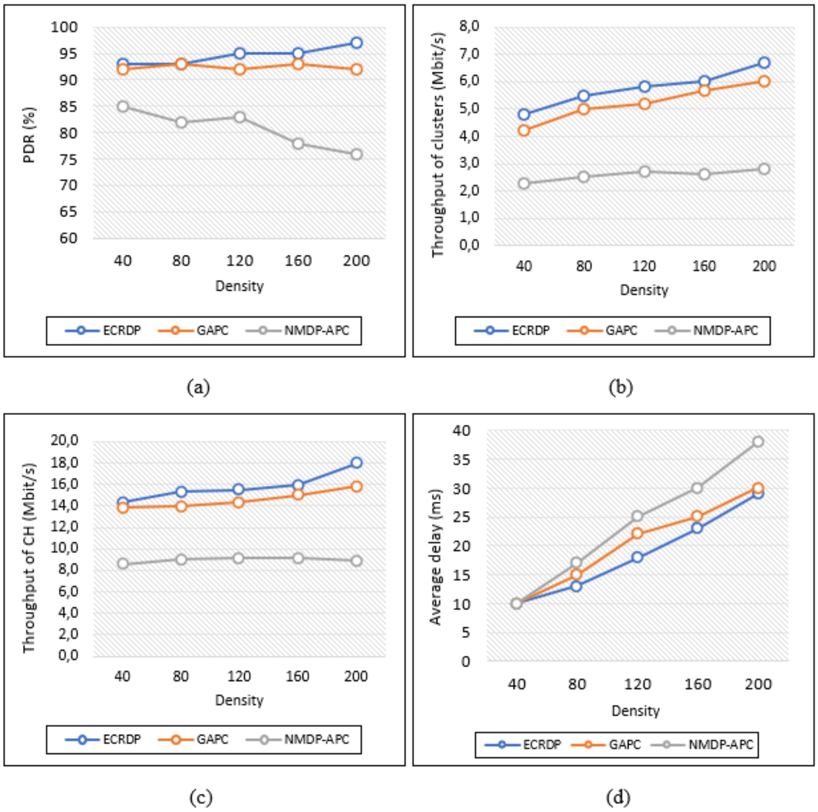


FIGURE 12. Communication performance: (a) PDR Vs Density; (b) Throughput of clusters Vs Density; (c) Throughput of CH Vs Density; (d) Average delay Vs Density.

show that the clustering efficiency in ECRDP is high compared to GAPC and NMDP-APC. In ECRDP, the efficiency exceeds 70% when the number of vehicles is 80, while the efficiency in the other two approaches does not exceed 60%. This is due to several factors such as the algorithms used cover areas with a high density of vehicles and ignore less dispersed areas. In this way, many vehicles will no longer be involved in the clustering process. To overcome this difficulty, ECRDP yields an algorithm that traverses the network environment and the set of vehicles, and uses the mobility characteristics of each vehicle to generate the different clusters.

In summary, thanks to the integration of the density notion in the fitness function formulation, and due to the proposed cluster maintenance phase, the ECRDP approach proves a great ability to maintain clustering stability in the case of a large number of vehicles.

2) COMMUNICATION PERFORMANCE ANALYSIS

Fig.12 (a) shows a comparison of the PDR for ECRDP, GAPC [16] and NMDP-APC [17]. It clearly highlights the superiority of the proposed approach in terms of PDR in dense traffic. During all the scenarios performed, the PDR of ECRDP keeps a value above 90%. This efficiency is mainly due to the stability of the clusters and the use of the mobility characteristics of the vehicles in the clustering process. Through ECRDP approach, we generate a lower number of clusters, which decreases the amount of data disseminated in the network even at high traffic density, while avoiding both network congestion, and minimizing packet drops at CHs. Therefore, ECRDP is resilient to dense traffic negative effects on PDR.

The two figures Fig.12 (b) and Fig.12 (c) show the impact of density on cluster throughput and cluster heads for the three approaches ECRDP, GAPC, and NMDP-APC. The results obtained show a performance of ECRDP in terms of throughput. This indicates the effectiveness of the proposed method in selecting appropriate cluster heads, which reduces the probability of link failures between vehicles, and increases the throughput.

Fig.12 (d) shows a comparison between ECRDP, GAPC and NMDP-APC in terms of average delay. Clearly, ECRDP provides less propagation time in low and in high density. In GAPC and NMDP-APC, a large number of clusters is generated, which increases the number of relay nodes between the source and destination node, and therefore affects the data transmission. A large number of intermediate nodes loads the bandwidth and increases the retransmission delay. To overcome this difficulty, ECRDP considers a smaller number of clusters, which improves the links between the source and destination node, and reduces the data retransmission time.

In summary, Fig.12 shows that the ECRDP approach largely improves the communication performance in VANETs for several reasons such as:

1- ECRDP approach ensures cluster better stability, by selecting appropriate vehicles to become cluster heads.

2- In the cluster formation phase, ECRDP exploits a link reliability model to form clusters.

3- A complete maintenance phase is introduced. This phase takes into consideration several cases in order to update the clusters.

VI. CONCLUSION

Due to the high density of vehicles in an urban scenario, routing is a major challenge for which clustering-based solutions have been presented in the literature. However, the stability of a clustering-based network structure can still be improved. To solve this problem, this paper suggests an efficient clustering routing scheme based on DPC and PSO for VANETs. This approach utilizes a clustering algorithm that combines the advantages of both DPC and PSO algorithms. Unlike other works in the literature, this scheme enriches the PSO algorithm with DPC algorithm to select CHs. It also accounts for the link quality instead of distance to create the clusters. Performance evaluation of ECRDP is given by comparing with the GAPC and NMDP-APC algorithms with different density combinations in an urban scenario. The results obtained prove the performance of the proposed approach to ensure cluster stability and routing efficiency.

However, we noticed that the cluster efficiency can be further improved in high density environments, and that Artificial Intelligence can be integrated in the maintenance phase to accommodate for complex traffic models.

REFERENCES

[1] A. Khalid, T. Umer, M. K. Afzal, S. Anjum, A. Sohail, and H. M. Asif, "Autonomous data driven surveillance and rectification system using in-vehicle sensors for intelligent transportation systems (ITS)," *Comput. Netw.*, vol. 139, pp. 109–118, Jul. 2018.

[2] S. Khatri, H. Vachhani, S. Shah, J. Bhatia, M. Chaturvedi, S. Tanwar, and N. Kumar, "Machine learning models and techniques for VANET based traffic management: Implementation issues and challenges," *Peer Peer Netw. Appl.*, vol. 14, no. 3, pp. 1778–1805, 2020.

[3] Y. Tang, N. Cheng, W. Wu, M. Wang, Y. Dai, and X. Shen, "Delay-minimization routing for heterogeneous VANETs with machine learning based mobility prediction," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3967–3979, Apr. 2019, doi: 10.1109/TVT.2019.2899627.

[4] R. Ghebleh, "A comparative classification of information dissemination approaches in vehicular ad hoc networks from distinctive viewpoints: A survey," *Comput. Netw.*, vol. 131, pp. 15–17, Feb. 2018.

[5] O. A. Wahab, A. Mourad, H. Otrok, and J. Bentahar, "CEAP: SVM-based intelligent detection model for clustered vehicular ad hoc networks," *Expert Syst. Appl.*, vol. 50, pp. 40–54, May 2016.

[6] M. Afrashteh and S. Babaie, "A route segmented broadcast protocol based on RFID for emergency message dissemination in vehicular ad-hoc networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 16017–16026, Dec. 2020, doi: 10.1109/TVT.2020.3041754.

[7] K. Kandali and H. Bennis, "An efficient routing protocol using an improved distance-based broadcasting and fuzzy logic system for VANET," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 6, pp. 81–93, Dec. 2020.

[8] W. Li, W. Song, Q. Lu, and C. Yue, "Reliable congestion control mechanism for safety applications in urban VANETs," *Ad Hoc Netw.*, vol. 98, Mar. 2020, Art. no. 102033.

[9] L. Wu, Q. Sun, X. Wang, J. Wang, S. Yu, Y. Zou, B. Liu, and Z. Zhu, "An efficient privacy-preserving mutual authentication scheme for secure V2V communication in vehicular ad hoc network," *IEEE Access*, vol. 7, pp. 55050–55063, 2019, doi: 10.1109/ACCESS.2019.2911924.

[10] T. S. J. Darwish, K. A. Bakar, and K. Haseeb, "Reliable intersection-based traffic aware routing protocol for urban areas vehicular ad hoc networks," *IEEE Intell. Transp. Syst. Mag.*, vol. 10, no. 1, pp. 60–73, Jan. 2018, doi: 10.1109/MITS.2017.2776161.

[11] B. K. Chaurasia, M. I. Alam, A. Prakash, R. S. Tomar, and S. Verma, "MPMAC: Clustering based MAC protocol for VANETs," *Wireless Pers. Commun.*, vol. 108, no. 1, pp. 1–28, 2019.

[12] M. Elhoseny, "Intelligent firefly-based algorithm with levy distribution (FF-L) for multicast routing in vehicular communications," *Expert Syst. Appl.*, vol. 140, Feb. 2020, Art. no. 112889.

[13] K. Kandali, L. Bennis, and H. Bennis, "A new hybrid routing protocol using a modified K-means clustering algorithm and continuous Hopfield network for VANET," *IEEE Access*, vol. 9, pp. 47169–47183, 2021, doi: 10.1109/ACCESS.2021.3068074.

[14] A. F. M. S. Shah, M. A. Karabulut, H. Ilhan, and U. Tureli, "Performance optimization of cluster-based MAC protocol for VANETs," *IEEE Access*, vol. 8, pp. 167731–167738, 2020, doi: 10.1109/ACCESS.2020.3023642.

[15] A. Katiyar, D. Singh, and R. S. Yadav, "State-of-the-art approach to clustering protocols in VANET: A survey," *Wireless Netw.*, vol. 26, no. 7, pp. 5307–5336, Oct. 2020.

[16] X. Bi, B. Guo, L. Shi, Y. Lu, L. Feng, and Z. Lyu, "A new affinity propagation clustering algorithm for V2V-supported VANETs," *IEEE Access*, vol. 8, pp. 71405–71421, 2020, doi: 10.1109/ACCESS.2020.2987968.

[17] T. Koshimizu, H. Wang, Z. Pan, J. Liu, and S. Shimamoto, "Normalized multi-dimensional parameter based affinity propagation clustering for cellular V2X," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6, doi: 10.1109/WCNC.2018.8377275.

[18] C. Lu, Z. Wang, W. Ding, G. Li, S. Liu, and L. Cheng, "MARVEL: Multi-agent reinforcement learning for VANET delay minimization," *China Commun.*, vol. 18, no. 6, pp. 1–11, Jun. 2021, doi: 10.23919/JCC.2021.06.001.

[19] M. S. Ayub, P. Adasme, D. C. Melgarejo, R. L. Rosa, and D. Z. Rodriguez, "Intelligent hello dissemination model for FANET routing protocols," *IEEE Access*, vol. 10, pp. 46513–46525, 2022, doi: 10.1109/ACCESS.2022.3170066.

[20] S.-W. Lee, S. Ali, M. S. Yousefpoor, E. Yousefpoor, P. Lalbakhsh, D. Javaheri, A. M. Rahmani, and M. Hosseinzadeh, "An energy-aware and predictive fuzzy logic-based routing scheme in flying ad hoc networks (FANETs)," *IEEE Access*, vol. 9, pp. 129977–130005, 2021, doi: 10.1109/ACCESS.2021.3111444.

[21] H. Riasudheen, K. Selvamani, S. Mukherjee, and I. R. Divyasree, "An efficient energy-aware routing scheme for cloud-assisted MANETs in 5G," *Ad Hoc Netw.*, vol. 97, Feb. 2020, Art. no. 102021.

[22] A. Mchergui, T. Moulahi, and S. Zeadally, "Survey on artificial intelligence (AI) techniques for vehicular ad-hoc networks (VANETs)," *Veh. Commun.*, vol. 34, Apr. 2022, Art. no. 100403, doi: 10.1016/j.vehcom.2021.100403.

[23] W. Ahsan, M. F. Khan, F. Aadil, M. Maqsood, S. Ashraf, Y. Nam, and S. Rho, "Optimized node clustering in VANETs by using meta-heuristic algorithms," *Electronics*, vol. 9, no. 3, pp. 1–14, 2020.

[24] X. Bao, H. Li, G. Zhao, L. Chang, J. Zhou, and Y. Li, "Efficient clustering V2V routing based on PSO in VANETs," *Measurement*, vol. 152, Feb. 2020, Art. no. 107306.

[25] G. Husnain and S. Anwar, "An intelligent cluster optimization algorithm based on whale optimization algorithm for VANETs (WOACNET)," *PLoS ONE*, vol. 16, no. 3, pp. 1–22, 2021.

[26] J. Cai, H. Wei, H. Yang, and X. Zhao, "A novel clustering algorithm based on DPC and PSO," *IEEE Access*, vol. 8, pp. 88200–88214, 2020, doi: 10.1109/ACCESS.2020.2992903.

[27] G. Liu, N. Qi, J. Chen, C. Dong, and Z. Huang, "Enhancing clustering stability in VANET: A spectral clustering based approach," *China Commun.*, vol. 17, no. 4, pp. 140–151, Apr. 2020, doi: 10.23919/JCC.2020.04.013.

[28] M. Ren, J. Zhang, L. Khoukhi, H. Labiod, and V. Vèque, "A unified framework of clustering approach in vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1401–1414, May 2018, doi: 10.1109/TITS.2017.2727226.

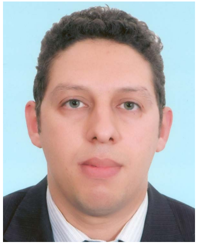
[29] M. Seo, S. Lee, and S. Lee, "Clustering-based data dissemination protocol using the path similarity for autonomous vehicles," *Symmetry*, vol. 11, no. 2, p. 260, Feb. 2019.

[30] M. S. Talib, A. Hassan, T. Alamery, Z. A. Abas, A. A.-J. Mohammed, A. J. Ibrahim, and N. I. Abdullah, "A center-based stable evolving clustering algorithm with grid partitioning and extended mobility features for VANETs," *IEEE Access*, vol. 8, pp. 169908–169921, 2020, doi: 10.1109/ACCESS.2020.3020510.
[31] H. W. Tseng, R. Y. Wu, and C. W. Lo, "A stable clustering algorithm using the traffic regularity of buses in urban VANET scenarios," *Wireless Netw.*, vol. 1, pp. 1–15, May 2019.
[32] W. Deng, R. Chen, B. He, Y. Q. Liu, L. F. Yin, and J. H. Guo, "A novel two-stage hybrid swarm intelligence optimization algorithm and application," *Soft Comput.*, vol. 16, no. 10, pp. 1707–1722, 2012.
[33] M. A. Hossain, H. R. Pota, S. Squartini, F. Zaman, and K. M. Muttaqi, "Energy management of community microgrids considering degradation cost of battery," *J. Energy Storage*, vol. 22, pp. 257–269, Apr. 2019.
[34] L. Rokach, "A survey of clustering algorithms," in *Data Mining and Knowledge Discovery Handbook*. Boston, MA, USA: Springer, 2009, pp. 269–298.
[35] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, Jun. 2014.
[36] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proc. IEEE Int. Conf. Evol. Comput., IEEE World Congr. Comput. Intell.*, May 1998, pp. 69–73, doi: 10.1109/ICEC.1998.699146.
[37] A. Ratnaweera, S. K. Halgamuge, and H. C. Watson, "Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 240–255, Jun. 2004, doi: 10.1109/TEVC.2004.826071.
[38] E. Moridi and H. Barati, "RMRPTS: A reliable multi-level routing protocol with tabu search in VANET," *Telecommun. Syst.*, vol. 65, no. 1, pp. 127–137, 2016.

KHALID KANDALI was born in 1985. He received the Ph.D. degree in computer science from Moulay Ismail University, Meknes, in 2021. His main research interests include vehicular ad hoc networks, wireless communications, and machine learning.

LAMYAE BENNIS received the Ph.D. degree in computer sciences from Moulay Ismail University, Meknes, Morocco, in 2020. She is currently a member of the Mathematics and Computer Sciences Laboratory, Faculty of Sciences of Meknes. Her research interests include educational technologies, intelligent systems, learning analytics, and serious games.

OMAR EL BANNAY received the Ph.D. degree in computer sciences from the Faculty of Science, Fes, in 2010. Since 2011, he has been a Professor at the National School of Applied Sciences, Khouribga. His research interests include computer sciences, distributed systems, VANET, and intelligent vision.



HAMID BENNIS received the Ph.D. degree in computer science and telecommunications from the University of Mohammed V Agdal, Rabat, Morocco, in 2011. He is currently a Professor of computer science and telecommunications with Moulay Ismail University, Meknes, Morocco. He is also the Head of Research Team "Communication Systems, Artificial Intelligence and Mathematics." His research interests include vehicular ad hoc networks, wireless communications, wireless power transmission, machine learning, and applications of the Internet of Things.