CSBR: A Cosine Similarity Based Selective Broadcast Routing Protocol for Vehicular Ad-Hoc Networks

Ankur Nahar
Computer Science & Engineering
Indian Institute of Technology
Jodhpur, India
nahar.1@iitj.ac.in

Himani Sikarwar
Computer Science & Engineering
Indian Institute of Technology
Jodhpur, India
himani.1@iitj.ac.in

Debasis Das
Computer Science & Engineering
Indian Institute of Technology
Jodhpur, India
debasis@iitj.ac.in

***Abstract*—Vehicular ad-hoc networks (VANETs) inflict smart traffic control using information exchange (i.e., warning messages, routing and network discovery messages) with nearby vehicles immediately in a highly dynamic environment. Prior approaches have proposed several routing computations and discovery techniques using prediction, direction aware, and hello message assisted algorithms. However, most of them do not perform adequately in real-life traffic situations and have various concerns i.e., broadcast storming, unpredictable network dynamics, fixed-route life-time problem, non-linear real-world traffic, high routing overhead, and low packet delivery ratio. Clustering architecture in vehicular scenarios improves resource utilization, network scalability and flow of data between communicating vehicles but stable cluster formations is also a challenging task in real-life traffic situations. Therefore, in this paper, we proposed a cosine similarity based selective broadcast routing protocol, also known as CSBR, which leverages non-linear cluster formation ability using cosine similarity index. Distinct clusters and the coordinating vehicles assist each other in finding the most suitable path to reach the destination. Additionally, a probabilistic forwarding approach is used to disseminate routing messages further in the network. The outcomes exhibit that the proposed scheme improved 5-10% packet delivery fraction (PDF), minimizes average delay approximately 25%, up to 10% low communication overhead, improved throughput upto 5-10%, and less neighbour discovery messages overhead compared to the existing broadcast routing protocols in VANETs.**

***Index Terms*—Vehicular Ad-Hoc Networks (VANETs), Clustering, Broadcasting, Routing.**

I. Introduction

The rapid growth in the transportation system increases eminence traffic on the streets [1]–[3]. In consequence of this, plenty of road accidents and traffic congestion has been noticed in recent times. In a report by the world health organisation (WHO) affirms that road traffic accidents touched 1.35 million deaths per year. Thus, the vehicular ad hoc networks (VANETs) aims to enhance road safety, information dissemination and traffic management applications [4]–[7]. Through VANETs, we can provide vehicle-to-vehicle (V2V) communication, which can avoid a crash and hazardous condition by propagating alert messages to other vehicles in the network. VANET is an intelligent traffic system, which is composed of interconnected vehicles and road-side units (RSUs) [8]–[10].

Fig. 1 represents a VANET scenario where the different vehicles in an urban environment are communicating with each other. RSUs are also placed in some sections of the city. Two nearby vehicles can transmit information to each other directly if they are in the communication range, but in case of a long-distance, they can communicate through RSUs. The angular similarity between vehicles is represented by the cosine angle, whereas linear distance is represented by Euclidean distance.

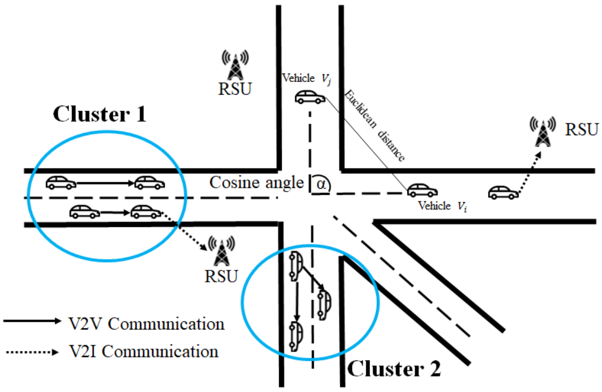

Fig. 1. *Example of A VANET V2V/V2I Network Scenario and Cosine Model*

The routing becomes a necessary step for network operations and accomplishing services [11]–[13]. Existing routing protocols in VANETs such as direction aware best forwarder selection routing (DABFS) [1] discover the route to the destination by broadcasting beacon messages to the neighbour vehicles. These broadcasting messages create a broadcast storm in the network, which degrades the network performance by creating a situation of congestion and high packet collision. Prediction based schemes such as Kalman prediction based neighbour discovery (KPND-AODV) [2] used a Kalman filter for predicting vehicular movement, which is not suitable for non- linear real-world environments. The CEG-RAODV [4] and CPB [5] used clustering architecture with a probabilistic forwarding approach, but the erroneous assumptions about the vehicular movement create a problem as the vehicle change its position before the

ISBN 978-3-903176-28-7 ©2020 IFIP

requesting vehicle discovers the path to the destination [14]–[18].

In this paper, the CSBR protocol aims to reduce the unnecessary broadcasts and provide an efficient approach for neighbour discovery and route discovery phase [19]–[22]. Our research aims to optimize routing processes using stable and robust clustering in real-world scenarios. The proposed scheme uses the concept of hybrid clustering and anycasting approach, in which we are limiting the broadcast messages only to the group of similar cosine indexed vehicles that form a cluster. Packet forwarding or dropping decisions are based on a probability factor and the coordination of the cluster members. The probabilistic approach provides a controlled environment to restrict propagating unwanted broadcast messages. This property helps to reach our prime objective to improve network performance in terms of throughput, less neighbour discovery message and control the broadcast storm in the network. A cluster head is responsible for the coordination among the members, which is chosen based on the distance, connectivity, and high connection probability metric. Different from previous clustering schemes, we use the angular movement of a vehicle and non-linear characteristics of real-world scenarios to form stable clusters [23], [24].

Implementing a real-world scenario is expensive to execute and test. Therefore, we use various mobility testbeds to simulate distinct existing routing protocols and CSBR. The simulation result demonstrates an improvement in the successful delivery ratio of a packet to an average of 97-98% and minimizes the average delay up to 25% in a dense network of 200 vehicles in an urban mobility scenario. Moreover, CSBR minimizes the routing overhead upto 10%, and enhanced the throughput of the network for the simulated environment up to 700-800 kbps, and showed an average of 10% improvement than other existing protocols used in this paper. The CSBR also decreases the number of broadcast packets needed to discover neighbouring vehicles.

The paper is organized in the subsequent sections: Section II, provides information about the literature and the past work done by different researchers in this area. In Section III, we discuss the objectives and contribution of the research. Section IV and V, contains the details about the network model and our proposed CSBR protocol in detail. Whereas, in section VI and VII, we present a performance analysis of simulated protocols and the simulation results.

II. Related Work

The area that is attracting lots of researchers is VANETs. Much prominent work has been carried out in this field, which studied routing approaches, issues involving in routing, and the solution to resolve these issues. Existing schemes based on prediction, opportunistic, hello messages, and clustering are described below:

A. Prediction and Opportunistic Based Routing Schemes

S. Haider et al. [1] introduced a direction based routing protocol for the bi-directional networks, such as VANETs. In their research, they have used relative position and the direction of a vehicle to determine the vehicle movements. C. Liu et al. [2] proposed a new technique using Kalman prediction for neighbour discovery and analyzed its effects on the routing. In their research, they found out that using Kalman prediction method neighbour discovery, the time was reduced, and the performance of the routing was also significantly improved. M. Naderi [3] proposed an adaptive routing technique to send beacon messages to the neighbours. They have used opportunistic routing strategy and the VANETs characteristics such as speed, position, and the direction of drift of the vehicles in the network to enhance throughput and the average successful packet transmission. In another research conducted by M. Toohani et al. [7] and F. Li and Y. Wang [8] discussed the challenges in VANETs routing and protocol designing. In their study, they discussed many routing protocols including geocast, hierarchical, opportunistic and positional protocols. S. Han [9] proposed an adaptive hello message technique which can reduce the overhead on periodic hello message-based routing protocols. T. Saini and S. Chandra [10] surveyed various routing protocols including proactive, reactive, geo-positional, QoS, anycast, multicast, and multipath routing. In their article, they covered various techniques and the protocols that can be useful in VANETs.

Y. Tang et al. [11] proposed a technique that uses machine learning and artificial intelligence techniques to predict the mobility of a vehicle. In their research, they found, if the movement of a vehicle can be determined, then a better routing scheme and improved performance can be achieved. D. Zhang et al. [12] proposed a substantial self-adaptive routing algorithm based on the link continuation model. In their research, they used mobility to design their routing protocol by using a heuristic search technique. A. Ghaffari [13] proposed a hybrid opportunistic technique for routing in VANETs, which used position based information and the greedy method, in which the highest geographic progressive vehicle chosen for forwarding the packet to the destination. H. Abbasi et al. [14] proposed an intelligent forwarding protocol and used handshake less communication between vehicles and also decoupled acknowledgements (ACKs) from the message dissemination process to reduce the delays. For faster retransmission, euclidean distance and the signal to noise ratio (SNR) has been used to optimize the contention window.

B. Cluster-Based Routing Schemes

Z. Khan et al. [4] introduced a cluster-based VANETs oriented evolving graph (CVoEG) model routing technique based on a bi-direction highway model. They used the Eigengap heuristic strategy for packet forwarding. L. Liu et al. [5] proposed a data dissemination technique that relied on the clustering and a probabilistic approach of broadcasting. In their research, they used conditional probability for packet forwarding decisions. In another research conducted by S. Haider et al. [6] simulated a cluster-based V2V environment to reduce collision on highways. A probability-based, direction-aware technique has been used to provide early warning messages to the other vehicle to reduce the chances of accidents.

In general, various routing protocols have been proposed for vehicular communication for performance enhancement, but designing an effective routing protocol is a challenging task for VANETs. Besides the vehicle's high mobility, the urban environment and traffic complexity is an influential parameter that influences the performance factor of a vehicular network. Improving cluster stability is another crucial factor that affects routing techniques in cluster-based routing. However, existing routing techniques suffer from local optimum cluster selection and less stability. Concerned by the stable clustering issue, our work aims for link stability, and efficient cluster formation and performance enhancement.

III. Motivation and Contribution

As described in the introduction section, optimized routing in clustered architecture depends on effective clustering and corporative message forwarding between the vehicles. Numerous researches attempted to solve the cluster stability, low coherence, and vehicle coordination problem, but they showed performance degradation in real-life traffic situations due to highly dynamic topology and incorrect assumptions. The correct analysis of vehicle characteristics such as velocity, direction, and angular movement assists in creating more stable clusters and better information forwarding behaviour. Using the message delivery probability factor helps to decrease the unwanted message broadcast in the network. Moreover, it increases the routing efficiency and optimizes the routing performance. Our design focuses on three challenges: i) How to efficiently create fully associate clusters of the vehicles; ii) How to choose the cluster head for each cluster and information forwarding relay toward the destination vehicle using coordinating vehicles in the cluster; iii) How to achieve high throughput and less overhead in VANETs.

CSBR makes use of the concept of cosine similarity indexing and cosine distance for the clustering in VANETs. Cosine similarity provides an efficient way to create a cluster of vehicles with the same characteristics in terms of the direction and movement of a vehicle. For the route discovery and path identification, CSBR identifies the cosine similarity between the communicating vehicles and finds cluster heads through cosine distance. Then it defines a relationship between cluster heads and the cluster members that coordinate among themselves and uses probability functions to reduce the broadcast density in the network. Our main contribution of the research are as follows:

- We propose a novel cluster-based routing protocol for the VANETs environment known as CSBR, which efficiently create stable clusters and use them for effectively route a packet to the destination.
- We introduce the cosine similarity concept to finding stable clusters and then use cosine distance to elect cluster heads among them to maximize the performance.
- CSBR protocol decrease unwanted broadcast messages by selectively sending, only to the cluster members based on a probability factor defined in proposed route selection algorithm. It allows the routing protocol to minimize broadcasts in the network and discover the optimum path to the destination.
- We evaluate the performance through extensive simulations and done comparison analysis with existing schemes DABFS, KPND, CEG-RAODV, and CPB in term of pdf, average delay, routing overhead, throughput, and neighbour discovery messages overhead.

IV. Network Model

For CSBR protocol, we develop a simulation environment which is considerable for VANETs infrastructure. Based on this model, we consider N numbers of vehicles such that $\{V : V_1, V_2, V_3 \ldots V_N\}$ with the M number of road-side units $\{R : R_1, R_2, R_3 \ldots R_M\}$. The mobility model for the connectivity can be represented by a graph G =(V, E), Due to the real-time compulsions and resource limitation, here we exercise short-range communication given according to dedicated short-range communications (DSRC) (i.e., 300 - 1000m) between vehicles. Each vehicle in a VANET broadcasts a traffic safety message every 100-300 ms, which keeps the vehicle's driving-related information, such as location, speed, turning intention, and driving status (e.g., regular driving, waiting for a traffic light, traffic jam, etc.), to other vehicles [20]. OBUs (on-board units) are used as the packet originator and the packet receiver. We assume that each vehicle's OBU contains geographical information by using the global positioning system (GPS). This location information can be exchanged with the other vehicles in the communication along with the routing table information by hello messages using triggered updates whenever there are any changes in the network.

The urban mobility scenario, which is a manhattan model [16] simulated for the block of Oslo city, Norway region with live traffic simulation using OsmWebwizard.

$$\begin{cases} \eta(t) - \eta(t^{'}) + 2\pi, & \text{for } -2\pi < \eta(t) - \eta(t^{'}) \leq -\pi \\ \eta(t) - \eta(t^{'}), & \text{for } -\pi < \eta(t) - \eta(t^{'}) \leq \pi \\ \eta(t) - \eta(t^{'}) - 2\pi, & \text{for } \pi < \eta(t) - \eta(t^{'}) \leq 2\pi \end{cases} \tag{1}$$

Here $\Delta\eta(t)$ is the directional difference between the new direction $\eta(t)$ and old direction $\eta(t^{'})$. In our experimental setup, the vehicles can move along the path of a territory. Each vehicle can communicate with another vehicle if both are in the communication range of each other. The vehicles can communicate straight to the road-side units if other vehicles are not in the communication range. The street structure enforces the vehicle to change its direction and its movement pattern. The intersections and traffic lights are used for traffic control management. The on-board units on the vehicle help to communicate with the other entities in the network.

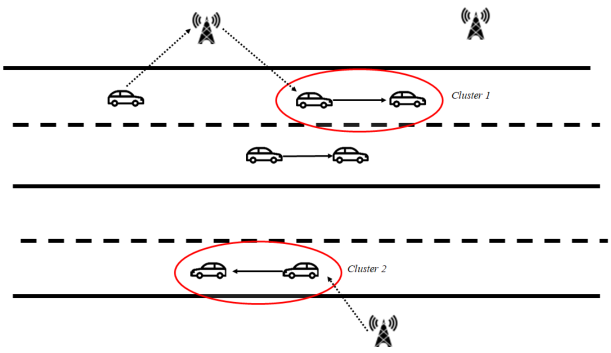

Fig. 2. Network Model for VANET Simulation

Fig. 2 shows the communication model between the vehicles and the road-side units (RSUs). If two vehicles are nearby each other, then they can communicate directly, but in case of a long-distance between two vehicles, they can communicate through RSUs. It can be assumed for the simulated urban city block that lane changing, acceleration, and direction-changing taking place on the road. The random distance between the vehicle can be given by $\hat{k}$ using eq. (3), which is the cosine distance between the vehicle. A vehicle V_i can move in the direction of another vehicle V_j, or it can move towards the opposite direction. The speed and the direction are varied according to the relative velocity of the other vehicles. Vehicles with the same direction and relative speed can form a cluster according to the cosine similarity (ε) using eq. (2). These clusters have been further utilized by the CSBR to discover the most suitable route towards the destination.

A. Basic of Cosine Similarity and Cosine Distance

The Cosine similarity is often used to analyze entities in a multidimensional space that measures the cosine angle between two vectors [17]. In our proposed protocol, we use cosine similarity to determine the directional relationship among the vehicles, where we measure the orientation of a vehicle and neglect the magnitude values. The resultant cosine similarity used to generate a metric that explains how two vehicles are associated with each other by considering the angle of movement. Fig. 3 Shows the mapping of a vehicle movement in a vector space that is the direction along the line in the multidimensional space.

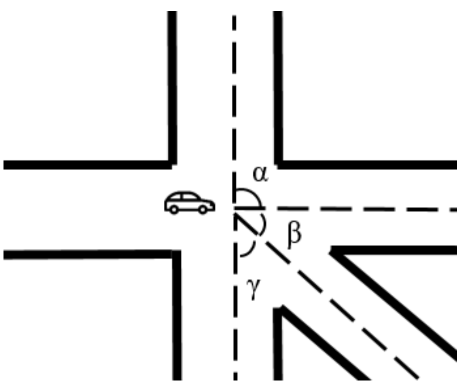

Fig. 3. Vehicle directional movement at intersection

Here, a vehicle $\vec{V_i}$ positioned at the intersectional roadblock where it can move in any direction. The angle of movement is given by an angle of α, β, and γ. If the two vehicles $\left\{\vec{V_i}, \vec{V_j}\right\}$ are moving in the same direction, then we can consider that the vehicles are moving parallel.

B. Cosine Similarity Measures

Since $\alpha \in [0, 2\pi]$, a vehicle can drive in any direction in front of an intersection, for example, left, right, straight or 180-degree turn.

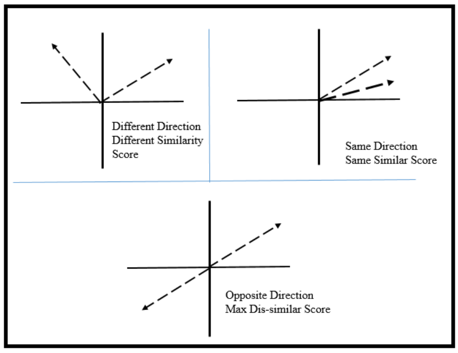

Fig. 4. Cosine similarity measures

Fig. 4 shows the similarity measure between the vehicular directional movements. Here, we can see if the movement of vehicles is in the same direction as the cosine similarity score is the same between the vehicles. In other conditions, the similarity score differs between the vehicles [18].

V. CSBR: A Cluster-Based Selective Broadcast routing protocol for vehicular Ad-Hoc Networks

This paper proposed a cluster-based selective broadcast routing technique that can enhance the neighbour selection method and reduce the broadcast storm prompted by the RREQ packets. CSBR can enhance network efficiency in terms of high packet delivery ratio, low average delays, low routing overhead, high throughput, and less neighbour discovery message overhead in VANETs. CSBR protocol operates in two phases: firstly, it formulates distinct clusters based on the cosine similarity index among vehicles and then selects the cluster head, which will communicate to the other cluster heads in separate clusters. The vehicle with high association probability and the lowest cosine distance metric selected as cluster head. Secondly, it discovers the optimum route from source to destination utilizing cluster heads.

CSBR protocol is inspired by the similarity clustering method. Firstly, it formulates primary clusters, which is based on the similarity index produced by the cosine similarity using vectored information about the vehicle i.e., velocity, direction, location, and the vehicle ID. Cosine similarity commonly refers to angular similarity or angular distance. The cluster formed by similarity index is independent of sequential ordering. It later handles the route discovery algorithm for each cluster individually. The clustering technique is also used for the neighbour discovery process. Using eq. (2) we can find the cosine similarity between two vehicles i.e., V_i and V_j in a vehicle set V_n. The cosine similarity can be defined as:

$$\varepsilon = \frac{\sum_{i=1}^{N} \vec{V_i}\vec{V_j}}{\sqrt{\sum_{i=1}^{N} \vec{V}_i^2}\sqrt{\sum_{i=1}^{N} \vec{V}_j^2}} \tag{2}$$

Here $\vec{V_i}$ and $\vec{V_j}$ is the i^{th} and j^{th} vehicle's vector information list. Each vehicle V_i has associated with a vector information metric values $\vec{V_i} = \left(\vec{V_1}, \vec{V_2}, ..\vec{V_j}..\vec{V_n}\right)$, where $\vec{V_i}$ represents a value which indicate associativity between vehicles. The cosine distance model is used to find the distance between the communicating vehicles. It can define as:

$$\hat{k} = \{1 - \varepsilon\} \tag{3}$$

Where $\hat{k}$ is the cosine distance and the ε is the cosine similarity. This distance is further used as one of the parameters for the cluster head selection by comparing it to the threshold distance ($\hat{\rho}$) defined in eq. (8). We used the following steps for identifying cluster heads in the network.

- Find the cosine similarity of the vehicles for clustering purpose
- Find the vehicle which has less distance ($\hat{k}$) than a threshold distance ($\hat{\rho}$) from the other vehicle
- Find the adjacency matrix ($neiMat_{ij}$) for every vehicle in the network

The connectivity among the vehicles in a cluster can be determined by the similarity index, in which the vehicles can communicate with each other. If the similarity index value between the vehicles is approximately equivalent to each other, then they are considered in the same cluster. In this method, we also employed an opportunistic approach [16] where unlike traditional protocols, a vehicle does not transmit its data packets to the most adjacent vehicle but prefer to transmit the packet to the cluster head or a selected vehicle in the network. Now, this vehicle concludes how to forward the packet further in the network. Our approach encourages the formulation of clusters of vehicles where some vehicles act as cluster heads to transmit information further. The cluster head is selected by calling the cluster head selection approach.

Table I, contains the symbol table used in the proposed routing scheme. The symbol table lists all the parameters and the details used in the algorithm.

TABLE I
SYMBOL TABLE FOR CSBR PROTOCOL ALGORITHM

Variables	
Name of Variable	***Description***
$V\{\}$	Vehicles set in the network V = $\{V_1, V_2, V_3...V_n\}$
ε	Cosine similarity for forming the clusters
$neiMat_{ij}$	The adjacency matrix of a vehicle
χ	Cluster head set $\chi = \{\chi_1, \chi_2, \chi_3...\chi_n\}$
$ClustM_\beta$	Member vehicles of a cluster $ClustM_\beta$ = $\{ClustM_1, ClustM_2...ClustM_n\}$
$\hat{k}$	cosine distance based on the cosine similarity between i^{th} and j^{th} vehicles
$\hat{\rho}$	Threshold distance
$OptimRoute_\gamma$	Minimum path to the destination
P_κ	Probability function
$\vec{V_i}$	Vector information metric of a vehicle V_ii.e., velocity, direction, location, and vehicle ID
$Rand()$	Random value generator function

A. CSBR Protocol Algorithm

As stated above, the proposed algorithm work in two steps.

1) Cluster Head Selection and Neighbour Discovery: In the first step, the cluster head selection is carried out by the vehicles in the network. The cluster head selection procedure assigns priority to each vehicle based on the distance metric and the connectivity parameters. The cluster head selection procedure starts with the computation of the cosine similarity index. The similarity index is used to determine the cluster size and the cluster members, which fall into the similar index values. The similarity index is calculated using eq. (2). After the similarity calculation, the distance connectivity matrix is calculated. This connectivity matrix is used to identify the adjacency (connectivity to other vehicles) of a vehicle. The adjacency matrix allows representing the network with the *N*N* matrix.

$$neiMat = [f(i,j)] \tag{4}$$

Where each element represents the connectivity of the vehicles. The formula calculates the adjacency matrix:

$$neiMat[i,j] = \begin{cases} 0, & \text{if } i = j \\ neiMat(i,j), & \text{if } i \neq jand(i,j) \in E \\ \infty, & \text{if } i \neq jand(i,j) \notin E \end{cases} \tag{5}$$

Here i and j is the i^{th} and j^{th} vehicle in network and G is the network graph with vehicles V. After calculating the adjacency matrix, distance form a vehicle to another vehicles in a cluster is calculated.

Before calculating this distance, some information is needed by the vehicle such as vehicle id, current position, and the velocity of the vehicle. This information is acquired by broadcasting the neighbour discovery packets. Once this information is obtained by a vehicle, it can calculate the distance from the other vehicles. The following eq. (6) and (7) can calculate the angular distance and the angular similarity of communicating vehicles:

$$\hat{k}_\theta = \frac{\cos^{-1}(\varepsilon)}{\pi} \tag{6}$$

$$\varepsilon_\theta = 1 - \hat{k}_\theta \tag{7}$$

Here the $\hat{k}_\theta$ is the angular distance between the vehicles and ε_θ is the angular similarity between the vehicles. ε is the similarity index value of a vehicle. After finding the coordinates of a vehicle, distance can

Algorithm 1: Cluster Head Selection Based on the Cosine Similarity Index in CSBR Protocol

Input: $neiMat_{ij}$, $V\{\}$
// The connectivity $neiMat_{ij}$ of i^th vehicle V_i in a defined vehicle set V
Output: cluster head χ_i
Data: vehicles set V_n
Calculate_connectivity_matrix$\{neiMat_{ij}\}$
Cluster_selection_procedure$\{neiMat_{ij},\ \hat{k}\}$
for *i=1 to N* **do**
    // Calculate cosine distance ($\hat{k}$) based on the cosine similarity index
    $\varepsilon = \frac{\sum_{i=1}^{N} \vec{V}_i \vec{V}_j}{\sqrt{\sum_{i=1}^{N} \vec{V}_i^2}\sqrt{\sum_{i=1}^{N} \vec{V}_j^2}}$
    // ε is the cosine similarity index
    $\hat{k} = \{1 - \varepsilon\}$
// Calculate threshold distance $\hat{\rho}$
$\hat{\rho} = 1/N_i \sum_{i=1}^{N} \hat{k}$
find_cluster_head $\{neiMat_{ij},\ \hat{\rho}\}$
if $neiMat_{ij} \leq \hat{\rho}$ **then**
    if $neiMat_{ij}$ == $maxneiMat_{ij}$ **then**
        return χ_i
    Exit

be calculated by the cosine distance formula given in eq. (3).

In the last step, the selection of the cluster head is performed by comparing each vehicles cosine distance to the other vehicles and the adjacency matrix of the vehicles in the cluster. Here the cosine distance is compared with the threshold distance, which is calculated by the formula:

$$\hat{\rho} = 1/N \sum_{i=1}^{N} \hat{k} \tag{8}$$

Here N is the number vehicle in the cluster and $\hat{k}$ is the distance between vehicles. The vehicle with the lowest cosine distance from other vehicles and has maximum adjacency values are selected as the priority vehicle or the cluster heads. Now, this vehicle will be responsible for forwarding the data packets to the destination or discovering the shortest path to the destination.

2) Route selection in CSBR: After selecting the cluster head for each cluster in the network, the next phase of the route discovery takes place. Priority vehicles in the cluster are responsible for the coordination and obtaining the most suitable path to the destination. One of the cluster member vehicles initiates route discovery message by generating RREQ packets. If the receiving vehicle is the cluster head, then it will transmit the RREQ packet to the succeeding hop or nearest cluster head with the probability of 1. But if the receiving vehicle is the member vehicle of the cluster, then it will decide to either forward the packet or drop the packet by a probability function. The packet forwarding probability function can be given by:

$$P_\kappa = 1 - e^{-\left(\frac{\hat{k}_i}{\hat{\rho}}\right)} \tag{9}$$

Here P_κ is the probability function, $\hat{k}_i$ is the cosine distance value of a vehicle V_i. Now, this probability value is compared with a random value generated by the *rand()* function. If the probability is higher than the random value, then the packet is forwarded to the nearest vehicle otherwise, the vehicle decides to drop the packet. This process continues until the request reaches the destination. After this, the path with the lowest end-to-end delay is chosen as the shortest path.

Algorithm 2: Route Selection in CSBR Protocol

Input: Initialize all clusters $Clust_\alpha$, cluster members $ClustM_\beta$
// $\alpha = 1, 2, ...N$, $\beta = 1, 2, ...M$
Output: Minimize path to the destination ($OptimRoute_\gamma$)
Data: Vehicle set V_n, Cluster Heads χ_i
// Route discovery using cluster heads
1 *Find_Route* $\{OptimRoute_\gamma\}$
// Send route discovery request to the cluster head
2 **for** *i=1 to* α **do**
    // Generate RREQ messages based on the probability function
3   $P_\kappa = \begin{cases} 1, & \text{if } cluster_{dist} = min \\ 1 - e^{-\left(\frac{\hat{k}_i}{\hat{\rho}}\right)}, & \text{otherwise} \end{cases}$
4   Fwd_route_Req $\{ClustM_\beta,\ Clust_\alpha\}$
5   **if** V_i == χ_i **then**
        // if receiving vehicle is cluster head
6   *go to step 11*
7   **if** V_i == $ClustM_\beta$ **then**
        // if receiving vehicle is Member vehicle
8   **if** $P_\kappa \leq rand(0, 1)$ **then**
9   *drop the request packet*
10   *forward to nearest vehicle within the cluster*
11 *forward the RREQ request with the (P) probability to the nearest cluster-head* (χ_i) *and wait for RREP messages*
12 *Cal_RREP_Delay* $\{\}$
13 Store: $OptimRoute_\gamma$ // Minimize path
14 **Exit**

The above presented *Algorithm 1* (i.e., Cluster head selection based on the cosine similarity index in CSBR Protocol) illustrates the selection of the cluster head and the coordinating cluster member vehicles. Cosine distance in used to find a more stable cluster based on the angular direction, which further solves the fixed-route lifetime problem by generating a new route based on cluster formations. *Algorithm 2* (i.e., Route selection in CSBR protocol) use a probabilistic approach for route discovery, which helps to achieve the solution to our prime concern i.e., broadcast storming in the network. Due to more stable cluster formation and optimum route discovery using cluster heads, the packet delivery ratio of the network also increases.

VI. Performance Analysis

In this paper, we have taken multiple ad-hoc protocols in consideration, such as DABFS, KPND-AODV, CEG-RAODV, CPB, and CSBR. DABFS [1] is based on the route discovery using beacon messages broadcasting to the neighbour vehicles. KPND-AODV [2] used a Kalman filter for predicting vehicular movement, whereas The CEG-RAODV [4] and CPB [5] used clustering architecture with a probabilistic forwarding approach to predict route to the destination. These protocols are examined in a VANET scenario in the simulation.

A. Simulation Environment

For simulation designs, we use the *ns2.35* tool, which is an OTcl based network simulator. Here, we take distinct bandwidth constraints and vehicle density in urban mobility. The system configuration, which is used for the simulation, is based on 9^{th} *generation intel i5* processor with 8 GB on-board ram on the Linux platform.

We created the scenario by the WebWizard tool, which is used for reproducing live traffic situations. Here we use the block of Oslo city, Norway, for the traffic simulation purpose. The Manhattan mobility model [16] is used as the simulation environment, as it is the most suitable model for urban mobility simulation. The implemented protocol simulates the dynamic characteristics and scenarios of real networks like congestion in the dense traffic areas, link breaks due to road conditions, drops behaviour due to the peak hour condition, and the limited link capacity. The dynamic behaviour due to the delays on different paths and acceleration of a vehicle, enables the vehicles to split the outgoing traffic among multiple distinct paths.

B. Vehicles Parameter Configurations

In this section, we define various parameters for the simulation purpose. We conducted simulations using ns2.35 tool using realistic parameters and scenarios. For real-life traffic generation, we integrate SUMO tool. The simulation parameters are listed in Table II depicts the parameters and the values used in the simulation.

TABLE II
Vehicles Configuration in Urban Environment

Parameter Name	*Value of the parameter*
Outgoing packet size	512 bytes (based on 1500-byte MTU)
Outgoing rate	As many packets as possible per second
Incoming packet size	0 bytes (off))
Incoming rate	0 (off)
Runtime	150 seconds
Receive buffer	50
Send buffer	50
Traffic Source	TCP/UDP
Area	5000*5000
MAC Standards	IEEE 802.11p, WAVE
Tool	NS-2.35
Mobility Model	Manhattan Model (Urban Scenario)
Classifier Model	Gaussian Mixture Mobility
Data Rates	2 Mbps
Transmission Range	300 – 1000 meters
Vehicle Density	25 - 200 Vehicles
Velocity of Vehicles	0 ∼ 20 (m/s)

C. Performance Measures

For simulation, we implemented our proposed protocol naming CSBR and also compared it with the recently proposed routing techniques in terms of following performance measures:

1) Packet Delivery Fraction: We can measure PDF in VANETs by determining the total packets received by a destination vehicle *(V_d)* in communication, ratio to the cumulative send packets by the communicating source vehicle *(V_s)*.

2) Average Network Delay: The average delay is specified as the average time (sec or ms) taken by the packets to arrive at the destination. The average delay depends upon the transmission mechanism and the intermediate vehicles.

3) Routing Overhead: VANETs routing load represented as the packets processed by the vehicles in communication. For this objective, we applied a counter to estimate the amount of routing packets.

4) Throughput: It is the amount of data packets successfully transmitted and received from origin to destination in a unit interval is called throughput of a network. It is typically measured in *bits/sec*.

5) Neighbour Discovery Message Overhead: Every vehicle in VANETs collects the information of neighbours by broadcasting neighbour discovery packets. The discovery message overhead signifies the number of messages transmitted to identify neighbour vehicles.

VII. Simulation Results

The simulation results based on the performance parameters outlined earlier are depicted in the subsequent illustrations.

Fig. 5 indicates the network density influence on the ratio of the successfully delivered packet. As represented, the CSBR protocol outperforms the other protocols and possesses the highest ratio of the successfully delivered packet due to less amount of packet drops in the network. CSBR uses similarity index and connectivity matrix for electing cluster heads, which produce more reliable routes to the destination. These cluster heads transmit packets to the destination using high forwarding probability, which increases the packet delivery ratio.

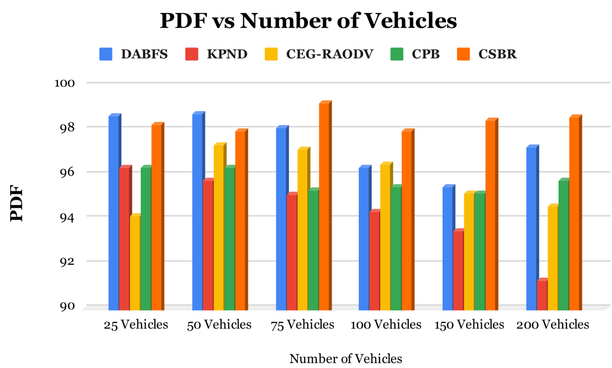

Fig. 5. Packet Delivery Fraction

As exhibited in Fig. 5, when we have scattered traffic situations such as 25 to 50 vehicles on the road, CSBR functions similar to DABFS and CEG-RAODV as the probability of forwarding a packet is less as there are fewer number of vehicles on the road. But as the traffic moves to denser conditions such as 100-200 vehicles on the road, CSBR performs more efficiently than all other compared protocols as the probability that a message will be forwarded towards the destination is increased. Here KPND represents the worst packet delivery ratio in high denser vehicular scenarios as Kalman filters are most suitable for the linear system. In contrast, highly dynamic real-world scenarios are mostly non-linear.

Fig. 6 indicates the average delay in simulated conditions. As represented, the CSBR protocol has a marginally moderate delay from the other protocols. CSBR uses cosine distance as one of the primary parameters for electing cluster heads, which further plays a primary role in packet forwarding. The selected cluster heads are used for the packet transmission, and these cluster heads work as the relay nodes to forward the information. Due to this, the average delay is quite low in the CSBR protocol.

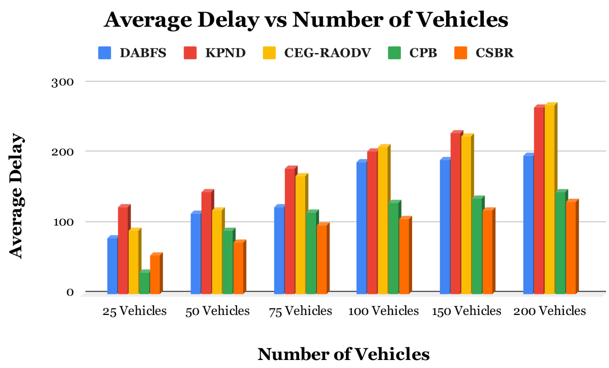

Fig. 6. Average Delay

As shown in the illustration, CSBR substantially outperforms all other existing protocols in terms of average delay. KPND-AODV possesses the highest delay amongst all other protocols as it demands substantial processing time in the neighbour discovery process, whereas CSBR has the lowest delay in the network due to passive clustering. All information regarding cluster members, position, direction, and speed sent in the packet header itself, which gradually decreases delays and processing time. Here, we can conclude that as the number of vehicles rises, the average delay also rises. As we already know, there are more packet collisions in the denser network. It increases the probability of the retransmission of the data packets as well as environmental interference incurs more transmission delays.

Fig. 7 indicates the comparison of routing overhead between broadcast and cluster-based routing schemes. We can see from the graph that the KPND-AODV protocol has the lowest routing overhead. In contrast, CSBR protocol performs respectively better then DABFS as vehicles need to process less number of broadcast packets during path discovery and neighbour discovery phases.

CSBR has low overhead in sparse traffic conditions as less number of packets need to be processed. But as the traffic increases, CSBR complexity increases resulting in high overhead. The clustering and forwarding technique

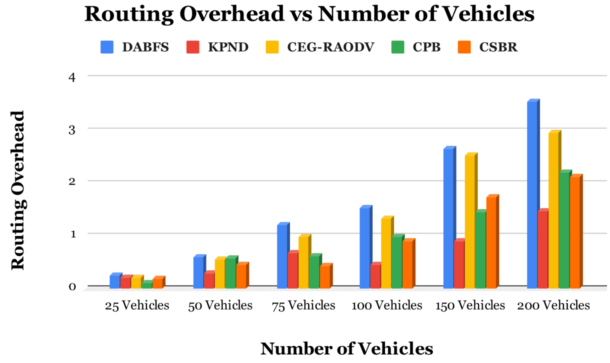

Fig. 7. Routing Overhead

has a significant impact on the routing overhead. Since, we know, CPB also uses a probabilistic approach to forwarding data packets further in the network. It has almost the equivalent routing overhead whereas CEG-ROADV has more routing overhead due to additional computation required to calculate an optimal number of clusters.

Fig. 8 indicates the effect of network density on the throughput. The graph represents a comparison between existing prediction schemes and cluster-based schemes. As depicted, both CSBR and CEG-RAODV protocol performs good and has a high throughput, whereas KPND-AODV emphasizes more on the neighbour discovery process and lacks performance enhancement. More bandwidth is utilized in the discovery process, and a low delivery ratio results in low throughput. DABFS is based on beacon message broadcasting for directional updates and creates broadcast storms with increasing numbers of vehicles hence does not show high incremental growth in throughput.

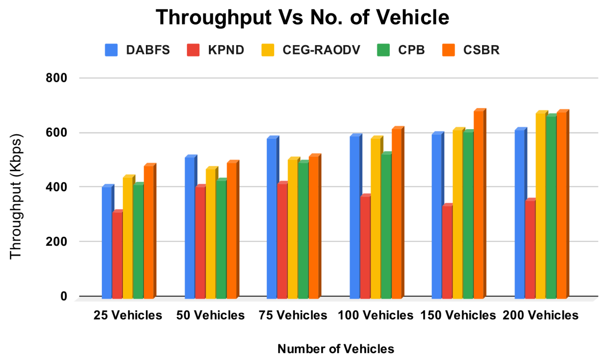

Fig. 8. Throughput

CSBR uses parameters such as connectivity matrix, cosine distance, and a probability function for forwarding packets, which improve the throughput of a network.

The cumulative amount of discovery messages generated throughout the entire simulation is also counted. Fig. 9 represents the influence of vehicle density on the neighbour discovery procedure. It can be analyzed that the message overhead rises with the rise of numbers of vehicles. Under different vehicle density simulation environments, CSBR needs a less number of neighbour discovery packets than other comparative protocols.

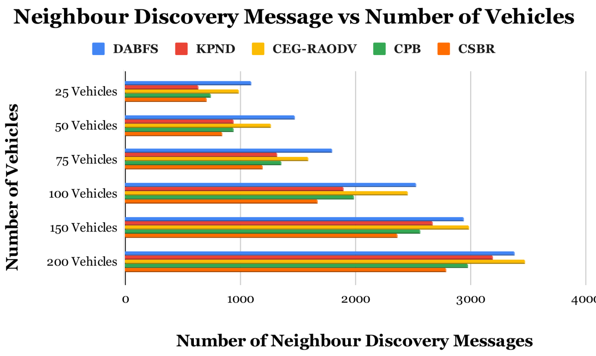

Fig. 9. Neighbour Discovery Message Overhead

Since we know, CSBR and CPB employ a probabilistic approach to forwarding data packets, and the amount of broadcast packets decreases as clusters stabilize in the network. That is why higher dense network performance of CSBR is improvised compared to the existing schemes.

VIII. Conclusion and Future Scope

In this paper, we proposed a cluster-based routing with the selective broadcast approach named CSBR. In this approach, the cluster heads are used to obtain an efficient route to the destination vehicle in a topological scenario. The cluster head is elected for each cluster based on the cosine distance, velocity, and neighbour adjacency. Cosine similarity provides more stable clusters in the network as it is based on the orientation of vehicles rather than just considering the magnitude of a vehicle. A probabilistic function is applied to determine the forwarding or dropping decisions of the data packets. For estimating the performance of the proposed CSBR protocol, we conduct numerous simulations and use distinctive density scenarios of vehicular networks. We observed that in the urban mobility scenario, existing routing protocols suffer from performance degradation. In contrast, CSBR performs more efficiently in given circumstances.

Our research work currently focuses more on the performance parameters of the vehicular networks in urban traffic circumstances. We can continue our research for the additional mobility models, i.e., highway and group mobility models, inter-sectional areas in the city, and complex traffic scenarios. We will also continue our research to enhance the stability of the clustering process with less overhead and self-learning architecture in VANETs routing.

References

[1] Haider, Shahab, Ghulam Abbas, Ziaul Haq Abbas, and Thar Baker. "DABFS: A robust routing protocol for warning messages dissemination in VANETs." *Computer Communications 147* (2019): 21-34.

[2] Liu, Chunfeng, Gang Zhang, Weisi Guo, and Ran He. "Kalman Prediction-Based Neighbor Discovery and Its Effect on Routing Protocol in Vehicular Ad Hoc Networks." *IEEE Transactions on Intelligent Transportation Systems* (2019).

[3] Naderi, Mohammad, Farzad Zargari, and Mohammad Ghanbari. "Adaptive beacon broadcast in opportunistic routing for VANETs." *Ad Hoc Networks 86* (2019): 119-130.

[4] Khan, Zahid, Pingzhi Fan, Sangsha Fang, and Fakhar Abbas. "An Unsupervised Cluster-Based VANET-Oriented Evolving Graph (CVoEG) Model and Associated Reliable Routing Scheme." *IEEE Transactions on Intelligent Transportation Systems* (2019).

[5] Liu, Lei, Chen Chen, Tie Qiu, Mengyuan Zhang, Siyu Li, and Bin Zhou. "A data dissemination scheme based on clustering and probabilistic broadcasting in VANETs." *Vehicular Communications 13* (2018): 78-88.

[6] Haider, Shahab, Ghulam Abbas, Ziaul Haq Abbas, Saadi Boudjit, and Zahid Halim. "P-DACCA: A Probabilistic Direction-Aware Cooperative Collision Avoidance Scheme for VANETs." *Future Generation Computer Systems 103* (2020): 1-17.

[7] Tahooni, Mohammad, Amir Darehshoorzadeh, and Azzedine Boukerche. "Mobility-based opportunistic routing for mobile adhoc networks." *In Proceedings of the 11th ACM symposium on Performance evaluation of wireless ad hoc, sensor and ubiquitous networks*, pp. 9-16. ACM, 2014.

[8] Li, Fan, and Yu Wang. "Routing in vehicular ad hoc networks: A survey." *IEEE Vehicular technology magazine 2*, no. 2 (2007): 12-22.

[9] Han, Seon Yeong, and Dongman Lee. "An adaptive hello messaging scheme for neighbor discovery in on-demand MANET routing protocols." *IEEE communications letters 17*, no. 5 (2013): 1040-1043.

[10] Saini, Trilok Kumar, and Subhash C. Sharma. "Prominent unicast routing protocols for Mobile Ad hoc Networks: Criterion, classification, and key attributes." *Ad Hoc Networks 89* (2019): 58-77.

[11] Tang, Yujie, Nan Cheng, Wen Wu, Miao Wang, Yanpeng Dai, and Xuemin Shen. "Delay-Minimization Routing for Heterogeneous VANETs With Machine Learning Based Mobility Prediction." *IEEE Transactions on Vehicular Technology 68*, no. 4 (2019): 3967-3979.

[12] Zhang, Degan, Ting Zhang, and Xiaohuan Liu. "Novel self-adaptive routing service algorithm for application in VANET." *Applied Intelligence 49*, no. 5 (2019): 1866-1879.

[13] Ghaffari, Ali. "Hybrid opportunistic and position-based routing protocol in vehicular ad hoc networks." *Journal of Ambient Intelligence and Humanized Computing* (2019): 1-11.

[14] Abbasi, Hamza Ijaz, Razvan Cristian Voicu, John Copeland, and Yusun Chang. "Towards Fast and Reliable Multi-hop Routing in VANETs." *IEEE Transactions on Mobile Computing* (2019).

[15] Li, Peng, Lu Guo, and Fang Wang. "A Multipath Routing Protocol with Load Balancing and Energy Constraining Based on AOMDV in Ad Hoc Network." *Mobile Networks and Applications* (2019): 1-10.

[16] Clausen, Thomas, Philippe Jacquet, Cédric Adjih, Anis Laouiti, Pascale Minet, Paul Muhlethaler, Amir Qayyum, and Laurent Viennot. "Optimized link state routing protocol (OLSR)." *inria-00471712* (2003).

[17] Yan, Gongjun, Stephan Olariu, and Michele C. Weigle. "Providing VANET security through active position detection." Computer communications 31, no. 12 (2008): 2883-2897.

[18] Wang, Leilei, Zhigang Chen, and Jia Wu. "An opportunistic routing for data forwarding based on vehicle mobility association in vehicular ad hoc networks." Information 8, no. 4 (2017): 140.

[19] Al-Anazi, Sumayia, Hind AlMahmoud, and Isra Al-Turaiki. "Finding similar documents using different clustering techniques." *Procedia Computer Science 82* (2016): 28-34.

[20] Zhang, Chenxi, Rongxing Lu, Xiaodong Lin, P-H. Ho, and Xuemin Shen. "An efficient identity-based batch verification scheme for vehicular sensor networks." In *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*, pp. 246-250. IEEE, 2008.

[21] Naumov, Valery, and Thomas R. Gross. "Connectivity-aware routing (CAR) in vehicular ad-hoc networks." In IEEE INFOCOM 2007-26th IEEE International Conference on Computer Communications, pp. 1919-1927. IEEE, 2007.

[22] Wang, Xin, Changle Li, Lina Zhu, and Chunchun Zhao. "An effective routing protocol for intermittently connected vehicular ad hoc networks." In 2013 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1750-1755. IEEE, 2013.

[23] Khan, Zahid, and Pingzhi Fan. "A novel triple cluster based routing protocol (TCRP) for VANETs." In 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), pp. 1-5. IEEE, 2016.

[24] Zhang, Kai, Jingjing Wang, Chunxiao Jiang, Tony QS Quek, and Yong Ren. "Content aided clustering and cluster head selection algorithms in vehicular networks." In 2017 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1-6. IEEE, 2017.