

Research Article

Deep Learning-Based Dynamic Stable Cluster Head Selection in VANET

Muhammad Asim Saleem ¹, Zhou Shijie,¹ Muhammad Umer Sarwar,²
Tanveer Ahmad ³, Amarah Maqbool,⁴ Casper Shikali Shivachi ⁵ and Maham Tariq⁴

¹School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, China

²Department of Computer Science, Government College University, Faisalabad, Pakistan

³Department of Electrical and Electronic Engineering, Auckland University of Technology, Auckland 1010, New Zealand

⁴Department of Computer Science, Government College Women University, Faisalabad, Pakistan

⁵South Eastern Kenya University, Kitui, Kenya

Correspondence should be addressed to Casper Shikali Shivachi; cshikali@seku.ac.ke

Received 26 March 2021; Accepted 1 July 2021; Published 12 July 2021

Academic Editor: Massimiliano Zanin

Copyright © 2021 Muhammad Asim Saleem et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

VANET is the spontaneous evolving creation of a wireless network, and clustering in these networks is a challenging task due to rapidly changing topology and frequent disconnection in networks. The cluster head (CH) stability plays a prominent role in robustness and scalability in the network. The stable CH ensures minimum intra- and intercluster communication, thereby reducing the overhead. These challenges lead the authors to search for a CH selection method based on a weighted amalgamation of four metrics: benefit factor, community neighborhood, eccentricity, and trust. The stability of CH depends on the vehicle's speed, distance, velocity, and change in acceleration. These all are included in the benefit factor. Also, the accurate location of the vehicle in changing the model is very vital. Thus, the predicted location with the Kalman filter's help is used to evaluate CH stability. The results have shown better performance than the existing state of the art for the benefit factor. The change in dynamics and frequent disconnection in communication links due to the vehicle's high speed are inevitable. To comprehend this problem, a graphing approach is used to evaluate the eccentricity and the community neighborhood. The link reliability is calculated using the eigengap heuristic. The last metric is trust; this is one of the concepts that has not been included in the weighted approach to date as per the literature. An adaptive spectrum sensing is designed for evaluating the trust values specifically for the primary users. A deep recurrent learning network, commonly known as long short-term memory (LSTM), is trained for the probability of detection with various signals and noise conditions. The false rate has drastically reduced with the usage of LSTM. The proposed scheme is tested on the real map of Chengdu, southwestern China's Sichuan province, with different vehicular mobilities. The comparative study with the individual and weighted metric has shown significant improvement in the cluster head stability during high vehicular density. Also, there is a considerable increase in network performance in energy, packet delay, packet delay ratio, and throughput.

1. Introduction

A vehicular ad hoc network, popularly known as VANET, is a particular type of Mobile Ad Hoc Network (MANET), where mobile nodes are considered vehicles moving on the road [1]. The advent of developments in VANET has paved a way for the growth of Intelligent Transportation System (ITS) applications. These are broadly classified into safety-oriented applications that intended to increase safety and

reduce fatal road accidents. The other is nonsafety that aims to provide additional services to passengers like traffic management, information sharing, and so on [2–4]. The vehicles communicate through the On-Board Units (OBUs), known as Vehicle to Vehicle (V2V). This becomes essential for most ITS application owing to low cost and availability [5]. Also, the vehicles can communicate with the auxiliary facilities (or installed infrastructure) like Road-Side Units (RSUs), using Vehicle-to-Infrastructure

(V2I) communications. The complete model design for VANET communication is shown in Figure 1.

Typically, in the V2V communication protocol, two approaches are used: flooding and relaying for data dissemination through the network. In flooding, each node broadcasts the received data packet to its neighbor received from the source. This process is repeated with the motive to reach the data packet to the source. In a flat, dense network, this approach leads to a storm problem [6]. In the relaying approach, the message broadcast from the source is forwarded to all the neighboring vehicles, and then a few are selected to forward it. The probability of success for data delivery is increased in the relaying approach but with high overhead and delay. To tackle such a problem, clustering in VANET is one of the prominent solutions. Some notations are listed in Table 1, which will be used for further gradation.

Through clustering in VANETs, there is information gathering, aggregation, and dissemination. Clustering is used to partition the network into smaller groups of moving vehicles. Typically, the dynamic clustering is done, which groups the vehicles on the fly as there is no physical connection among them, and they all are moving. There are several benefits like efficient bandwidth, proper distribution of resources, and scalability that the clustering approach offers [7]. There are many methods employed for clustering in VANETs [8, 9]. The urban roads are complicated as vehicle positions change frequently; they have unevenly distributed; thus, their routing and the forwarding capabilities change with their position. These issues result in an unstable network and a need for a cluster model that offers high stability in the dynamic VANET scenario. The methods especially designed for the stability of the cluster are compared in Table 2.

Through this analysis, we can find that most of the schemes conceived for stability are formulated on the effect of a stability parameter. The weighted scheme has also considered the parameters like speed, distance, acceleration, or link time. These schemes involve no information about the type of vehicle present in the traffic scenario. The connectivity with the neighbor is also not explored. These have been designed for highway where the vehicle's speed is too high, so the static cluster formation concerning speed and position will change dynamic, which will affect the stability of the cluster head. With all these literature analyses, the authors have designed a weighted cluster head selection based on multimetric. The benefit factor, community neighborhood, eccentricity, and trust value are introduced. Also, the concept of trust has not to be utilized to date. The following are the contributions made by the authors in this paper:

- (i) The benefit factor designed by the authors in [19] is based on T_{leave} , the time required by the vehicle to complete the remaining segment of the lane. This is calculated based on the speed of the vehicle. The speed of any vehicle is a variable quantity; in such a scenario, obtaining the accurate speed is essential for cluster head stability. This issue has been addressed in this paper; the authors have predicted each vehicle's

current location, and this position has been used to calculate the benefit factor. The authors showcase the validation through the comparative results.

- (ii) The authors also intended to improve the VANET performance using an auxiliary facility, Road-Side Units (RSUs), improving the network's performance. However, the trade-off between the number of RSUs for better coverage and the high installation and maintenance costs is resolved in this paper by considering the angle suspended by the lanes. The effect of crossroads is also included to avoid ambiguity in the direction of the lane.
- (iii) The authors have presented spectrum sensing as a classification problem and proposed the sensing method based on the recurrent neural network by employing the LSTM. The signal's power spectrum is given as input to the LSTM and uses various signal noise data types to train the network. The decision is made according to the confidence of the noise class. As the designed method is based on deep learning, the network automatically learns the parameters and can adapt well to different noise levels. This concept has been further extended to segregate the vehicles into two categories: the primary (PU) and secondary users. The primary includes vehicles like ambulance, a civic service that needs immediate assistance. The PU are given access to the network bandwidth and improve the stability of the VANET.
- (iv) VANET has a dynamic topology where the vehicles travel at high speed and have frequent arrival with an irregular interval between them. To have a stable cluster, they have to be evolving in nature. The authors have designed the cluster-based evolving graph model that calculates the metrics that could select a stable cluster head to address this use.

The authors designed a new multimetric weighted CH selection scheme that considers different selection metrics that increase the cluster stability and efficiency. The designed scheme is well tested on the real data map, taken for the region Chengdu, southwestern China's Sichuan province. The clustering efficiency is tested for various vehicle densities, and also the network performance is evaluated using the parameters. The rest of the paper is organized as follows: In Section 2, the network model and the position prediction using the Kalman filter and the RSU placement are presented. In Section 3, the multimetric weighted CH selection scheme is discussed with all its metrics. Section 5 presents the simulation work and compares the proposed scheme with the existing state of the art. In the last section, the authors conclude the work with the prospects of the work.

2. Network Model with Prediction and RSU Deployment

The proposed scheme is mainly concentrated on the highway environment. Here, the model is specially designed for vehicles that are well equipped with a positioning system

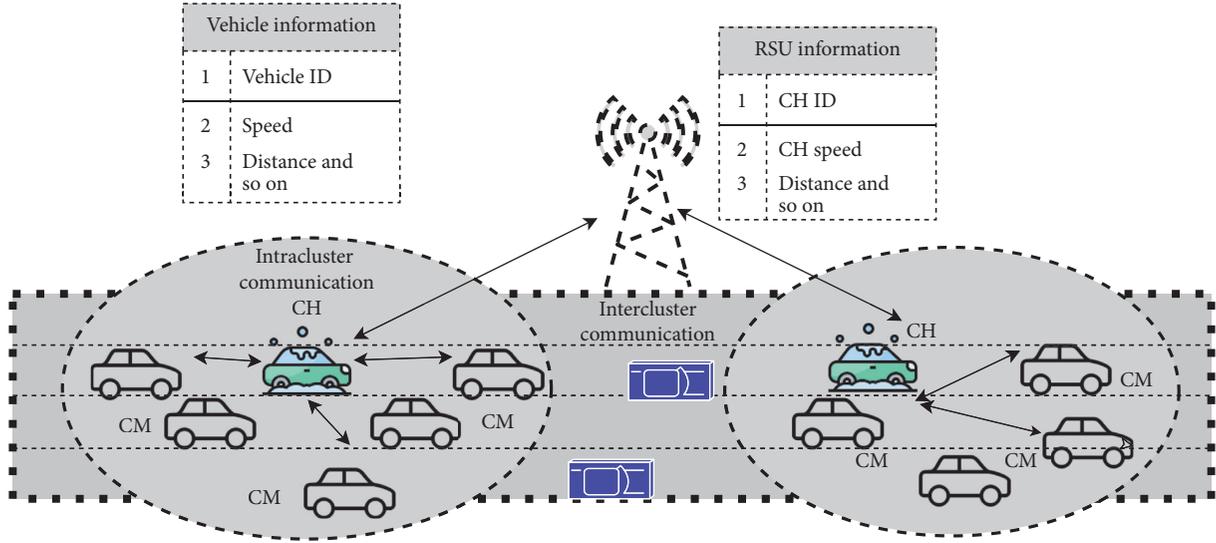


FIGURE 1: The complete system model.

TABLE 1: List of symbols with the nomenclature.

Symbol	Nomenclature
A	Affinity matrix
BF	Befit factor
C	Cluster
CN	Community neighbors
D	Distance
Dig	Diagonal matrix
d_{ij}	Diagonal elements
ΔD_{thr}	Threshold distance
b_{μ}, b_f, b_o	The bias of each gate in LSTM
NH_D	Neighborhood
Ψ_{vehi}	Relative velocity
ID	Vehicular ID
L	Length of the lane
Ecc	Eccentricity
$G(M, E, rl)$	Graph tuple
E	Energy
M	Total number of vehicles
N	Total number of clusters
NCC	Neighbor connection centrality
h	Channel
\mathcal{H}_0	Hypothesis
P_d	Probability of detection
t_p	Delay in p th path
P_f	Probability of false alarm
q	Wavelength
rl	Link reliability
RSU_{tran}^{L-ID}	RSU
$Q(\cdot)$	Complementary distribution function
S_{thr}	Speed threshold
T_{leave}	Time to leave
TRF	Transmission range factor
T	Total simulation time
TR	Transmission range
χ^2	Chi-square probability distribution function
t	An instant of time
\tanh	Activation function
V_{vehi}	Velocity of the vehicle
ΔV_{thr}	Velocity threshold

TABLE 1: Continued.

Symbol	Nomenclature
V_{avg}	Average velocity
w	A random number [0, 1]
σ_x^2	Standard deviation
γ	SNR
\hat{x}	Predicted coordinates
$[x, y]$	Coordinates of vehicle
\emptyset	Angle of the vehicle
no_lane	Total number of lanes
δ	Penalty
λ	Eigenvalues
$\sigma_f, \sigma_u, \sigma_0$	Gates in LSTM
A_p	Channel gain
P_r	Probability of receiving
G	Antenna gain
\odot	Hamdard function
$T(Y)$	Test statistics

TABLE 2: Comparison of various state-of-the-art schemes designed for CH stability.

Year	Parameters for CH selection							Clustering scheme		Network and utility scenarios
	Speed	ID	Signal strength	Relative speed	Neighbor count	Link lifetime	Position	Individual	Weighted	
2011 [10]	Yes	Yes	Yes	—	—	—	—	Yes	—	Multihop with random walk mobility
2012 [11]	—	—	—	Yes	—	—	—	PDR	—	Multihop with manhattan mobility scenario
2013 [12]	—	—	—	Yes	—	—	—	Least mobility	—	Multihop with two lanes
2015 [13]	—	—	—	Yes	Yes	—	—	Neighbor count	—	Multihop with a dynamic scenario
2018 [14]	Yes	—	—	—	Yes	Yes	Yes	Priority	—	Multihop with two lanes
2018 [15]	Yes	—	—	—	—	Yes	Yes	The conditional probability of link reliability	—	Singel-hop with an urban scenario
2019 [16]	Yes	—	Yes	Yes	—	Yes	—	—	Yes	Singel-hop with an urban scenario and highway
2019 [17]	Yes	—	—	—	—	—	Yes	—	Yes	Singel-hop with highway
2019 [18]	Yes	—	—	Yes	—	—	Yes	—	Yes	Singel-hop with highway

(e.g., a GPS), which acquires information about the current location and an IEEE 802.11p-compliant radio transceiver that enables the communication between the other vehicles and the RSU (Road-Side Unit). The cars typically share different geographical proximity information, which can be harnessed for effective traffic management and organization. The designed weighted, dynamic, adaptive, and fuzzy clustering algorithm is manifested on the same ideology [20–22]. Here, the vehicles are grouped in small clusters based on their proximity concerning the RSU deployed near the lanes. A cluster head (CH) is selected based on the amalgamation of four components benefit, trust, community neighborhood, and eccentricity. The rest are cluster members (CM). There is an exchange of information among the members using Vehicle-to-Vehicle (V2V) communication.

The complete knowledge is like the parameter speed, acceleration, distance travelled, location, and so on, which are maintained in each CM. The CH plays its role in broadcasting important information to the in-range RSU and the CM.

In this section, a detailed discussion of the designed model is done. Each parameter for the CH selection is examined and modeled without increasing the network's complexity and overhead delay.

2.1. Network Model. The network model considered is shown in Figure 2(a) that shows the nodes in the VANET in an unclustered fashion. In Figure 2(b), the nodes are now under each cluster head in continuous communication

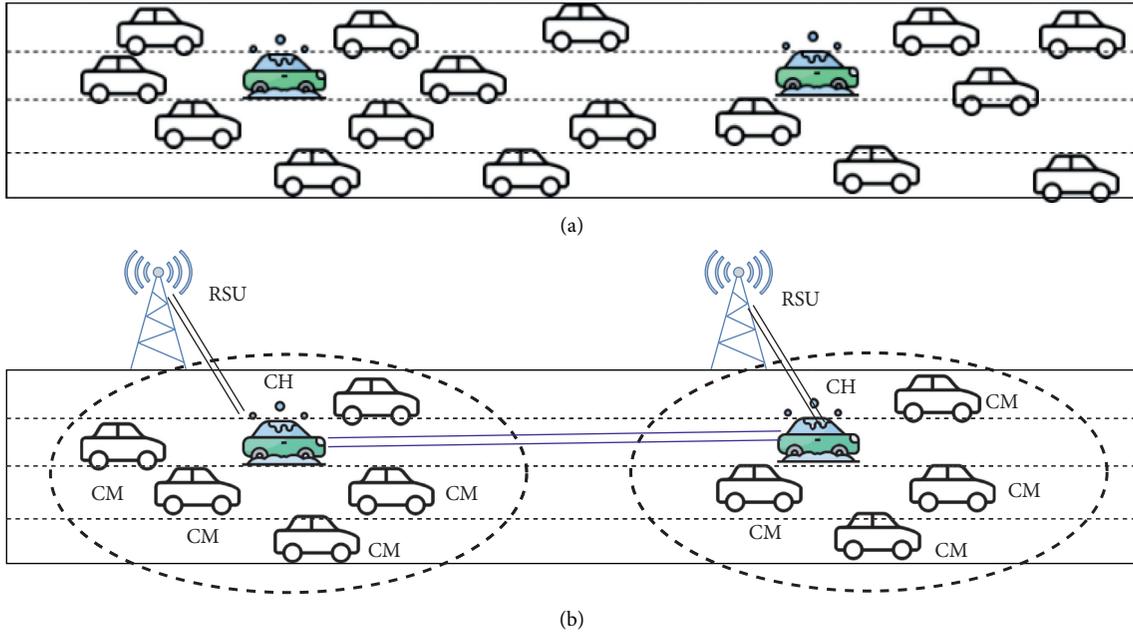


FIGURE 2: The proposed analogy for VANET clustering: (a) flat structure and (b) cluster concept with CH and RSU.

with the RSU. The network model's consideration is as follows: (1) The real-time road map is taken for the study. (2) The vehicle mobility is considered for a fixed time interval for study. (3) Each vehicle is considered as a node that is variable with time, and each car possesses a unique ID. (4) RSUs are effectively placed in each lane with proper analysis of the road with an identical communication range. (5) Each vehicle and the RSU are equipped with buffer memory and a GPS device for real-time data collection and synchronization. (6) The communication among the CM and CH-CM is enabled through V2V. (7) Each vehicle's position is estimated through the Kalman filter that intimates the RSU about the range of vehicles and ensures noninterruptible communication with smooth transitions among the RSU. (8) The network over time evolves into several dynamic clusters; the maximum hops between the cluster head and its CM are maxing hop. (9) Packet data of size 64 bits is taken.

2.2. RSU Deployment. To overcome the shortcomings of overhead delay and fewer storage facilities, the vehicles need an auxiliary facility like Road-Side Units (RSUs) to improve network performance. There is always a trade-off between the areas' total coverage with a sufficient number of RSUs with minimum installation and maintenance cost. This leads to the search for optimal locations where the RSU can maintain an efficient communication channel at a low price. The location affects communication as any void area will cause a drop of packets transferred among the vehicle and RSU. In this deployment, the lane coordinates are obtained from the SUMO simulation. Each RSU is assumed with a fixed circular transmission range RSU_{tran}^{L-ID} . There are placed on either of the roadsides. This ensures no interference in the coverage range of RSU and each lane as sufficient coverage

without any void area [23]. The deployment coordinates are calculated using the particular direction of the lane through the geometric methodology. At the beginning of any lane, an RSU is placed to get the next location lane curvature $\cos(\emptyset)$ which is used. Let the location of any point on the lane be expressed as $[x, y]$. The next location is calculated as

$$\begin{aligned} x_{next} &= x_{old} + RSU_{tran}^{L-ID} * \cos(\emptyset), \\ y_{next} &= y_{old} + RSU_{tran}^{L-ID} * \sin(\emptyset). \end{aligned} \quad (1)$$

That takes into consideration the RSU transmission range. With the introduction of the angle factor, the cross-lane will not suffice to formulate such conditions. There are few checkpoints designed that could ensure that cross-lane will be utilized to place the RSU continuously. The deployment of the RSU is shown in Figure 3.

2.3. Location Prediction. After optimally placing the RSU on the desired area map, the next task is to reduce the drop rate in the communication between the RSU and CH. This will ensure better overall network performance and increase the reliability of the VANET. To make this possible, the RSU must be aware of the vehicle's next stamp position beforehand. In this work, the authors have utilized the work proposed by Kalman [24, 25] that predicts the vehicle's location with geographical routing. The prediction of the vehicle's location depends on the direction and the velocity. Also, it has been observed that the vehicle's angle plays a crucial role in predicting the position as it varies over time. The prediction is made using the Kalman filter. Each vehicle runs a Kalman filter prediction, which predicts the vehicles' position, velocity, and direction. The position vector at the moment t is $x(t) = [x(t), y(t), V_{vehi}(t), \phi(t)]^T$ and the predicted values for each parameter at the next instant are

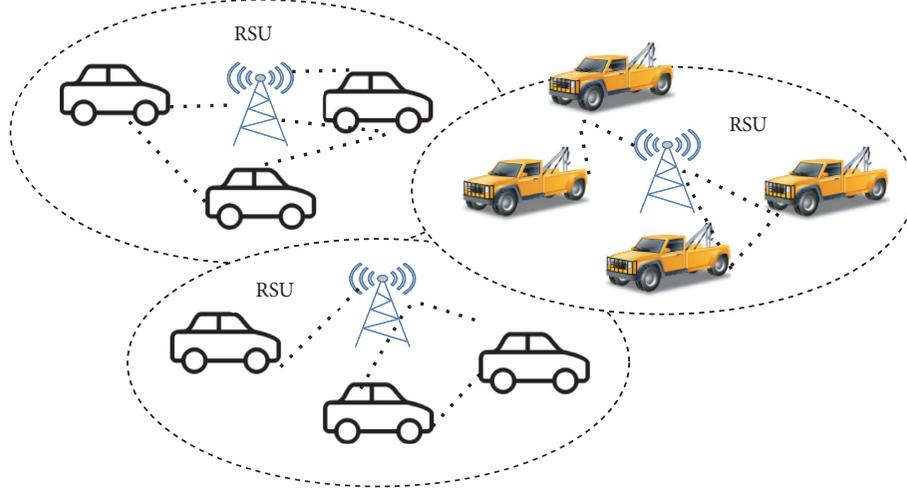


FIGURE 3: The deployment of RSU on the lane for the circular transmission range.

$$\begin{aligned} \tilde{x}(t) &= |\hat{x}(t), \hat{y}(t), \widehat{V}_{\text{vehi}}(t), \hat{\phi}(t)|^T, \\ \hat{x}(t+1) &= \left| \hat{x}(t) + \widehat{V}_{\text{vehi}}(t) \times \cos(\hat{\phi}(t)) \times \Delta T \right|, \\ \hat{y}(t+1) &= \left| \hat{y}(t) + \widehat{V}_{\text{vehi}}(t) \times \sin(\hat{\phi}(t)) \times \Delta T \right|. \end{aligned} \quad (2)$$

The individual RSU scans for vehicles in its range; then using the following judgment equation, new predicted positions are derived. The difference between the past and the predicted position is used to get the new update value with the information on the angle suspended by the lane.

3. Dynamic Weighted Cluster Head Stability Algorithm

The designed clustering model has been broadly classified into two parts: (1) cluster formation and (2) CH selection. In the cluster formation, the vehicles are segregated into small groups and then based on a weighted approach, a weighted formulation of four metric benefit factor (BF), community neighborhood (CN), eccentricity (Ecc), and trust (T) is designed to select the cluster head, and also these parameters ensure the stability of the cluster head over the time. All of these metrics are normalized to avoid the overriding effect.

3.1. Cluster Formation. The RSU guides the cluster formation. Algorithm 1 encapsulates the cluster formation where each vehicle enters the lane for the first timestamp broadcast and communicates with all the RSU in its communication range. To form a stable cluster, two conditions are observed for that timestamp, the distance D and relative change in the vehicle's speed V_{vehi} in that timestamp. A distance between the RSU and that car is calculated. If this distance is less than a threshold ΔD_{thr} , the vehicle is attached to that RSU temporarily. This process is carried out for all the vehicles present in the vicinity and moving in the same direction. In the real-time scenario,

each vehicle's speed level may be different, and this variation can seriously affect the cluster formation. For stable cluster formation, the change in relative speed is also observed if it is less than a threshold ΔV_{thr} and then that vehicle gets permission to get permanently attached to that RSU. Now, the RSU stores the ID of that vehicle and becomes a cluster member (CM).

3.2. Cluster Head Selection. The next step is selecting the cluster head, and it is a node in VANET that coordinates or heads the cluster. It takes the responsibility of broadcasting, discovery, and maintenance of the routing path. It remains in the contact of the RSU to sustainably maintain the intra- and intercommunication channels. The main motive of designing any clustering algorithm is the stability of the cluster head in VANET. The vehicles' high dynamic mobility can lead to frequent reclustering and eventually decreases the cluster stability. The methodology designed in this paper for the cluster head selection is a weighted combination of four factors, as mentioned above. All these parameters direct towards the search for a stable CH. All these and their amalgamation for selecting a suitable CH are discussed in the following.

3.2.1. Benefit Factor (BF). This parameter is defined to maximize the stability of the cluster structure. To ensure this, the elected CH is expected to stay connected with all the cluster members for a longer duration. Thus, the BF is derived from the three metrics as designed in [19]

$$BF = w_1 \times T_{\text{leave}} + w_2 \times \psi_{\text{vehi}} + w_3 \times NH_D, \quad (3)$$

where w_1, w_2 , and w_3 are the corresponding weights that vary in the range $[0, 1]$ satisfying the condition that $(w_1 + w_2 + w_3 = 1)$ and can be reformulated by the local authority based on the road conditions and the cluster member behavior. The first metric is T_{leave} , which is time to leave; this is the time required for a vehicle to complete the lane's

```

Input: Velocity  $V_{\text{vehi}}$  and location  $[x, y]$ ; the number of the lanes in a map (no_lane); time_span
Output: Cluster (C),  $N$  no of clusters
For  $t = 1:\text{time\_span}$ 
For  $i = 1:\text{no\_lane}$ 
  If any vehicle detected in  $\text{RSU}_{\text{tran}}^{\text{L-ID}}$ 
    D: RSU location  $[x_{\text{RSU}}, y_{\text{RSU}}]$  and vehicles location  $[x_{\text{vehi}}, y_{\text{vehi}}]$ 
     $V_{\text{vehi}}$ : Speed of the vehicle
    If  $D < \Delta D_{\text{thr}} \parallel V_{\text{vehi}} < \Delta V_{\text{thr}}$ 
       $\text{RSU} \leftarrow \text{ID}$ 
    Endif
     $C \leftarrow \text{RSU}, \forall \text{ID}$ 
  Endif
Endif
End
 $N \leftarrow \forall C; C \neq \text{Null}$ 

```

ALGORITHM 1: Cluster formation.

remaining segment. This factor ensures selecting a CH with enough time left to complete the lane that makes it head for a longer time. It is calculated using L , the length of the lane, and D , the distance covered by a vehicle on the road segment, and t , the time consumed by the vehicle to drive that D section of the lane.

$$T_{\text{leave}} = \frac{L - D}{t}. \quad (4)$$

The second metric is ψ_{vehi} defined as relative average speed; this parameter determines how close a vehicle's velocity is to its neighbors. A reward function is conceived that takes into account the velocity of vehicles over the long term. Speed of each vehicle (V_{vehi}) is evaluated. Accordingly, their speed is either rewarded or penalized with an absolute value (δ); correspondingly, the relative average speed is incremented or decremented, as shown in

$$\begin{aligned} \psi_{\text{vehi}}(t+1) &= \psi_{\text{vehi}}(t) + \delta ; |V_{\text{vehi}} - V_{\text{avg}}| \leq S_{\text{thr}}, \\ \psi_{\text{vehi}}(t+1) &= \psi_{\text{vehi}}(t) - \delta ; |V_{\text{vehi}} - V_{\text{avg}}| > S_{\text{thr}}, \end{aligned} \quad (5)$$

where S_{thr} is the parameter that ensures that the vehicle moving with the velocity V_{vehi} is almost travelling at the same speed as that of the neighbors. The initial value of ψ_{vehi} is calculated using the TraCI parameters, and δ is taken as 0.01. The last metric is neighborhood degree (NH_D), and it is defined as the number of the neighbors whose speed difference with the vehicle is less than a threshold S_{thr} .

3.2.2. Eccentricity (Ecc). The next parameter is eccentricity; in real time, communication links break more frequently due to the vehicles' high speed. To maintain the link, there is a requirement for an evolving cluster model. Usually, reclustering will become inevitable once the CH resigns or loses its suitability to continue as a CH. To ensure stability, the concept of eccentricity is introduced. Here, an evolving graph-based model is designed by using spectral clustering. A vehicular graph topology is intended to be

$G(M, E, rl)$, where M is the number of vehicles present in the timestamp t , E is the ordered pair of the links among the vehicles, and rl is the link reliability. The affinity matrix is constructed to represent the graph topology for dimensions $M \times M$.

$$A = \begin{cases} r_{ij}, & \text{if } (D_i, D_j) \in E, \\ 0, & \text{if } i = j, \\ \infty, & \text{otherwise,} \end{cases} \quad (6)$$

$$A = \begin{bmatrix} r_{11} & \cdots & r_{1M} \\ \vdots & r_{22} & \vdots \\ r_{M1} & \cdots & r_{MM} \end{bmatrix}.$$

The tool employed for the spectral clustering is the Laplacian graph. The Laplacian graph is calculated for the affinity matrix:

$$G_L = \text{Dig} - A, \quad (7)$$

where Dig is the diagonal matrix with elements. Also,

$$d_{ij} = \sum A_{ij}. \quad (8)$$

In spectral clustering, eigenvectors of a similarity/affinity matrix are derived from the original dataset. The eigendecomposition of the graph will be serving as the model for dimensionality reduction of mobile vehicles. The optimal number of clusters is calculated by using the eigenvalue of the Laplacian graph. Based on the eigenmap heuristic [26], the eigenvalues $\lambda_i; i = \{1, 2, \dots, M\}$ get sorted in the ascending order out of which k is picked that serves as the clusters in that timestamp.

$$k = \arg \max (\lambda_{i+1} - \lambda_i). \quad (9)$$

After the number of the clusters is obtained, then k eigenvectors are extracted in a matrix with dimensionality $M \times k$. In the last, K-means is applied to get the optimal number of clusters. Ecc is the mean/average eigenscore of each group that is calculated as in [27]

$$\text{Ecc}_i = \frac{1}{|N_i|} \sum_{\lambda_i \in N_i} \lambda_i. \quad (10)$$

The maximum value of Ecc ensures a stable cluster head selection designed based on the evolving graph.

3.2.3. Community Neighborhood (CN). The evolving Laplacian graph also provides information about neighbors. The importance of neighbor ensures the CH's stability as the cluster member will not change for a given timestamp that will establish a reliable link among them. There is designing of the CN using the transmission factor (TRF); it represents the reliability of the connection between two vehicles that satisfy the following condition:

$$\text{TRF}(r_{ij}) = \begin{cases} 0, & \text{if } r_{ij} < \text{TR}, \\ \frac{\text{TR}^2 - r_{ij}}{\text{TR}^2}, & \text{if } 0 < r_{ij} \leq 1, \end{cases} \quad (11)$$

where TRF is the maximum transmission range of the vehicle and r_{ij} is the distance between those two vehicles at the timestamp t . There is a negative correlation between the distance and transmission range. This enables that if two vehicles are closer, then a more reliable connection is bound. The neighbor nodes are defined as those vehicles that satisfy the condition $\text{TRF}(r_{ij}) > 0$. Then, the next step is to count neighbor connection centrality that is defined as follows:

$$\text{NCC} = \sum_{r_{ij} \in NE} \text{TRF}(r_{ij}). \quad (12)$$

The last step is to get CN, the weighted average of NCC over the timestamp t [28].

$$\text{CN} = w_i \times \text{NCC}, \quad (13)$$

where w_i is the weighted associated at each timestamp.

3.2.4. Trust (T). The type of vehicle also plays an essential role in cluster stability; this is an efficient technique for dealing with malicious and compromised nodes [29]. To address this problem, the authors have included the notion of spectrum sensing. Here, concepts are addressed that could remarkably enhance the cluster head's stability in the network. The first is the spectrum sensing technique that helps to utilize the spectrum efficiently. Here, the spectrum sensing is taken as a classification problem using the long short-term memory (LSTM). The network is trained using various types of signal and noise data. Data can be handled by using the latest technology called big data handling [30, 31], though this decision is made on the confidence of the noise class. Since the method includes a recurrent neural network, it automatically learns the energy features and adapts to any untrained noise or signal in a real and dynamic environment. This enables the detection of the primary users' (PU), like medical vans, police vans, or any other type of civil service, to use the spectrum as a priority if there is an emergency. The rest is considered the secondary users (SU).

In the following section, the formulation to address the above-stated issues is discussed.

LSTM Model Design. The internal structure of LSTM is shown in Figure 4. Here, $Y^{(t)}$ is the input to the cell structure, and output is denoted by $a^{(t)}$. The previous cell input is taken as $a^{(t-1)}$, current and previous cell states are represented by $c^{(t)}$ and $c^{(t-1)}$. σ_f , σ_u , and σ_0 show the value of three gates forget gate, update gate, and output gate, respectively. The Hadamard function is denoted by \odot , and \tanh shows the activation function with elementwise addition \oplus .

Any LSTM cell constitutes three main gates: update gate, forget gate, and output gate. The function of each gate is stated as follows:

- (i) Update gate (σ_u): decide when to update the current cell state
- (ii) Forget gate (σ_f): discard function of the current cell
- (iii) Output gate (σ_0): provided to the output through the output gate

The updated vector for each state is obtained through

$$\tilde{c}^{(t)} = \tanh(w_c [a^{(t-1)}, Y^{(t)}] + b_c), \quad (14)$$

where b_c is the bias term and \tanh is the activation function. All the three gates get updated using the individual bias and the sigmoid function.

$$\begin{aligned} \Gamma_\mu &= \sigma(w_\mu [a^{(t-1)}, Y^{(t)}] + b_\mu), \\ \Gamma_f &= \sigma(w_f [a^{(t-1)}, Y^{(t)}] + b_f), \\ \Gamma_0 &= \sigma(w_0 [a^{(t-1)}, Y^{(t)}] + b_0), \end{aligned} \quad (15)$$

$$\sigma_m = \frac{1}{1 + e^{-m}},$$

where w_μ , w_f , and w_0 are the weight matrices. The bias terms are denoted by b_μ , b_f , and b_0 . An elementwise product is taken among the previous cell state $c^{(t-1)}$ and forget gate Γ_f and among the update gate and candidate vector updating $\tilde{c}^{(t)}$. The elementwise product among the output gate Γ_0 and hyperbolic tangent vector $c^{(t)}$ is as follows:

$$\begin{aligned} c^{(t)} &= \Gamma_\mu \odot \tilde{c}^{(t)} + \Gamma_f \odot c^{(t-1)}, \\ a^{(t)} &= \Gamma_0 \odot \tanh(c^{(t)}). \end{aligned} \quad (16)$$

The architecture designed in this study has two-bit layer with every 100 nodes and a fully connected layer followed by a SoftMax layer as the decision is no binary bases. The network is trained for 500 epochs, having a learning rate of 0.01, and the batch size of the data chunk at a time is 500.

Data Modeling. The data is captured through the ideal probability of the primary user. We consider that each vehicle takes part in the Cognitive Radio (CR) networking to transfer the information. The bandwidth allotted to the

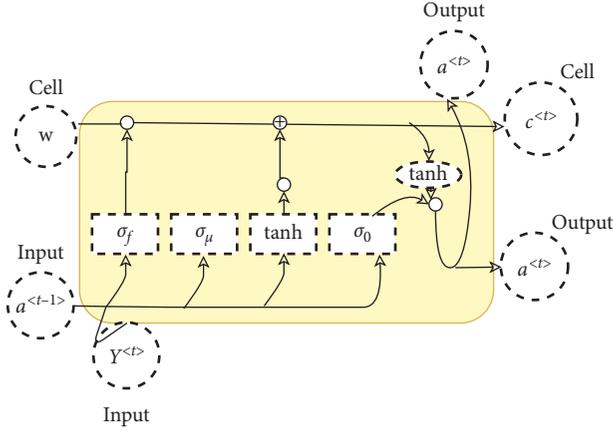


FIGURE 4: Internal structure LSTM network.

network is limited, so limited spectrum uses CR. The free spectrum of the PU can be utilized by secondary users, reducing the overhead problem. In PU's presence, the spectrum will be released for that vehicle if it receives a signal with an energy higher than the threshold signal energy. The factor trust (T) is increased if it a primary user and belongs to that neighborhood; else, the trust value is decremented. The behavior of the PU is mapped through the quantized energy vector called the sensing report. The actual status of the PU is estimated through the acknowledgment signal along with the fusion of reliable local decision of CR. The clean PU signal is acquired, and its power is measured as σ_x^2 . The Gaussian noise is added to the computed power signal of PU to achieve signal-to-noise ratio (SNR) γ and evaluate by using the relation $\sigma_w^2 = \sigma_x^2/\gamma$.

Adaptive sensing at the individual vehicle level is as follows:

$$\begin{aligned} \mathcal{H}_0: Y^{(t)} &= w^{(t)}, \\ \mathcal{H}_1: Y^{(t)} &= h^{(t)} X^{(t)} + w^{(t)}, \end{aligned} \quad (17)$$

where $Y^{(t)}$ is the signal envelop received at the t -th time instant by the sensing vehicle, the PSU signal is distorted using $w^{(t)}$ an Additive White Gaussian noise with zero mean and variance σ^2 , $X^{(t)}$ is the SNR signal transmitted from the PU transmitter, and $h^{(t)}$ is the channel gain [32]. This hypothesis can be viewed as a classification problem where there is a signal from the primary user or the presence of noise. The signal received at any instant is represented in general form with the previous sensing event of sample size M being fed to the current sensing event:

$$Y = [Y^{(1)} Y^{(2)}, \dots, Y^{(M)} Y^{(M+1)} Y^{(M+2)}, \dots, Y^{(2M)}]^T, \quad (18)$$

where M is the sample size taking into consideration these number of vehicles at timestamp t and the transpose of a vector is denoted by $[.]^T$. Here, the probability of detection and the probability of false alarm are used to detect the performance of the spectrum sensing algorithm, which are defined as

$$\begin{aligned} P_d &= P_r\{H_1 | H_1\}, \\ P_f &= P_r\{H_1 | H_0\}. \end{aligned} \quad (19)$$

The presence of the PU signal is denoted by H_1 while absence by H_0 in that case, the only noise is detected. At a given M sensing sample, the energy test statistic is represented by the energy that calculates test statistics $T(Y)$ in the time domain to compare with, which follows the Neyman-Pearson criteria.

$$T(Y) = \frac{1}{2N} \sum_{i=1}^N |Y_i^{re} + Y_i^{im}|^2. \quad (20)$$

The test statistic $T(Y)$ is the random variable with a chi-square probability distribution function (χ^2) with k degrees of freedom. It can also be represented as $Q = \sum_{i=1}^k |z_i|^2$. Here, $k = 2N$ for the complex-valued case and $k = N$ for absolute values. The threshold ε can be defined using the central limit theorem. The detection probability can be defined as [33]

$$P_d(\varepsilon, t) = \mathbb{Q}\left(\left(\frac{\varepsilon}{\sigma_u^2} - \gamma - 1\right) \sqrt{\frac{t f_s}{2\gamma + 1}}\right), \quad (21)$$

where $\mathbb{Q}(\cdot)$ is the complementary distribution function and is Gaussian in nature; that is,

$$\mathbb{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{t^2}{2}\right) dt. \quad (22)$$

If we take the inverse of equation (21), then the threshold for probability detection can be calculated as

$$\mathbb{Q}^{-1}(P_d) = \left(\frac{\varepsilon}{\sigma_u^2} - \gamma - 1\right) \sqrt{\frac{t f_s}{2\gamma + 1}}. \quad (23)$$

The false detection P_f is calculated as

$$\mathbb{Q}^{-1}(P_f) = \left(\frac{\varepsilon}{\sigma_u^2} - 1\right) \sqrt{t f_s}. \quad (24)$$

In the Rayleigh transmission channel, the converted carrier band signal is passed:

$$h(t) = \sum_{p=1}^{N_p} A_p \delta(t - t_p), \quad (25)$$

where N_p are several paths, A_p is channel gain, and t_p is the delay in p th path. The signal moves in the line of sight (LOS) considered by a free-space path loss propagation model and receives power calculated as follows:

$$P_r = \frac{G_l \times \varrho^2}{(4\pi d)^2} \times P_t. \quad (26)$$

The received and transmitted power are notated as P_r and P_t with wavelength ϱ . The LOS distance is denoted by d , and the factor dependent defines field radiation of PU and SU upon the antenna G_l . The square of the distance between PU and SU is inversely proportional to the received power. The noise variance is added to the received power and

updates the actual power value. The noisy channel is modeled as the convolution of Rayleigh channel response and noise variance where $n(t)$ is noise variance.

$$g(t) = h(t) * n(t). \quad (27)$$

This completes the design of the spectrum sensing structure.

In the training of LSTM, the behavior of CR users to the changing activity of PU in the operating environment is learned. The sensing report is generated by the CR user and makes a local decision based on its energy. Based on the outcomes, the acknowledgment signal status, the report is assigned to a sensing class. The information of the primary user is forwarded to the data center that decides spectrum sensing. The bandwidth of the transmitted signal is divided into N_s subcarriers and transmitted in chunks. These subcarriers are frequency spaced by $\Delta f = 1/T_d$ where T_d is time to transmit a signal. The signals are multiplexed using inverse discrete Fourier transform (IDFT) as

$$y_b(t) = \sum_{k=1}^{N_s} Y_k e^{j2\pi k \Delta f t}. \quad (28)$$

Every secondary node senses the energy of the transmitted signal by PU, and based on the comparison with a threshold \mathbb{Q}^{-1} from equation (23), it decides whether it is from the primary user and to vacate the channel. This energy received enables the calculation of the trust value. Figure 4 portrays the complete model for the training using the concept of spectrum sensing and then obtaining the trust value. P_d and P_f are evaluated in the proposed scheme LSTM-based spectrum sensing detection. The signals collected are fed to the LSTM network one by one, and the corresponding values of the detection and false alarm probability are computed. The primary signal vector of each SNR value is processed to the LSTM network. The number of times it correctly classified the signal H_1 divided by the total number of primary user signals fed to the network determines P_d . Similarly, the noise sequence is forwarded to the LSTM network and calculates the false alarm probability P_f . The hypothesis H_0 is divided by the total number of noise sequences used in the prediction. The complete model designed for predicting the vehicle's trust value using LSTM is shown in Figure 5.

As discussed above, adaptive spectrum sensing is implemented for forming trust based on the energy detector's threshold. The threshold is set to the target that is the desired constant probability of detection.

The limits are as obtained from equation (23). The energy of the vehicle received is given to the trained LSTM network. The network decides whether the vehicle is a primary user or not, and in turn, the value of the trust is assigned if it successfully vacates the spectrum. This adaptive spectrum sensing is done using the trained data using the LSTM. In Figure 6, the training curve of LSTM is elucidated, where the loss function, along with the accuracy, is plotted for 3,500 epochs. The loss defines the difference between the actual and predicted values of the primary users' signals in various data and noise signals.

The analysis of trained LSTM is depicted in terms of the confusion matrix, as shown in Figure 7. Here, the adaptive spectrum sensing problem is formulated as a classification problem illustrated in equation (17). This prediction is made for a total of 4,000 vehicles. This binary classifier's accuracy is 89% for the correct detection of the absence of any PU and 83.5% for any PU presence in the network. This designed LSTM model is efficient and can predict with sufficient accuracy, even in various noise conditions and mixed signals. A comparison between P_d generated from the theoretical analysis as designed in [34] and P_d for the trained LSTM is shown in Figure 8. Since $P_{d,\max} = 1$, the higher the probability, the better the trust score, and accordingly, the vehicle is considered trustworthy. This completes the design and evaluation of the trust metric. Now, the proposed weighted CH selection is discussed in the following in detail.

The complete algorithm for the CH selection is explained in Algorithm 2. All the above parameters discussed are incorporated, and a weighted metric is formulated for selecting a stable cluster head for a maximum period [35]. The number of vehicles in a time instant is firstly clustered using the RSU as discussed in Algorithm 1. Then, for all the members (CM), the above four parameters are calculated, and then a weighted CH_{score} is calculated to select the cluster head as discussed in Algorithm 2.

3.3. Computational Complexity of the Designed Complete Algorithm. In this section, the computational complexity of the dynamic weighted algorithm is discussed. The clustering algorithm is divided into two parts, the cluster formation and cluster head selection. Thus, the total time complexity of the algorithm can be stated as

$$O_{\text{tot}} = O_{\text{CF}} + O_{\text{CHS}}, \quad (29)$$

where O_{CF} is the time complexity of cluster formation and O_{CHS} is the cluster head selection.

In cluster formation, as already discussed in Section 3.1, the authors find only the distance between the vehicles and the RSU. M is the maximum number of vehicles taken in the worst-case analysis. Thus, the worst-case time complexity for this is

$$O_{\text{CF}} = O((\log M)^2). \quad (30)$$

In the cluster head selection, there are four parameters taken into consideration. The total time complexity for the cluster head selection is

$$O_{\text{CHS}} = O_{\text{BF}} + O_{\text{Ecc}} + O_{\text{CN}} + O_T. \quad (31)$$

The complexity of BF is a linear equation, where three things have been calculated, the time to leave, relative average speed, and the neighborhood degree. The time to leave and the relative average speed are only the constant values fetched and calculated. In contrast, in the neighborhood degree, the distance between the vehicles in the vicinity is calculated through their coordinates. Thus,

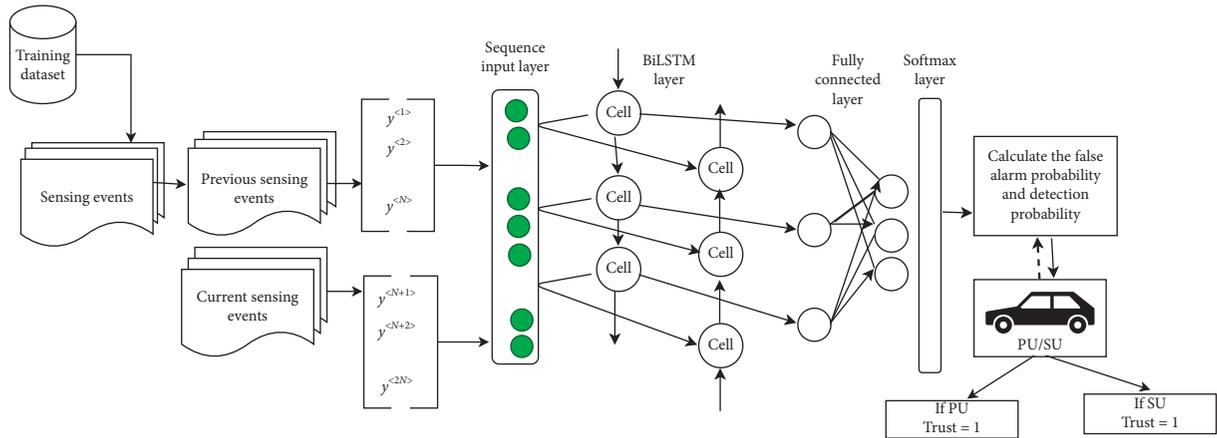


FIGURE 5: Adaptive spectrum sensing using LSTM.

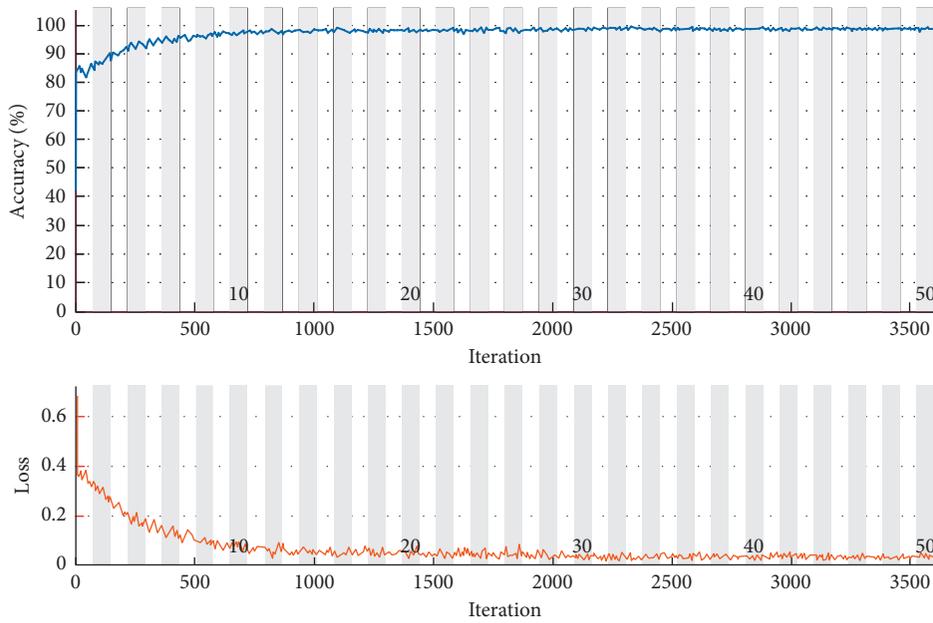


FIGURE 6: Training progress of designed LSTM for 3,500 iterations.

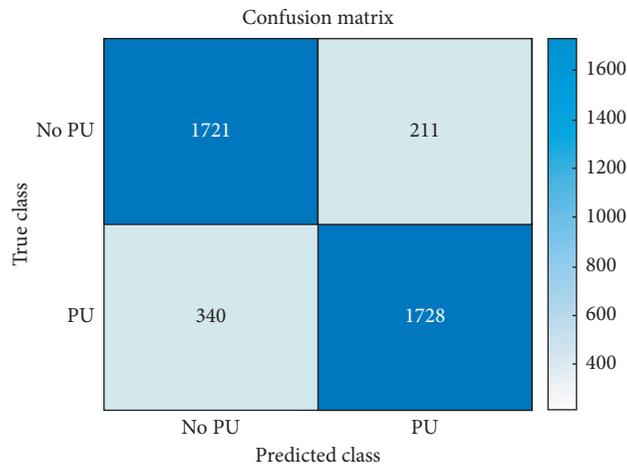


FIGURE 7: Confusion matrix for performance analysis of the classifier designed using LSTM.

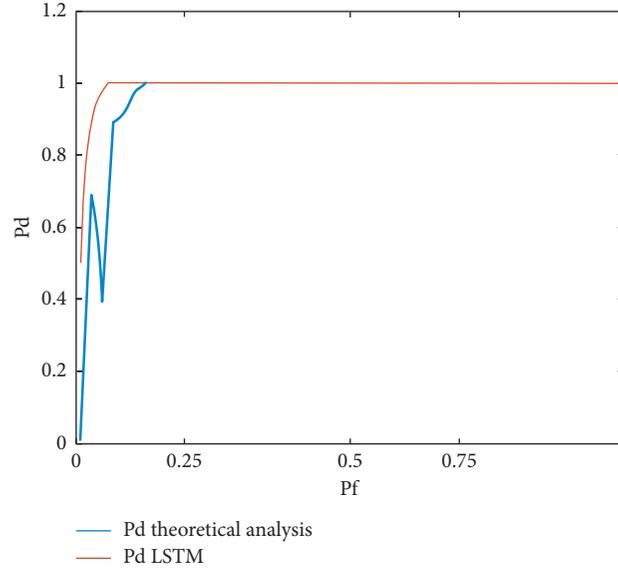


FIGURE 8: Comparison between the probability of detection and probability of false alarm using the proposed spectrum sensing LSTM method.

```

Input: Relative Speed  $V_{\text{vehi}}$  and Location  $[x, y]$ ; Number of the lane in a map (no_lane); time_span; Cluster (C),  $N$  no of clusters.
Output: CH
For  $t = 1:\text{time\_span}$ 
For  $i = 1:N$ 
  For  $j = 1:CM$ 
    Obtain the BF using the equation (3)
    Calculate the  $r_{ij}$ ; affinity matrix
    Fabricate the evolving graph  $G$  for each cluster
    Find the maximum eigenvalues  $\lambda$ ; calculate  $Ecc_j$  using equation (10)
    Determine the neighbor and find the CN using equation (12)
    From the LSTM trained network, calculate the  $T$ 
    Obtain the  $CH\_score$  for each CM
     $CH\_score = w_1 \times BF + w_2 \times Ecc + w_3 \times CN + w_4 \times T$ 
    End  $(w_1 + w_2 + w_3 + w_4 = 1)$ 
  End
   $CH = \max(CH\_score) \forall j$ 
End
End

```

ALGORITHM 2: CH selection.

$$O_{BF} = O(1) + O(1) + O((\log M)^2). \quad (32)$$

The next is the eccentricity, calculated using the spectral clustering methods that involve the affinity matrix and the eigenvalue decomposition. The complete complexity of spectral clustering is

$$O_{Ecc} = O(M)^2 + O(M)^3. \quad (33)$$

In the community neighborhood, an affinity matrix is generated for the near adjacent vehicles. Thus, the complexity for this is given as

$$O_{CN} = O((\log M)^2). \quad (34)$$

In the metric trust, the main role is played by the LSTM for spectrum sensing; the theoretical time complexity of LSTM is given as

$$O_{TH} = 4IH + 4H^2 + 3H + HK, \quad (35)$$

where I is the number of inputs, K is the number of outputs, and H is the number of hidden layers. In this study, as the model is trained once and then for a given vehicles signal, the LSTM, through its spectrum sensing, senses the vehicle either being a primary or secondary user. Thus, the time complexity bottles down to

$$O_T = O(4H^2). \quad (36)$$

Thus, the complete time complexity is reduced to removing all the terms with less complexity than cubic and quadratic terms.

$$O_{CH} = 2O((\log M)^2) + O(M)^3. \quad (37)$$

Thus, the total time complexity is

$$O_{\text{tot}} = 3O((\log M)^2) + O(M)^3. \quad (38)$$

4. Simulation Results and Discussion

This section discusses the results achieved at different stages of the cluster head model designing. It is bifurcated into the following sections: (1) Simulation Environment and Tool; (2) Network Performance Evaluation; (3) Experimental Evaluation with the existing state of the art and similar other cases.

4.1. Simulation Environment and Tool. All the simulation experiments are performed using MATLAB (R2020a), with processor Intel® Core TM i3, 1.98 GHz, Simulation of Urban Mobility [36] (SUMO 0.25.0; SUMO 0.12.0), and the TraCI [37]. SUMO is an open-source microscopic road traffic simulator licensed under the General Public License (GPL). It was developed through a collaboration between the Center for Applied Informatics Cologne (ZAIK) and the Institute of Transportation Systems (ITS) at the German Aerospace Center (DLR). While TraCI stands for the Traffic Control Interface specifically designed to get access to the traffic running on the road simulated, the embedded feature of extracting simulated objects' values. The network performance metrics like throughput, energy, packet delay, and packet delay ratio are evaluated through MATLAB. The simulation area is Chengdu, the capital of southwestern China's Sichuan province. The area taken for the simulation has the latitude = 30.6598628°N and longitude = 104.0633717°E. The area is busiest as there is a tourist place Chairman Mao statue. The total traffic environment summary is provided in Table 3.

The simulated section of the original region is shown in Figure 9; it is a vast area with a high urban and highway mobility model. The region also consists of the famous tourist spot that ensures dense vehicle movement around the peak working hours. The authors have deployed RSU, the additional facilities for network stability as per the proposed

algorithm discussed in Algorithm 3, given in Section 2.2. The map after the deployment is shown in Figure 10. It is evident from the figure that the total area has been covered efficiently. Also, enough RSU are concentrated at the cross-lane to provide sufficient coverage without any overhead delay or congestion problem, which can eventually lead to less drop in the packets and affect the network's stability.

4.2. Network Performance Evaluation. The proposed scheme's importance is tested using the four network performance metrics discussed in the following. The communication between vehicles is modeled through the Nakagmi channel, representing the obstacle medium to match real-life data transmission. Here a two-hop model is taken, where the packet transfer is initiated through any randomly selected vehicle that acts as a source to the CH and the CH sent the packet to the intended designation. The delay of this packet through the network enables the authors to calculate the desired network parameters. The size and data rate are as specified in the following table. The time taken for the packet transfer is measured using the MATLAB internal clock. Packet contains a random sequence of 1's and 0's. The communication network parameters are listed in Table 4.

The evaluation parameters for the proposed VANET stability performance assessment are as follows:

- (1) Energy (E): the amount of energy consumed at each node for communication is measured in Joules. The consumption of energy is directly proportional to the distance between the hops.

$$E = \alpha_1 + \alpha_2 \times D + \alpha_3. \quad (39)$$

- (2) Packet delivery ratio (PDR): PDR is the average ratio of successfully received packets at the destination vehicle over the total generated packets on the source vehicle. The delivery ratio decreases with increasing data rates.

$$\text{PDR} = \frac{\sum_{i=1}^N \text{packet received} \times (\text{datarate} \times \text{packet size})}{\text{packet generated at source} \times (\text{datarate} \times \text{packet size})}. \quad (40)$$

- (3) Packet delay (PD): it refers to the time taken for a packet to be sent through the transmission media from the source to the destination vehicle. The delay

in packet delivery depends on network congestion, noise, and hop travel distance.

$$\text{PD} = \frac{\sum_{i=1}^N \text{packet transmitted} \times (\text{datarate} \times \text{packet size}) - \text{packet received} \times (\text{datarate} \times \text{packet size})}{t}. \quad (41)$$

- (4) Throughput: it represents the amount of data successfully transferred from source vehicle to destination vehicle in a given period, typically

measured in Kilobits per second (kbps). The higher throughput can be achieved with less hop count and network stability.

TABLE 3: Simulation specification.

Parameter	Value
Scenario	Urban
Lanes	302
Average speed	12.71 m/sec
Simulation time	1500 sec
RSU placement	500 m
RSU transmission range	350
Vehicle transmission range	200
Vehicle length	5 m and 10 m
Vehicle speed	0–30 m/sec

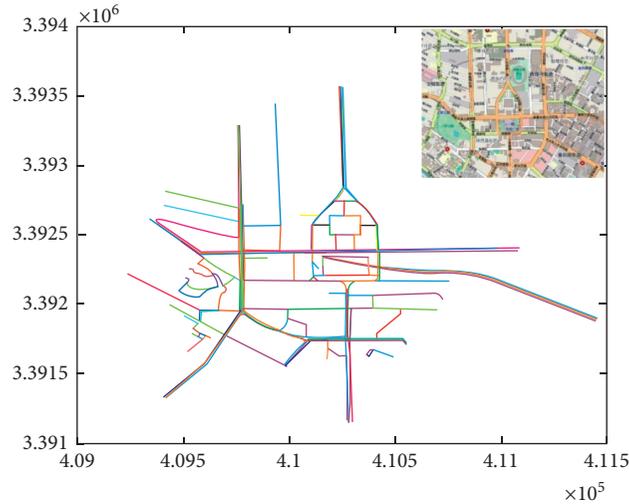


FIGURE 9: The simulated map area along with the original Chengdu, capital of southwestern China's Sichuan province.

```

Input: Lane coordinates  $[x, y]$ ; angle ( $\emptyset$ ); number of the lanes in a map (no_lane)
Output: Change in angle as per Crossroad
Place the RSU in each lane.
For  $i = 1$ : no_lane
  If ( $[x_{old}, y_{old}] < [x_{next}, y_{next}]$ )
    Lane is in [+ , +]
  Elseif ( $[x_{old}, y_{old}] > [x_{next}, y_{next}]$ )
    Lane is in [+ , -]
  Elseif ( $[x_{old}, y_{old}] > [x_{next}, y_{next}] \parallel \emptyset$ )
    Lane is in [- , +]
    Change the  $\emptyset$ 
    If  $\emptyset < 90^\circ$ 
       $\emptyset = \emptyset + 180^\circ$ 
    Elseif  $\emptyset > 90^\circ \parallel \emptyset < 270^\circ$ 
       $\emptyset = \emptyset + 90^\circ$ 
    Elseif  $\emptyset > 270^\circ \parallel \emptyset < 360^\circ$ 
       $\emptyset = \emptyset + 180^\circ$ 
    Endif
  Else
    Lane is [- , -]
  Endif
End

```

ALGORITHM 3: RSU deployment in the presence of crossroad.

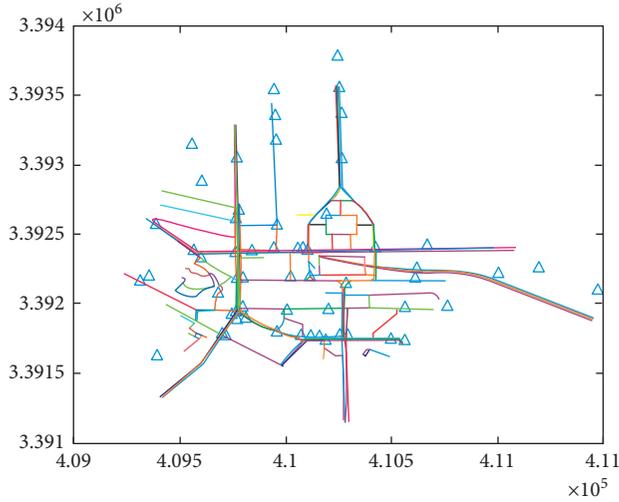


FIGURE 10: The deployment of RSU.

TABLE 4: Integrated network parameters.

Parameters	Value
Channel	Wireless
Propagation model	Propagation/Nagakami ($m = 3$)
MAC	Mac/802_11Ext
Data rate	4, 8, 10, 12, and 14 kbps
IEEE 802.11p transmission range	350 m
Packet sending interval	0.2 sec
Packet size	64 bytes

$$\text{Throughput} = \frac{\sum_{i=1}^N \text{packet received} \times (\text{datarate} \times \text{packet size})}{t} \quad (42)$$

- (5) Cluster head stability: it represents the number of times the same vehicle is chosen as cluster head in the total solution's time span.

$$\text{CHS} = \text{mode} \left(\sum_{i=1}^t \text{ID} \right). \quad (43)$$

4.3. Experimental Evaluation. The proposed scheme is tested with the following simulation environment settings:

- (i) Different vehicles' densities
- (ii) The vehicle position in a cluster is estimated by Kalman filter and calculated based on its current location

The complete evaluation for different vehicular densities is tabulated in Table 5. This study is done to evaluate the cluster stability in the different vehicular densities and the dynamic scenario. Vehicles are clustered on the ground of several RSUs. Each RSU covered area is considered as a cluster, and vehicles under that area are cluster members. The RSUs with empty clusters are omitted from the evaluation of the parameters. The dissipation of the energy

has a similar pattern for vehicular density. This ensures the stability of the cluster head. Also, the clusters with more number of cluster members or denser clusters pose a challenge as cluster head stability frequently changes in these conditions. The throughput is analyzed, two two-hop models. The throughput will be more for denser clusters as compared to the sparsely populated cluster. The other two parameters, packet delay and packet delay ratio, solemnly depend on the density and the distance of the nodes selected.

The authors have compared the evaluation of the befit factor in two ways. First, the authors have used the vehicles' current location to calculate the befit factor as suggested in the [19] and in the other way predicted and corrected location of the vehicles using Kalman filter as discussed in Section 2.3 employed. Figure 11 showcases the current location and the vehicles' corrected location for a single timestamp in the total simulation. We can analyze that, with dynamic evolving networks and frequent changing vehicle velocity, vehicles' exact location is trivial to know for the stability of the VANET. Also, the estimation of the vehicle is dependent on the direction of movement and current velocity. It has been observed that the vehicle's moving angle may vary instead of being constant at some points. These factors have significantly affected the befit factor analysis for selecting the cluster head and its stability. This assessment can be done from Figure 12, where 1,000 vehicle densities and 11 clusters are formed. The cluster head's stability is counted as the count of continuous timestamps for which any vehicle ID is constantly served as cluster head. Here, more stability is provided by the predicted method as we can observe that the frequency of a single vehicle in becoming the cluster head is about 180 times for a single period evaluation.

4.4. Baseline Comparisons. The baseline algorithms are the algorithms designed in the literature for the individual metric. In this analysis, three baseline algorithms are included as mentioned and discussed in Section 3. The analysis is carried out for the 100 vehicular density. The results are shown in Figure 13 at an instant of the total simulation time; in this, we can observe that CH selection is highest for the proposed scheme compared to the individual one as designed by the different arts state. All the simulation has been conducted on the same platform and done as described as in the literature for the comparison. As the model is urban, the vehicle's speed and density eventually lead to dynamic network changes. The cumulative effect of all the three metrics is evident compared to the individual as just change in velocity, community neighborhood, or eccentricity standalone cannot reflect the changes that a weighted approach can do. Thus, there is a need for a weighted approach that can evaluate the evolving changes in the network from time to time and make the transition of CH less.

The different network metrics are also tested on the 1,000 vehicular densities for the proposed scheme. The data packet is transmitted from a cluster member to the cluster head, and

TABLE 5: Comparison of various network parameters for different vehicular densities.

Network parameters	Clusters										
	1	2	3	4	5	6	7	8	9	10	11
<i>Vehicular density: 200</i>	12	17	1	5	19	40	65	10	5	8	7
Energy (J)	3.59	3.59	—	3.59	2.62	3.59	3.59	3.59	4.46	3.59	3.59
PDR	2	3	4	3	4	3	4	5	4	5	2
PD (s)	5.82	2.19	4.36	2.85	2.32	1.14	2.17	4.91	4.60	1.91	4.90
Throughput (kbps)	1.65	2.25	0.20	1.54	2.35	2.24	1.72	1.07	1.40	2.11	1.18
<i>Vehicular density: 400</i>	15	18	141	80	50	19	12	71	17	33	12
Energy (J)	3.59	3.59	3.59	5.36	2.62	1.15	3.59	3.59	1.77	3.59	3.59
PDR	1	3	2	5	2	1	3	4	2	3	2
PD (s)	1.51	1.36	4.89	2.31	3.09	3.76	1.82	6.64	1.24	1.36	1.69
Throughput (kbps)	2.02	1.93	8.26	4.81	2.78	2.95	1.32	1.51	1.11	1.55	1.32
<i>Vehicular density: 600</i>	21	19	21	30	85	100	29	212	45	10	28
Energy (J)	3.59	3.59	3.59	8.78	2.62	1.75	3.59	3.59	1.77	3.59	3.59
PDR	4.04	5.29	2.18	1.30	2.27	3.50	2.04	1.67	2.66	6.35	2.29
PD (s)	2	4	3	2	4	3	3	2	2	2	2
Throughput (kbps)	2.11	2.51	2.53	2.73	3.25	3.99	2.32	7.47	4.38	1.42	2.42
<i>Vehicular density: 800</i>	35	135	40	30	50	145	45	96	110	62	52
Energy (J)	3.59	3.59	3.59	4.90	2.62	6.16	3.59	3.59	6.82	3.59	3.59
PDR	3	1	2	5	4	3	1	3	3	3	3
PD (s)	8.16	2.05	2.21	3.67	2.48	2.95	1.38	2.34	7.12	6.27	1.13
Throughput (kbps)	1.05	7.59	1.82	1.25	2.00	7.07	2.19	6.13	4.45	2.33	2.34
<i>Vehicular density: 1,000</i>	25	45	42	45	152	22	170	145	89	170	95
Energy (J)	14.37	1.39	4.307	204.07	0.91	140.74	77.86	1.39	1.39	79.95	0.65
PDR	1	0.31	1.41	0.70	1.26	1.04	1.33	0.66	0.9	1	0
PD (s)	0.87	1.86	2.70	1.33	0.05	1.96	4.12	3.06	3.35	1.88	2.54
Throughput (kbps)	0.13	1.25	1.35	1.35	2.05	0.57	2.27	2.57	1.93	2.45	1.93

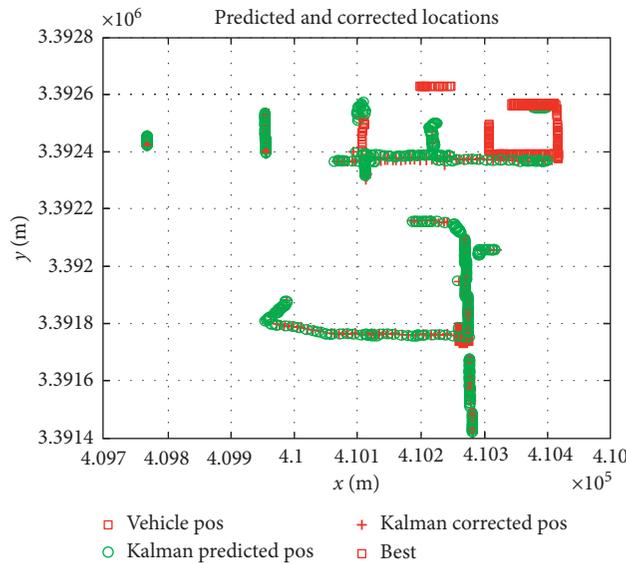


FIGURE 11: The corrected and the original location of the 1,000 vehicles for a single timestamp in the total simulation time.

the CH transmits it to the destination cluster member. The average amount of energy consumed at each node in a cluster for various data rates is shown in Figure 14. The consumption is observed to decrease or be constant for specific clusters for each data rate; this can be as the cluster topology has not changed, or the CH has not changed. The increase in energy is where the number of clusters is with

fewer members and a larger distance, which requires more dissipation of the energy.

The packet delay ratio and the packet delay are demonstrated in Figures 15 and 16, respectively. The delay in a packet can be associated with the scant cluster formation; the distance among the chosen nodes is far apart, the network's congestion due to the dense

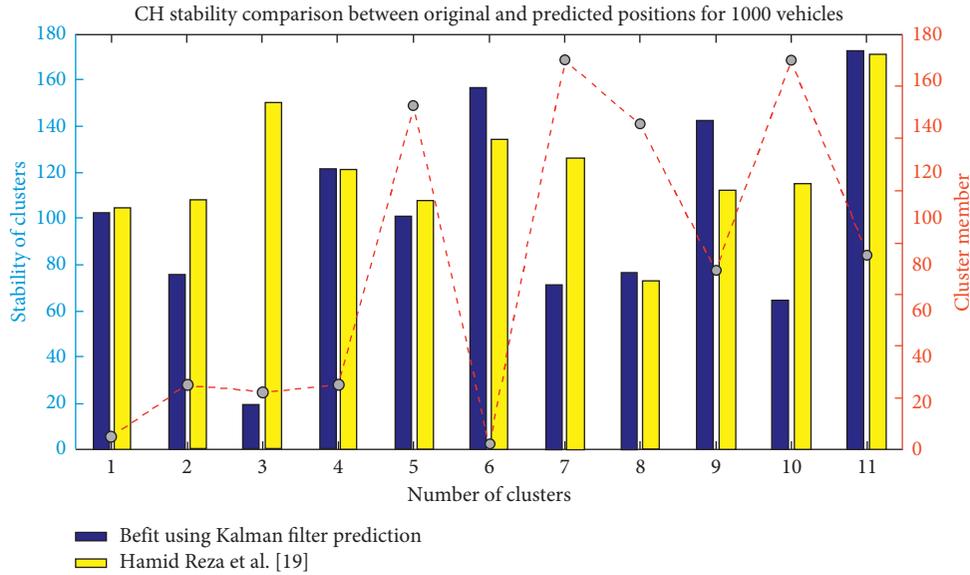


FIGURE 12: The comparison of the cluster stability for the befif factor using original and predicted 1,000 vehicles' position.

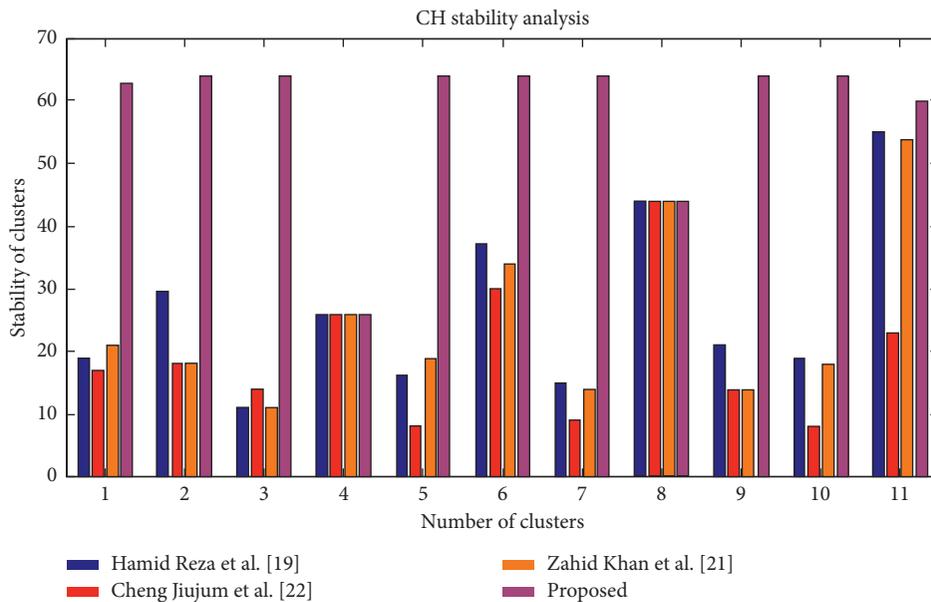


FIGURE 13: Comparative analysis of the CH stability for 1,000 vehicular densities.

population. The throughput is also analyzed, as shown in Figure 17. The throughput is the metric for the efficiency of the designed network. The high value of throughput resembles better performance with better communication

among the cluster members. The drop and loss of packets eventually affect the throughput of the network. The clusters with only one cluster member ought to drop the packet.

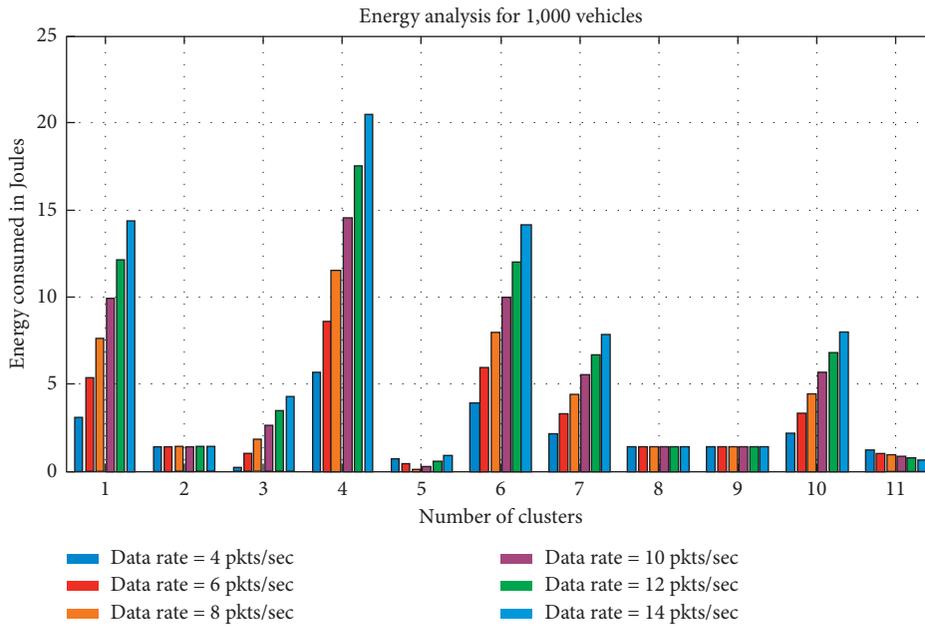


FIGURE 14: The energy consumption for various data rates at 1,000 vehicular densities.

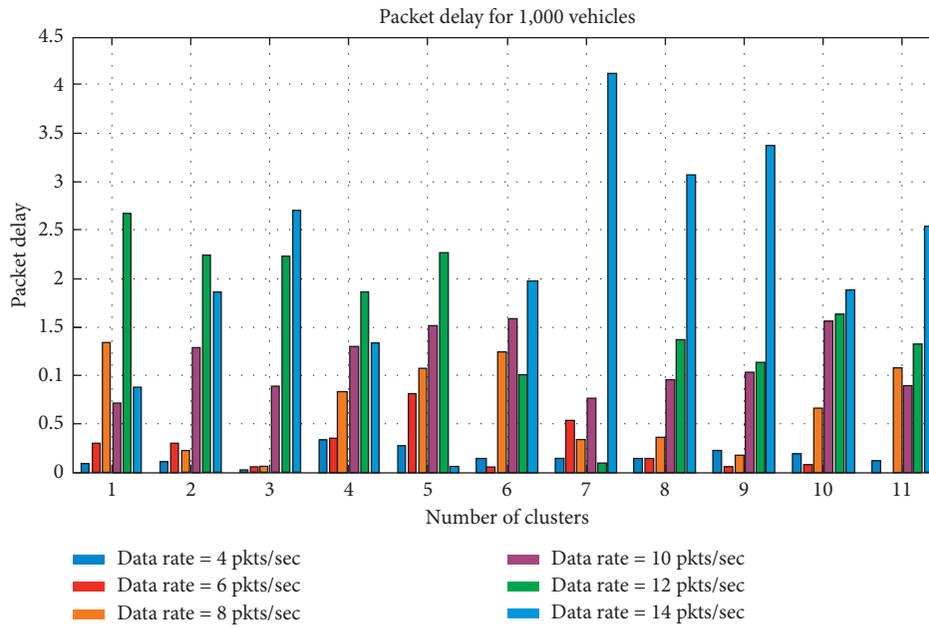


FIGURE 15: The packet delay for various data rates at 1,000 vehicular densities.

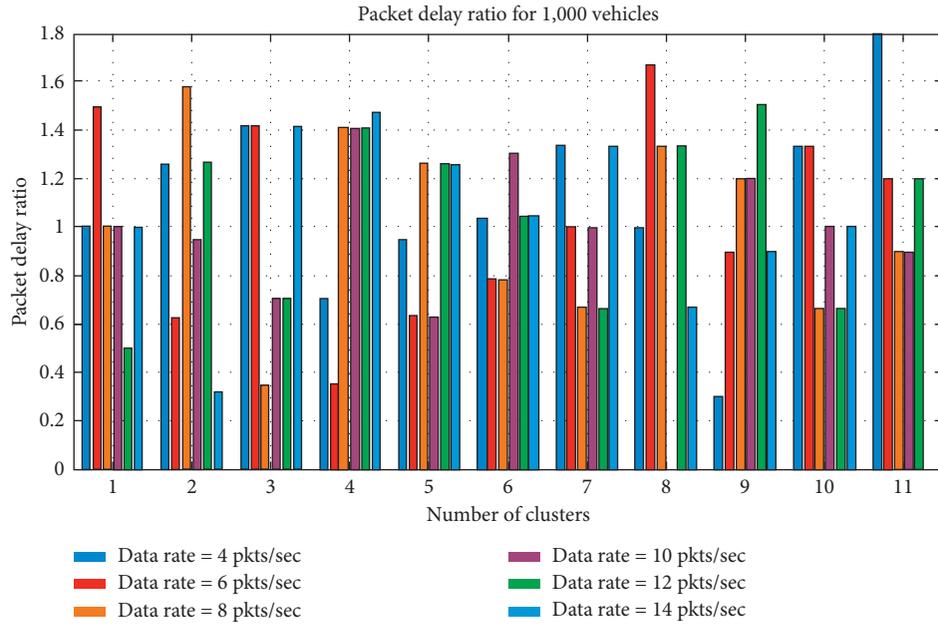


FIGURE 16: The packet delay ratio for various data rates at 1,000 vehicular densities.

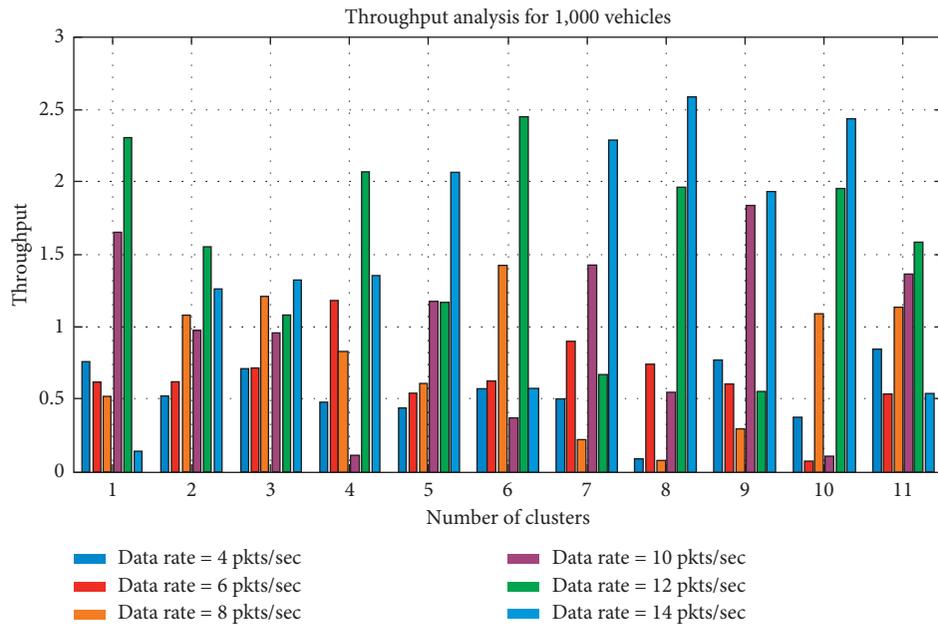


FIGURE 17: The throughput for various data rates at 1,000 vehicular densities.

5. Conclusion

In this paper, the authors have designed a new scheme for selecting a stable cluster head using the weighted approach. This formulation is crafted by including the four different metrics included to address all the parameters needed to enhance the dynamic network’s stability. The authors have designed a practical methodology for the deployment of the RSU working as an additional facility to enhance the network based on the angle of the lane. The clusters are formed concerning RSU. These clusters are further improved to select a stable cluster head using metrics: befit factor,

eccentricity, community neighborhood, and trust. The befit factor is dependent on the speed of the vehicle. VANET is a network with high-speed vehicles and dynamic topology. The precision in the exact location and speed in such networks is trivial to combat such a situation. The authors have employed the Kalman filter to predict the vehicle’s location at the next instant, which can improve the befit calculation and, in turn, suffice for the cluster head’s stability. The results have demonstrated superior results as compared to the original formulated befit factor. The next two metrics are designed on the evolving graph structure as they are scalable and eliminate the need for recalculating in case of a change

in the network's topology. The last is the trust value included for the primary users for multiple reasons; the primary users get hidden in the network. The detection of the primary users' energy is resolved using an LSTM, deep learning trained for different signals and noises. The accuracy is around 80%, with a misclassification rate of around 14%. The cluster head's stability is measured in terms of mode, where the number of times a vehicle is selected as cluster head is noted. The results show the weighted approach's supremacy compared to the cluster head stability achieved through a single metric. The designed method is tested on the real map for the region of Chengdu, southwestern China's Sichuan province, for the different vehicle mobility densities. Also, the scheme is tested for various integrated network parametric analyses with a two-hop structure. The results regarding throughput, packet delay, energy, and packet delay ratio have shown the proposed scheme's domination for different data rates.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors acknowledge the individuals who provided help during the research and preparation of the manuscript.

References

- [1] S. I. Boucetta, C. Z. Johanyák, and L. Károly Pokorádi., "Survey on software defined VANETs," *Gradus*, vol. 4, no. 1, pp. 272–283, 2017.
- [2] A. Paranjothi, M. S. Khan, and S. Zeadally, "A survey on congestion detection and control in connected vehicles," *Ad Hoc Networks*, vol. 108, Article ID 102277, 2020.
- [3] M. Akhtar and S. Moridpour, "A review of traffic congestion prediction using artificial intelligence," *Journal of Advanced Transportation*, vol. 2021, Article ID 8878011, 2021.
- [4] A. Ghansiyal, M. Mittal, and A. K. Kar, "Information management challenges in autonomous vehicles," *Journal of Cases on Information Technology*, vol. 23, no. 3, pp. 58–77, 2021.
- [5] S. Vodopivec, J. Bešter, and A. Kos, "A multihoming clustering algorithm for vehicular ad hoc networks," *International Journal of Distributed Sensor Networks*, vol. 10, no. 3, Article ID 107085, 2014.
- [6] S. Latif, S. Mahfooz, B. Jan, N. Ahmad, Y. Cao, and M. Asif, "A comparative study of scenario-driven multi-hop broadcast protocols for VANETs," *Vehicular Communications*, vol. 12, pp. 88–109, 2018.
- [7] R. T. Goonewardene, F. H. Ali, and E. Stipidis, "Robust mobility adaptive clustering scheme with support for geographic routing for vehicular ad hoc networks," *IET Intelligent Transport Systems*, vol. 3, no. 2, pp. 148–158, 2009.
- [8] P. Thakur and A. Ganpati, "A comparative study of cluster-head selection algorithms in VANET," in *Proceedings of International Conference on IoT Inclusive Life (ICIIL 2019)*, Chandigarh, India, December 2019.
- [9] A. Katiyar, D. Singh, and R. S. Yadav, "State-of-the-art approach to clustering protocols in VANET: a survey," *Wireless Networks*, vol. 26, no. 7, pp. 5307–5336, 2020.
- [10] Z. Zhang, A. Boukerche, and R. Pazzi, "A novel multi-hop clustering scheme for vehicular ad-hoc networks," in *Proceedings of the 9th ACM International Symposium on Mobility Management and Wireless Access*, Miami Beach, FL, USA, October 2011.
- [11] L. Zhang and H. El-Sayed, "A novel cluster-based protocol for topology discovery in vehicular ad hoc network," *Procedia Computer Science*, vol. 10, pp. 525–534, 2012.
- [12] S. Ucar, S. C. Ergen, and O. Ozkasap, "VMaSC: vehicular multi-hop algorithm for stable clustering in vehicular ad hoc networks," in *Proceedings of the 2013 IEEE Wireless Communications and Networking Conference (Wcnc)*, April 2013.
- [13] Y. Chen, M. Fang, S. Shi, W. Guo, and X. Zheng, "Distributed multi-hop clustering algorithm for VANETs based on neighborhood follow," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, p. 98, 2015.
- [14] D. Zhang, H. Ge, T. Zhang, Y.-Y. Cui, X. Liu, and G. Mao, "New multi-hop clustering algorithm for vehicular ad hoc networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1517–1530, 2019.
- [15] X. Ji, H. Yu, G. Fan, H. Sun, and L. Chen, "Efficient and reliable cluster-based data transmission for vehicular ad hoc networks," *Mobile Information Systems*, vol. 2018, Article ID 9826782, 2018.
- [16] G. H. Alshuhli, A. Khattab, and Y. A. Fahmy, "Double-head clustering for resilient VANETs," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 2917238, 2019.
- [17] A. Bello Tambawal, R. Md Noor, R. Salleh, C. Chembe, and M. Oche, "Enhanced weight-based clustering algorithm to provide reliable delivery for VANET safety applications," *PLoS One*, vol. 14, no. 4, Article ID e0214664, 2019.
- [18] X. Cheng and B. Huang, "A center-based secure and stable clustering algorithm for VANETs on highways," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 8415234, 2019.
- [19] H. R. Arkian, R. Ebrahimi Atani, A. Pourkhalili, and S. Kamali, "Cluster-based traffic information generalization in vehicular ad-hoc networks," *Vehicular Communications*, vol. 1, no. 4, pp. 197–207, 2014.
- [20] M. Mittal, R. K. Sharma, and V. P. Singh, "Modified single pass clustering with variable threshold approach," *International Journal of Innovative Computing Information and Control*, vol. 11, no. 1, pp. 375–386, 2015.
- [21] M. Mittal, R. K. Sharma, and V. P. Singh, "Validation of k-means and threshold based clustering method," *International Journal of Advancements in Technology*, vol. 5.2, pp. 153–160, 2014.
- [22] S. Kaur, R. K. Bansal, M. Mittal et al., "Mixed pixel decomposition based on extended fuzzy clustering for single spectral value remote sensing images," *Journal of the Indian Society of Remote Sensing*, vol. 47, no. 3, pp. 427–437, 2019.
- [23] H. Yang, Z. Jia, and G. Xie, "Delay-bounded and cost-limited RSU deployment in urban vehicular ad hoc networks," *Sensors*, vol. 18, no. 9, p. 2764, 2018.
- [24] Y. Mo, D. Yu, J. Song, K. Zheng, and Y. Guo, "Vehicle position updating strategy based on Kalman filter prediction in VANET environment," *Discrete Dynamics in Nature and Society*, vol. 2016, Article ID 1404396, 2016.

- [25] J. Liu, X. Wang, Y. Li, X. Kang, and L. Gao, "Method of evaluating and predicting traffic state of highway network based on deep learning," *Journal of Advanced Transportation*, vol. 2021, Article ID 8878494, 2021.
- [26] H. Zhou, D. Zhang, and K. Xie, "Accurate traffic matrix completion based on multi-Gaussian models," *Computer Communications*, vol. 102, pp. 165–176, 2017.
- [27] Z. Khan, P. Fan, S. Fang, and F. Abbas, "An unsupervised cluster-based VANET-oriented evolving graph (CVoEG) model and associated reliable routing scheme," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3844–3859, 2019.
- [28] J. Cheng, Q. Ma, R. Yu et al., "Research on the prediction-based clustering method in the community of medical vehicles for connected health," *IEEE Access*, vol. 7, pp. 71884–71896, 2019.
- [29] A. Saidi, K. Benahmed, and N. Seddiki, "Secure cluster head election algorithm and misbehavior detection approach based on trust management technique for clustered wireless sensor networks," *Ad Hoc Networks*, vol. 106, Article ID 102215, 2020.
- [30] M. Mittal, V. E. Balas, L. M. Goyal, and R. Kumar, *Big Data Processing Using Spark in Cloud*, Springer, Berlin, Germany, 2019, <https://www.springer.com/gp/book/9789811305498>.
- [31] A. Srivastava, A. Prakash, and R. Tripathi, "Fuzzy-based beaconless probabilistic broadcasting for information dissemination in urban VANET," *Ad Hoc Networks*, vol. 108, Article ID 102285, 2020.
- [32] X. Qian and H. Li, "On the performance of spectrum sensing in cognitive vehicular networks," in *Proceedings of the 2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, August 2015.
- [33] M. A. Saleem, S. Zhou, A. Sharif et al., "Expansion of cluster head stability using fuzzy in cognitive radio CR-VANET," *IEEE Access*, vol. 7, pp. 173185–173195, 2019.
- [34] X. You, C.-X. Wang, and Y.-C. Liang, "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Science China Information Sciences*, vol. 64, no. 1, pp. 1–74, 2020.
- [35] M. S. Kakkasageri and S. S. Manvi, "Multiagent driven dynamic clustering of vehicles in VANETs," *Journal of Network and Computer Applications*, vol. 35, no. 6, pp. 1771–1780, 2012.
- [36] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, "SUMO—simulation of urban mobility: an overview," in *Proceedings of the Third International Conference on Advances in System Simulation. SIMUL 2011*, pp. 1–6, Barcelona, Spain, October 2011.
- [37] A. Wegener, M. Piórkowski, M. Raya, H. Hellbrück, S. Fischer, and J.-P. Hubaux, "TraCI: an interface for coupling road traffic and network simulators," in *Proceedings of the 11th Communications and Networking Simulation Symposium*, pp. 155–163, Ottawa Canada, April 2008.